JOURNAL OF BACTERIOLOGY, Apr. 2006, p. 3024–3036 Vol. 188, No. 8
0021-9193/06/$08.00+0 doi:10.1128/JB.188.8.3024–3036.2006

Functional Domains of the *Bacillus subtilis* Transcription Factor AraR and Identification of Amino Acids Important for Nucleoprotein Complex Assembly and Effector Binding†

Irina Saraiva Franco,[1] Luís Jaime Mota,[1]‡ Cláudio Manuel Soares,[2] and Isabel de Sá-Nogueira[1,3]*

Laboratory of Microbial Genetics, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Portugal[1]; Laboratory of Protein Modeling, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Portugal[2]; and Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal[3]

Received 28 December 2005/Accepted 8 February 2006

The *Bacillus subtilis* AraR transcription factor represses at least 13 genes required for the extracellular degradation of arabinose-containing polysaccharides, transport of arabinose, arabinose oligomers, xylose, and galactose, intracellular degradation of arabinose oligomers, and further catabolism of this sugar. AraR exhibits a chimeric organization comprising a small N-terminal DNA-binding domain that contains a winged helix-turn-helix motif similar to that seen with the GntR family and a larger C-terminal domain homologous to that of the LacI/GalR family. Here, a model for AraR was derived based on the known crystal structures of the FadR and PurR regulators from *Escherichia coli*. We have used random mutagenesis, deletion, and construction of chimeric LexA-AraR fusion proteins to map the functional domains of AraR required for DNA binding, dimerization, and effector binding. Moreover, predictions for the functional role of specific residues were tested by site-directed mutagenesis. In vivo analysis identified particular amino acids required for dimer assembly, formation of the nucleoprotein complex, and composition of the sugar-binding cleft. This work presents a structural framework for the function of AraR and provides insight into the mechanistic mode of action of this modular repressor.

The transcription factor AraR controls the expression of several *Bacillus subtilis* genes encoding enzymes and permeases involved in the degradation of arabinose-containing polysaccharides, uptake of L-arabinose (and possibly arabinose oligomers), xylose, and galactose, and further intracellular catabolism of arabinose and arabinose oligomers. The arabinose (*ara*) regulon comprises at least 13 genes (Fig. 1), the *araAB DLMNPQ–abfA* operon (43), the divergently arranged *araE*/*araR* genes (42, 45) located in distinct regions of the *B. subtilis* chromosome, and the genes *abnA* and *xsa* positioned immediately upstream and 23 kb downstream from the operon, respectively (57). The first three genes of the arabinose metabolic operon, *araA*, *araB*, and *araD*, encode the enzymes required for the intracellular conversion of L-arabinose into D-xylulose 5-phosphate, which is further catabolized through the pentose phosphate pathway (41). The function of *araL* and *araM* is unknown, and genes *araNPQ* encode components of an ABC-type transporter most likely involved in the uptake of arabinose oligomers (43, 45). The product of the *araE* gene is a permease, the main transporter of arabinose into the cell (45), that is also responsible for the uptake of xylose and galactose (14). The *araR* gene encodes the regulatory protein of the system (27). The last gene of the metabolic operon, *abfA*, and the *xsa* gene most probably encode arabinofuranosidases involved in the intracellular degradation of arabinose oligomers (36). The *abnA* gene encodes an extracellular endo-arabinanase that degrades the arabinose homoglycan arabinan (16).

The pathways of L-arabinose utilization in *B. subtilis* and *Escherichia coli* are identical, and the catabolic enzymes are functionally homologous (7, 18, 41). However, in these two model microorganisms for gram-positive and gram-negative bacteria, respectively, the regulators have distinct phylogenetic origins and well-characterized different modes of action, outlining the divergent evolutionary pathways of catabolic enzymes and regulatory proteins. The *B. subtilis ara* regulon, comprising five arabinose-responsive promoters (*ParaABD LMNPQ-abfA*, *PabnA*, *Pxsa*, *ParaE*, *ParaR*), is under the negative control of AraR (36, 42, 45), whereas in *E. coli* the regulatory protein, AraC, activates four arabinose-responsive promoters (*ParaBAD*, *ParaE*, *ParaFGH*, *ParaJ*) and represses the expression of its own promoter *ParaC* (46, 47). The AraC protein, composed of an N-terminal arabinose-binding and -dimerization domain and a C-terminal DNA-binding domain (DBD), uses a binary switch mechanism of allosteric regulation. In the absence of arabinose, a peptidyl arm from the effector-binding and -dimerization domain binds to the DBD. Consequently, in the dimer the orientation of the DBDs with respect to each other favors their binding to two distant DNA sites causing the formation of a loop. When the effector is present, the arms bind over arabinose, freeing the DBDs, which bind to two adjacent DNA sites (references 46 and 47 and references therein).

The *B. subtilis* AraR protein recognizes and binds to eight specific DNA operator sites within the five different promoters

* Corresponding author. Mailing address: Isabel de Sá-Nogueira, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Av. da República, Apt. 127, 2781-901 Oeiras, Portugal. Phone: 351 21 4469524. Fax: 351 21 4411277. E-mail: sanoguei@itqb.unl.pt.

† Supplemental material for this article may be found at http://jb.asm.org/.

‡ Present address: Imperial College London, Center for Molecular Microbiology and Infection, Armstrong Road, Flowers Building—2nd floor, London SW7 2AZ, United Kingdom.

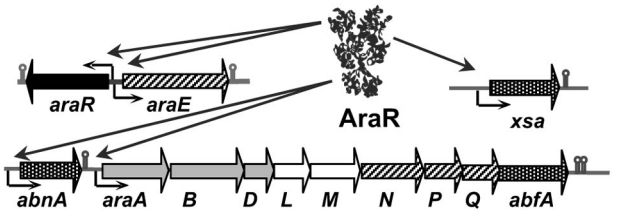


FIG. 1. Schematic representation of the *B. subtilis* arabinose regulon. Expression of all *ara* genes is repressed by AraR and induced by arabinose. Distinct mechanisms of AraR binding to its operator regions in the different promoters allow a tight or flexible control of the system. AbfA, AbnA, and Xsa are enzymes involved in the degradation of arabinose-containing polysaccharides. AraE is a permease, the main transporter of arabinose into the cell, also responsible for the uptake of xylose and galactose. AraNPQ are components of an ABC-type high-affinity transport system for arabinose and/or arabinose oligomers. Intracellular catabolism of arabinose into xylulose 5-P, which is further catabolized through the pentose phosphate pathway, is carried out by AraA, AraB, and AraD. The function of AraL and AraM is unknown.

(*ParaABDLMNPQ-abfA*, *PabnA*, *Pxsa*, *ParaE*, *ParaR*), as determined by DNaseI footprinting studies and mutation analysis (26, 27, 36). However, two distinct modes of transcriptional repression are observed that might reflect different physiological requirements. A high level of repression is achieved by cooperative binding of AraR to two in-phase operators concomitant with DNA looping within the promoter region of the *ara* metabolic operon, and the genes *xsa* and *araE*, that ensures tight control of expression of intracellular enzymes and transport systems. On the other hand, binding to a single operator that autoregulates *araR* expression and represses *abnA* expression is less effective, resulting in flexible control that allows basal transcription of both genes (26, 27, 36). Binding of arabinose to AraR is presumed to result in a conformational change that inhibits or prevents AraR-DNA binding, leading to transcriptional derepression of the regulon. AraR is a 362-residue and 40-kDa protein that exhibits a mosaic organization comprising two domains of different origins (27, 42). The N-terminal region (residues 1 to 70) contains a helix-turn-helix (HTH) consensus signature sequence of the GntR family of bacterial regulatory proteins (10), while the C-terminal region (residues 71 to 362) displays significant similarity to that of the LacI/GalR family of bacterial regulators (56). This modular architecture, the predicted HTH motif, and the amino acid similarities suggest that AraR is composed of two functional domains: a smaller N-terminal domain responsible for the DNA-binding specificity and a larger C-terminal domain involved in effector binding and oligomerization. To date, in addition to AraR, 21 proteins presenting this mosaic modular association in different bacteria have been sequenced. Together they constitute one of the six GntR subfamilies typified by AraR (17, 38). However, no functional studies have been reported for these AraR-like proteins (http://www.sanger.ac.uk/Software/Pfam). Here, we combined random mutagenesis of the *araR* gene and site-directed mutagenesis guided by molecular modeling techniques to identify the residues within AraR that are involved in DNA binding, dimer assembly, and arabinose binding. The functional relevance of the identified amino acids was evaluated by genetic analysis. This structure-function analysis of AraR provides insights into the roles of amino acid residues involved in AraR nucleoprotein complex assembly and effector binding.

MATERIALS AND METHODS

Strains and growth conditions. *B. subtilis* strains used in this study (Table 1) were grown in liquid Luria-Bertani (LB) medium (25) or C minimal medium (32) and solid sugar-free agar (SFA) medium (LabM) or LB broth solidified with 1.6% agar. Chloramphenicol (5 μg ml^{-1}), kanamycin (10 μg ml^{-1}), erythromycin (1 μg ml^{-1}), and neomycin (1 μg ml^{-1}) were added when appropriate. The Amy phenotype was tested on tryptose blood agar base medium (Difco) plates containing 1% of potato starch after overnight incubation, and starch hydrolysis was detected by flooding the plates with an I_2-KI solution as described previously (42). *Escherichia coli* DH5α (Gibco BRL) was used for routine molecular cloning work, and *E. coli* SU101 (5) was used as the host for the analysis of chimeric LexA-AraR fusions. *E. coli* strains were grown on LB broth and LB broth solidified with 1.6% agar or MacConkey agar medium (Difco) plates, with ampicillin (100 μg ml^{-1}), chloramphenicol (20 μg ml^{-1}), kanamycin (30 μg ml^{-1}), tetracycline (12 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 (26).

DNA manipulations and sequencing. DNA manipulations were carried out as described previously (40). 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 MERmaid kit (Bio101). PCRs were performed in a GeneAmp PCR system 2400 (Perkin-Elmer), and PCR products were purified using a QIAquick PCR purification kit (QIAGEN). DNA was sequenced using a Sequenase V 2.0 kit (USB) or an ABI PRISM BigDye terminator ready-reaction cycle sequencing kit (Applied Biosystems).

Random and site-directed mutagenesis of *araR*. Plasmid pLS30 was used as a DNA template in random mutagenesis experiments and was constructed as follows: a 2,920-bp PvuII-PvuII region from pSN32 (27) containing the *lacZ* gene was removed by restriction and autoligation, producing pLS21. The unique BglII site present in this plasmid was eliminated by fill-in to yield pLS22. Insertion of a 1,446-bp EcoRI-BamHI fragment from pLM3 (43), containing an entire copy of the *araR* gene and a truncated copy of *araE*, into pLS22 EcoRI-BamHI yielded pLS24. Finally, the creation of two additional restriction sites (BglII or KpnI) in *araR* by silent mutagenesis was accomplished by PCR amplification using pLS24 as a template and the pairs of primers ARA200-ARA203 and ARA201-ARA202 (Table 2) in two separate experiments; the resulting 187-bp and 8.1-kb PCR products, respectively, were digested with BglII or KpnI and ligated to yield pLS30, at which point the presence of both restriction sites and the absence of further mutations were checked by sequencing.

To facilitate cloning procedures and sequence analysis, the *araR* allele was divided into three fragments flanked by unique restriction sites (162-bp BglII-KpnI, 324-bp KpnI-MluI, and 664-bp MluI-EcoRI); random mutagenesis of *araR* was accomplished by PCR amplification of each fragment with flanking primers ARA200-ARA203, ARA202-ARA10, and ARA95-ARA1 (Table 2), respectively, under conditions that increased the frequency of *Taq* polymerase error to 0.4% (19). The PCR contained 1× PCR buffer (MBI Fermentas), 0.5 mM primers, 0.1 to 1 mg ml^{-1} pLS30 DNA, 1 mM deoxynucleoside triphosphates, 6.6 mM $MgCl_2$, 0.5 mM $MnCl_2$, and 1 μl of *Taq* polymerase (MBI Fermentas) in a 100-μl total volume. PCR products were cleaved with the appropriate restriction enzymes, and the resulting fragment was cloned back into pLS30 to replace the equivalent region of the wild-type *araR* allele. The recombinant plasmids were transformed into *E. coli* DH5α competent cells, yielding three distinct libraries of *araR* mutations. Ten isolates from each group of approximately 800 colonies were picked randomly to evaluate the efficiency of mutagenesis by sequencing the mutagenized insert. The determined value of 0.13% is lower than the 2% found by Leung et al. (19). The remaining *E. coli* colonies of each library were pooled, and linearized plasmid DNA was used in separate experiments to transform *B. subtilis* IQB350 (Table 1). This procedure allowed the integration of the *araR* mutant alleles into the chromosome of *B. subtilis*, at the *amyE* locus, via double recombination. Phenotypic analysis of constitutive AraR$^-$ mutants and AraR^s superrepressor mutants was accomplished, respectively, through screening for Lac$^+$ colonies in SFA plates with X-Gal (5-bromo-4-chloro-3-indolyl-β-D-thiogalactopyranoside) (0.02%) and for Lac$^-$ colonies in the same medium with 0.2% L-arabinose. Chromosomal DNA from colonies displaying the expected phenotype was used as a template to amplify the mutagenized region of the *araR* allele, which was subsequently cloned back into pLS30 as described

TABLE 1. *B. subtilis* strains used in this work

Strain[a]	Genotype	Relevant phenotype	Source
IGCg701	*metB10 lys3 araR* E142K	Ara^-	Paveia and Archer (33)
IGCg711	*metB10 lys3 araR* G215V	$Ara^{+/-}$	Paveia and Archer (33)
IGCg715	*metB10 lys3 araR* G215D	$Ara^{+/-}$	Paveia and Archer (33)
IGCg735	*metB10 lys3 araR* G138E	Ara(Con)[b]	Sá-Nogueira et al. (44)
IQB101	*araAB′-lacZ erm*	Ara^- $LacZ^-$	Sá-Nogueira et al. (43)
IQB350	Δ*araR*::*kan araAB′-lacZ erm*	Ara^- $LacZ^+$	pLM8→IQB101[c]
IQB351	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR cat*	$LacZ^-$	pLS24→IQB350
IQB352	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR cat*	$LacZ^-$	pLS30→IQB350
IQB355	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* F37S *cat*	$LacZ^+$	pIF1→IQB350
IQB356	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* Q61R *cat*	$LacZ^+$	pIF2→IQB350
IQB357	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* L33S *cat*	$LacZ^+$	pIF3→ IQB350
IQB358	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* Δ13–65 *cat*	$LacZ^+$	pIF8→IQB350
IQB364	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* T87I *cat*	$LacZ^-$	pIF13→IQB350
IQB368	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* S53P *cat*	$LacZ^+$	pIF17→IQB350
IQB371	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* H226R *cat*	$LacZ^+$	pIF27→IQB350
IQB372	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* I308T *cat*	$LacZ^+$	pIF28→IQB350
IQB373	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* P319L *cat*	$LacZ^+$	pIF29→IQB350
IQB374	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* C271R *cat*	$LacZ^+$	pIF30→IQB350
IQB375	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* R285K *cat*	$LacZ^+$	pIF31→IQB350
IQB376	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* I82T *cat*	$LacZ^+$	pRS1→IQB350
IQB377	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* L157R *cat*	$LacZ^+$	pRS2→IQB350
IQB378	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* S146R *cat*	$LacZ^-$	pRS3→IQB350
IQB379	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* T117A *cat*	$LacZ^-$	pRS4→IQB350
IQB398	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* E142K *cat*	$LacZ^-$	pIF32→IQB350
IQB399	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* G215V *cat*	$LacZ^-$	pIF33→IQB350
IQB500	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* G215D *cat*	$LacZ^-$	pIF34→IQB350
IQB501	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* D211G *cat*	$LacZ^-$	pIF35→IQB350
IQB502	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* G138E *cat*	$LacZ^+$	pIF36→IQB350
IQB510	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* F94A *cat*	$LacZ^-$	pSC9→IQB350
IQB511	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* D212A *cat*	$LacZ^-$	pSC10→IQB350
IQB512	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* Q214A *cat*	$LacZ^-$	pSC11→IQB350
IQB516	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* R218L *cat*	$LacZ^-$	pIF46→IQB350
IQB517	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* D301N *cat*	$LacZ^-$	pIF47→IQB350
IQB518	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* I89A *cat*	$LacZ^+$	pVAB1→IQB350
IQB519	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* L114A *cat*	$LacZ^+$	pVAB2→IQB350
IQB520	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* R99A *cat*	$LacZ^-$	pVAB3→IQB350
IQB528	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* Y92F *cat*	$LacZ^-$	pVAB5→IQB350
IQB529	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* F305A *cat*	$LacZ^+$	pVAB6→IQB350
IQB563	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*Δ*araR* R45A *cat*	$LacZ^+$	pLS30(mut)[d]→IQB350
IQB569	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* E102Q *cat*	$LacZ^+$	pIF68→IQB350
IQB570	Δ*araR*::*kan araAB′-lacZ erm* Δ*amyE*::*araR* H318A *cat*	$LacZ^-$	pLS30(mut)[d]→IQB350
IQB579	*araAB′-lacZ ery* Δ*amyE*::*Pspac Ppcn-lacI neo*	$LacZ^-$	pGR63→IQB101
IQB580	*araAB′-lacZ ery* Δ*amyE*::*Pspac-araR Ppcn-lacI neo*	$LacZ^-$	pIF81→IQB101
IQB581	*araAB′-lacZ ery* Δ*amyE*::*Pspac-araR* R45A *Ppcn-lacI neo*	$LacZ^+$	pIF82→IQB101
IQB584	*araAB′-lacZ ery* Δ*amyE*::*Pspac-araR* F94A *Ppcn-lacI neo*	$LacZ^-$	pIF86→IQB101
IQB585	*araAB′-lacZ ery* Δ*amyE*::*Pspac-araR* L114A *Ppcn-lacI neo*	$LacZ^-$	pIF87→IQB101
IQB591	*araAB′-lacZ ery* Δ*amyE*::*Pspac-araR* I89A *Ppcn-lacI neo*	$LacZ^-$	pIF103→IQB101
IQB592	*araAB′-lacZ ery* Δ*amyE*::*Pspac-araR* F305A *Ppcn-lacI neo*	$LacZ^-$	pIF105→IQB101

[a] All strains derive from *B. subtilis* $168T^+$ (prototroph [43]) except IGCg701, IGCg711, IGCg715, and IGCg735, which derive from BR151 (43).
[b] Con, constitutive.
[c] The arrows indicate transformation and point from donor DNA to recipient strain. Transformation was always carried out with linearized DNA.
[d] Mutagenized (mut) pLS30 DNA was used as the donor DNA (see Materials and Methods).

above and sequenced. The resulting plasmids bear the following mutations leading to single-amino-acid substitutions F37S, Q61R, L33S, T87I, S53P, H226R, I308T, P319L, C271R, R285K, I82T, L157R, S146R, and T117A (see Table 1).

In a previous work a collection of *B. subtilis* strains carrying mutations that mapped on the *araR* region was characterized (33, 44). Chromosomal DNA from these strains was amplified by PCR, using oligonucleotides ARA13 and ARA18, and sequenced. Four *araR* alleles containing single missense mutations (E142K, G215V, G215D, and G138E) were subsequently amplified by PCR with primers ARA202 and ARA10 for E142K and G138E or primers ARA15 and ARA93 for G215V and G215D. The resulting KpnI-MluI or MluI-EcoRI fragments were subcloned into pLS30 (see Table 1).

To create an in-frame deletion in the 5′ region of *araR* the KpnI site in pLS30 was changed for a second BglII site by amplification of the *araR* allele by use of oligonucleotides ARA10 and ARA94 (Table 2). The PCR product was digested with MluI and BglII and cloned back into pLS30 MluI-BglII to generate pIF8, which carries an in-frame deletion comprising amino acids (aa) 13 to 65.

Amino acid substitutions in AraR were made by the QuikChange (Stratagene) site-directed method using plasmid pIF38 as the template. This plasmid is a pLS30 derivative obtained by digestion with EcoRI and Eco47III to remove a 3.4-kb fragment that includes the *cat* gene and *amyE* back sequences. Mutagenic oligonucleotides used (see Table 2) carried the modified codon in the center. The regions containing the target nucleotides (296-bp BamHI-BglII, 486-bp BglII-MluI, or 1,001-bp MluI-PvuII fragments) were subcloned into pLS30 and the mutations confirmed by sequencing. The linearized plasmids were used to transform *B. subtilis* receptor strain IQB350 (Table 1), and the Lac phenotype was checked as described above. Substitutions R45A and H318A were generated by a similar method using pLS30 as the template. After ligation and linearization, the mutagenized plasmids were also used to transform *B. subtilis* strain IQB350.

TABLE 2. Oligonucleotides used in this work

Oligonucleotide	Site-directed substitution[a]	Sequence (5′→3′)[b]
ARA1		(−39)TAAGGGTAACTATTGCCG(−22)
ARA4		(−184)TTCTTCATTTCCCTGCCCTCCCG(−162)
ARA10		(+622)GAGAAAGCAAATGCTCCGC(+604)
ARA13		(−210)CATTTGGTTCTAATTGAGTTGG(−189)
ARA15		(+430)CATATTGACGGACTCATCG(+448)
ARA18		(+1133)CCCAAAGCTTGCTGAATTTATTCATTCAGTTTTCGTGC(+109)6
ARA93		(+1197)GACAGAATTCGTTCGTTG(+1214)
ARA94		(+215)GAGAGATCTTTGTCGCTTCAC(+235)
ARA95		(+460)AAAAGCGCCCTTCAAACC(+477)
ARA96		(+17)GAGGACTCGAGATGTTACC(+35)
ARA98		(+1127)CATTGCTGCAGTTATTCATTCAG(+1105)
ARA142		(+247)TCAGCGCTCGAGTCCAATAAAACGATC(+273)
ARA143		(+502)TTGGAGCTCAACGGCATTCCTTTTGCG(+528)
ARA144		(+796)ACACTCGAGAAAAACAGCAAGCACATGCC(+824)
ARA155	R218L (fwd)	(+661)GACACACAAGGCGTGAAACTGATGAACGG(+689)
ARA156	R218L (rev)	(+687)GTTCATCAGTTTCACGCCTTGTGTGTCATCAG(+656)
ARA157	D301N (fwd)	(+921)CGGGTACAATGATTCACATTTCGC(+944)
ARA158	D301N (rev)	(+944)GCGAAATGTGAATCATTGTACCCG(+921)
ARA159	F94A (fwd)	(+298)GACTATATTGCCCCGAGCATCATC(+321)
ARA160	F94A (rev)	(+318)GATGATGCTCGGGGCAATATAGTC(+298)
ARA161	D212A (fwd)	(+655)GCTGATGCAACACAAGGCGTGAAACG(+680)
ARA162	D212A (rev)	(+680)CGTTTCACGCCTTGTGTTGCATCAGC(+655)
ARA163	Q214A (fwd)	(+658)GATGACACAGCAGGCGTGAAACG(+680)
ARA164	Q214A (rev)	(+680)CGTTTCACGCCTGCTGTGTCATC(+658)
ARA169	L114A (fwd)	(+359)CTATGCTTGCGACAAGCACAAACAAC(+384)
ARA170	L114A (rev)	(+375)GCTTGTCGCAAGCATAGAATACCC(+352)
ARA171	I89A (fwd)	(+284)CAACTTACGCATCAGACTATATTTTCCC(+311)
ARA172	I89A (rev)	(+310)GGAAAATATAGTCTGATGCGTAAGTTGTC(+282)
ARA173	R99A (fwd)	(+314)GCATCATCGCAGGAATCGAGTCC(+336)
ARA174	R99A (rev)	(+334)ACTCGATTCCTGCGATGATGCTC(+312)
ARA175	Y92F (fwd)	(+291)CATATCAGACTTTATTTTCCCGAGC(+315)
ARA176	Y92F (rev)	(+315)GCTCGGGAAAATAAAGTCTGATATG(+291)
ARA177	F305A (fwd)	(+928)GATGATTCACATGCCGCCCAAATC(+951)
ARA178	F305A (rev)	(+951)GATTTGGGCGGCATGTGAATCATC(+928)
ARA181	H318A (fwd)	(+968)CCTCTGTCAAAGCTCCGAAATCAGTGC(+994)
ARA182	H318A (rev)	(+992)ACTGATTTCGGAGCTTTGACAGAGG(+968)
ARA183		AAATCTAGAATTTTGGAGGAATGGATATGTTACC(+35)
ARA200		(+49)GTAAAAGAAGAGATCTCGTCTTGGATTAATCAAGG(+83)
ARA201		(+72)CCAAGACGAGATCTCTTCTTTTACTTGCGCG(+42)
ARA202		(+214)GGAGGTACCTTTGTCGCTTC(+233)
ARA203		(+235)GTGAAGCGACAAAGGTACCTCCG(+213)
ARA242	R45A (fwd)	(+148)CGGCATACCATCGCGAAAGCGATCGGAGAC(+177)
ARA243	R45A (rev)	(+177)GTCTCCGATCGCTTTCGCGATGGTATGCCG(+148)

[a] Amino acid substitution obtained and direction of amplification of oligonucleotide, forward (fwd) or reverse (rev).
[b] The number(s) within the primer sequence refers to the position of the sequence relative to the transcription start site of *araR*, except in ARA1, where it indicates the position in pLS30 relative to the EcoRI site (+1). Restriction sites or modified codons used in site-directed mutagenesis are underlined.

The presence of the mutation was verified by sequencing the *araR* allele in the resulting strains by PCR amplification of chromosomal DNA with oligonucleotides ARA1 and ARA4.

Construction of chimeric LexA-AraR. The *araR* coding sequence comprised in pLS30 was amplified with oligonucleotides ARA96 and ARA98, containing XhoI and PstI sites, respectively. After cleavage with these enzymes, the resulting 1-kb PCR fragment was inserted into pSR658 XhoI-PstI (5) to yield pSC1, which carries an in-frame fusion of the entire AraR with LexA-DBD (LexA-DBD $AraR_{1–362}$). To construct chimeras containing truncated versions of AraR, *araR*-specific oligonucleotides ARA142 to ARA144 (Table 2) were synthesized to engineer XhoI or SacI sites and used in amplifications with pSC1 as the template. Insertion of the PCR products into pSR658, digested with the appropriate enzymes, generated plasmids pSC3 (LexA-DBD $AraR_{76–362}$), pSC4 (LexA-DBD $AraR_{160–362}$), and pSC5 (LexA-DBD $AraR_{258–362}$). Construction of the pSR658 derivatives containing *araR* alleles harboring the single mutation L114A or the double mutation I89AL114A was accomplished as follows. Elimination of the StyI site from pSR658 by fill-in resulted in pVAB7, where the wild-type allele was subsequently cloned as described above for pSC1, creating pVAB8. Substitution of the 514-bp StyI-BglII region of *araR* in pVAB8 for the corresponding region of the alleles contained in pVAB1 or pVAB2 yielded, respectively, pIF109 or pVAB10, which carry mutated allele I89A or L114A. To obtain the double mutant I89AL114A, the 286-bp AvaI-AvaI fragment from pVAB10 was exchanged for the same region in pIF109, yielding pIF122. All the new constructs bearing the missense mutations and in-frame fusions were confirmed by sequencing. The Lac phenotype of the *E. coli* strains harboring these plasmids was checked after growth in MacConkey agar plates (Difco) with the appropriate antibiotics.

Construction of *Pspac-araR* fusions for *B. subtilis* transdominance assays. The wild-type *araR* allele was amplified by PCR using pLS30 as the template and oligonucleotides ARA183 and ARA95, which create XbaI and EcoRI sites at the 5′ and 3′ end of the gene. The DNA fragment was digested with these enzymes and inserted into the same sites of pBKSII(−), generating pIF79. DNA fragments from *araR* alleles containing the mutation R45A, I89A, L114A, or F94A were amplified using chromosomal DNA from strain IQB563, pVAB1, pVAB2, or pSC9, respectively, with primers ARA183 and ARA10, digested with XbaI and MluI, and subcloned into pIF79. For mutation F305A, primers ARA93 and ARA95 were used and the fragment MluI-EcoRI was subcloned into pIF79. All these alleles were placed under the control of the *Pspac* promoter (11) by digestion of the pIF79 derivatives with XbaI and ClaI and insertion of the fragments into pGR63 XbaI-ClaI (37). All mutations were confirmed by DNA

sequencing, and linearized plasmids were used in separate experiments to transfer the *Pspac-araR* fusions into the chromosome of *B. subtilis* IQB101 (Table 1), at the *amyE* locus, via double recombination.

β-Galactosidase assays. *B. subtilis* strains were grown in C minimal medium supplemented with 1% casein hydrolysate in the presence and absence of 0.4% (wt/vol) L-arabinose as previously reported (43). For the transdominance assays of mutant *araR* alleles, induction of the *Pspac* promoter was accomplished with the addition of 0.5 mM IPTG. Samples of cell culture were collected and analyzed 2 h after the addition of L-arabinose. β-Galactosidase activity was measured and expressed in Miller units as described previously (43). The ratio of the levels of β-galactosidase activity determined in the presence and absence of inducer was taken as a measure of AraR repression in the analyzed strains (repression factor). Growth of *E. coli* strains for quantification of repression was carried out as described previously (5), and samples of cell culture were collected and analyzed 2 h after they reached an optical density at 600 nm of ~0.2.

Immunoblotting of cell extracts. *B. subtilis* strains were grown as described for the β-galactosidase assays (see above), and 8 ml of cell culture was harvested 2 h after induction. After resuspension in lysis buffer (500 mM KCl, 20 mM HEPES-K^+, 10 mM EDTA, 1 mM dithioerythritol, 10% glycerol, 1 mg ml^{-1} lysozyme) and incubation at 37°C for 10 min, cells were subjected to 3 cycles of freezing in liquid N_2 and thawing at 37°C. Benzonase (Merck) (5 U) was added, and incubation was continued for 10 min. After 45 min of centrifugation at 13,000 rpm, samples of the soluble fraction containing 20 μg of protein were resolved on 12.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis gels. Gels were transferred for 1 h at 100 V to nitrocellulose membranes, and blots were developed with anti-AraR-MBP2* serum (27) by use of an ECL detection system (Amersham Biosciences) as described by the manufacturer. Protein concentrations were determined using a Bio-Rad kit.

Multiple sequence alignment. We used the ClustalW 1.82 program (53) with the default parameters to perform multiple alignments of amino acid sequences of AraR and nine AraR-like proteins. Similarly, the sequences of N-terminal regions of these 10 AraR-like proteins and 20 other members of the GntR family (Pfam PF00392) or the C-terminal regions of the AraR-like proteins and 20 other members of periplasmic binding proteins and sugar binding domain of the LacI/GalR family (Pfam PF00532 [Peripla_BP_1]) were aligned to score amino acid similarities. Identification of homologous residues was carried out as described by Kraus et al. (13). Residues were considered characteristic for the GntR family or the Peripla_BP_1 family when at least 75% of the 30 entries shared identical residues in one position. Amino acids that are typical of the AraR-like proteins are identical in 75% of the 10 AraRs and are found in less than 25% of the other members of the GntR family or the Peripla_BP_1 family.

Modeling AraR. Since no significant hits for the whole AraR sequence were found when searching the structural databases for sequence homologues of the complete AraR, the N- and C-terminal domains cannot be modeled together. However, the situation is different when one tries to find homologues for the two domains separately. For the N-terminal domain of AraR we found that the N-terminal domain of the structurally characterized FadR (54), a fatty acid-responsive transcription factor from *E. coli*, displays some homology (28% identity). For this protein there is also a structure bound to DNA (58) which is quite helpful in characterizing potentially important residues for the DNA interaction on AraR. Therefore, this part of the structure of FadR (PDB code 1HW1; without DNA) was used to derive the structure of the N-terminal domain of AraR (the first 68 residues). The C-terminal domain of AraR shows some homology (28% identity) with the structurally characterized C-terminal domain of the purine repressor PurR (29, 49, 50), which was also determined in the DNA-bound and -free forms. In this case, in contrast to the results seen with the N-terminal domain of FadR, the DNA-bound and DNA-free conformations are substantially different, even in the C-terminal domain (which does not contact DNA). Therefore, we considered use of the DNA-bound form (PDB code 1QPZ) to model the C-terminal part of AraR to be safer. The program MOD ELLER (39), version 6.1, was used for deriving the structures of both domains of AraR (examining the dimers). The initial alignments were optimized through several modeling cycles until a good quality model for the unknown structure was achieved. The quality was assessed by examining the restraint violations reported by MODELLER (39) and a Ramachandran analysis performed by the program PROCHECK (15). The final model of the N-terminal domain has 94.6% of the residues in the most-favored regions and 5.4% in additional allowed regions. There were no residues in generously allowed or in disallowed regions. The final model of the C-terminal domain has 91.1% of the residues in the most-favored regions, 7.6% in additional allowed regions, and 0.8% in generously allowed and 0.6% (3 residues) in disallowed regions. Some of the residues presenting problems already had them in the template structure. In the modeled C-terminal domain, given that the binding of arabinose depends heavily on side chain conformations, we decided to use a better method to predict them with more accuracy. For that method we used side chain prediction methods developed previously (22, 24), tailored specifically for structures obtained using comparative modeling (and therefore containing some backbone conformation errors) (23). These methods use a backbone-dependent rotamer library, and in this case we used the "Direct" strategy (see reference 23 for details). These methods were applied to the C-terminal dimer. The resulting conformation was then energy minimized using GROMOS96 (51, 55), keeping the backbone fixed to produce the final model that was going to be used in the analysis.

Arabinose was built in the active site by use of the structure of the ribose-binding protein (PDB code 2DRI) (3) that shows some homology with AraR and contains ribose in the binding site. The arabinose was optimized on the binding site of AraR after some manual adjustments and energy minimization.

RESULTS

AraR-like proteins and AraR modeling. AraR exhibits significant similarity to the LacI/GalR family of bacterial regulators (56). The similarity does not extend to the N-terminal region, which is related to the HTH consensus signature sequence of the GntR family of bacterial regulatory proteins (10). Nine orthologs displaying more than 40% amino acid identity with *B. subtilis* AraR occur in eight other gram-positive bacteria. To obtain a framework for the evaluation of AraR functional analysis, we aligned the sequences of these 10 AraR-like proteins and 20 other members of the GntR family (Pfam PF00392 [N-terminal region]) or 20 members of periplasmic binding proteins and sugar-binding domain of the LacI/GalR family (Pfam PF00532 [Peripla_BP_1]) and scored amino acid similarities (see Materials and Methods). The results are summarized in Fig. 2, where only three members of the non-AraR sequence group are shown. Amino acids that are typical of the AraR-like proteins are identical in 75% of the 10 AraRs and are found in less than 25% of the other members of the GntR family or the Peripla_BP_1 family. The identical residues in AraR-like proteins were found scattered throughout the entire sequence, although some clusters are located in the C-terminal domain of the proteins (positions 90 to 218 in AraR in *B. subtilis*) (Fig. 2).

To obtain a more detailed structure-function correlation, a model for the AraR dimer was built that allowed the identification of amino acids potentially involved in effector binding, dimerization, and DNA binding (Fig. 3A, B, and C) (coordinates of the models are available; see the supplemental material). However, due to the chimeric organization of AraR the N- and C-terminal domains cannot be modeled together. The N-terminal domain of the structurally characterized FadR (54), a fatty acid-responsive transcription factor from *E. coli*, was used to derive the structure (the first 68 residues) of the N-terminal domain of AraR. The C-terminal domain of the *E. coli* purine repressor PurR (29, 49, 50) was used to model the C-terminal part of AraR (see Materials and Methods). The effector-binding domain of the members from the LacI/GalR family is structurally analogous to those of periplasmic-binding proteins despite the generally low level of sequence identity between these proteins (9, 15, 20). The definition of the arabinose-binding pocket in AraR was based on the structure of D-ribose-binding protein from *E. coli* (28).

LexA-AraR chimeras identify the dimerization domain of AraR. We previously established that AraR is indeed a DNA-binding protein (27) and that arabinose is the specific effector that modulates its DNA-binding activity and its ability to

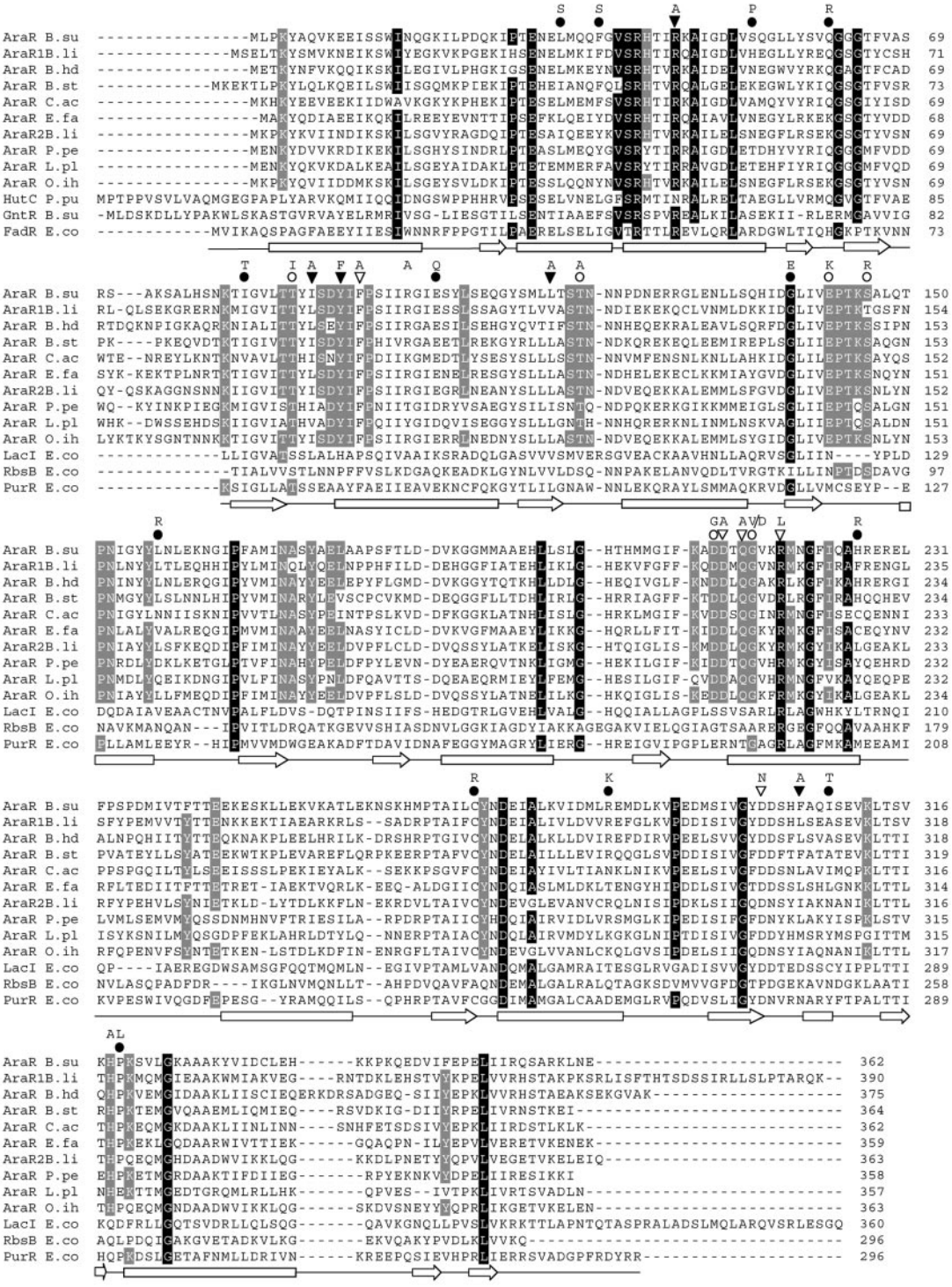


FIG. 2. Sequence alignment of AraR-like proteins and three members of the GntR family (N-terminal region) or two members of LacI/GalR family (C-terminal region) and one periplasmic binding protein. Residues that are typical of the entire GntR or LacI/GalR family are depicted with gray characters on a black background; residues characteristic of AraR homologous proteins are highlighted with white characters on a gray background (see Materials and Methods). Positions of AraR mutations obtained by random mutagenesis (circles) or site-directed mutagenesis (triangles) are shown. Black triangles or circles represent mutations leading to a constitutive phenotype, and open triangles or circles denote changes that resulted in an AraR superrepressor phenotype. Letters representing the introduced residues are shown above the symbols. The secondary structures (arrows represent beta strands; bars represent alpha-helices) of FadR (amino acid residues 1 to 73) and PurR (positions 60 to 296) are shown below the alignment according to van Aalten et al. (54) and Schumacher et al. (50), respectively. The microorganisms of source and accession numbers are as follows: B. su, *B. subtilis* (P96711); B. li, *B. licheniformis* (Q62R80 and Q62UH0); B. hd, *B. halodurans* (Q9KBQ0); B. st, *Geobacillus stearothermophilus* (Q9S470); C. ac, *Clostridium acetobutylicum* (Q97JE6); E. fa, *Enterococcus faecium* (gi48825728); P. pe, *Pediococcus pentosaceus* (gi48870639); L. pl, *Lactobacillus plantarum* (Q88S80); O. ih, *Oceanobacillus iheyensis* (Q8EMP1); HutC P. pu, histidine utilization repressor of *Pseudomonas putida* (P22773); GntR, B. su gluconate utilization repressor of *B. subtilis* (P10585); FadR E. co, fatty-acid metabolism regulator of *E. coli* (P09371); LacI E. co, lactose repressor of *E. coli* (P03023); RbsB E. co, ribose-binding protein of *E. coli* (P02925); PurR E. co, purine repressor of *E. coli* (P15030).

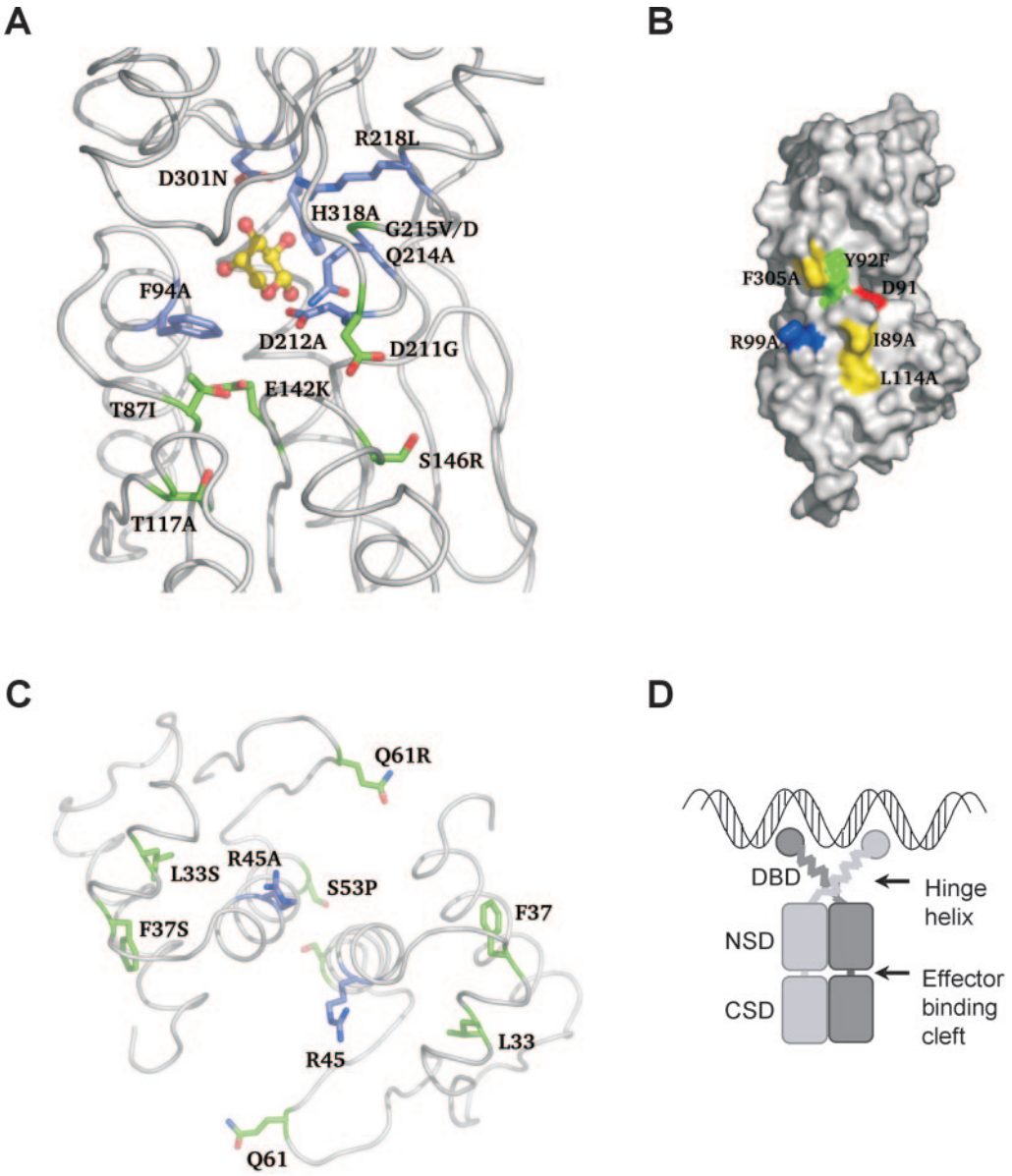
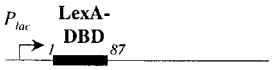
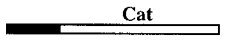
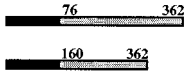
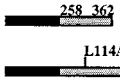
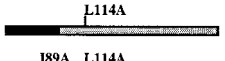
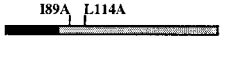


FIG. 3. AraR modeling. A) Close-in view of the arabinose binding site on the C-terminal domain of AraR, highlighting in blue shades the residues mutated by site-directed mutagenesis using the model and in green shades the residues found by random mutagenesis. Arabinose is represented by balls and sticks, and the fold of the protein is represented by a tube. B) Colored molecular surface of the C-terminal domain of AraR highlighting the residues potentially involved in the dimer contact and the mutations performed on them. The dimer contact is made up of hydrophobic residues, surrounded by polar and charged residues. C) Close-in view of the N-terminal region of AraR, highlighting in blue shades the residues mutated by site directed mutagenesis and in green shades the residues found by random mutagenesis. The fold of the protein is represented by a tube. The structure is oriented so that the part interacting with the DNA (using the similarity with FadR) (29) is facing the viewer. D) Schematic drawing of the domain structure of LacI/GalR repressors adapted from Lewis et al. (20). The DBDs, the N-terminal subdomains (NSD) and the C-terminal subdomains (CSD), the effector binding cleft, and the hinge helix, which is formed upon operator binding and connects the DBD to the core of the repressor, are indicated. Panels A, B and C were rendered with PyMol (6).

repress transcription from the *ara* promoters (26, 27). AraR is presumed to act as a dimer due to the dyad symmetry of its target sequences and the existence of a potential HTH motif, both features commonly observed in bacterial transcription repressors that bind DNA as homodimers or homotetramers (30). Gel filtration and glutaraldehyde cross-linking experiments carried out previously suggested that AraR is able to multimerize in solution, but the results were not conclusive (27; L. J. Mota and I. de Sá-Nogueira, unpublished data). To address this issue and to further distinguish functional domains of AraR, chimeric fusions between the amino-terminal DBD of LexA from *E. coli* and different portions of AraR were constructed and analyzed. LexA is a repressor that binds to an operator site in the promoter of *sulA* gene and represses transcription. Binding to the DNA is dependent upon dimerization of the C-terminal domain, which can be removed and replaced

TABLE 3. Repression of *sulA-lacZ* by LexA-AraR chimeras in *E. coli*[a]

Plasmid-encoded LexA-AraR chimeras	β-Galactosidase activity
	2252.0 ± 31.0
	23.8 ± 4.1
	6.3 ± 1.6
	5.3 ± 0.5
	967.0 ± 68.0
	96.9 ± 7.3
	8.9 ± 0.6
	95.7 ± 9.3

[a] Rectangles in black represent the LexA DNA-binding domain (LexA-DBD), in white the chloramphenicol acetyl-transferase (Cat), and in grey the different regions of AraR (numbers indicate amino acid residues). Full-length AraR chimeras harboring single and double mutations (L114A and I89AL114A), obtained by site-directed mutagenesis, are indicated. *E. coli* strains carrying the correspondent plasmid-encoded chimeras were grown in LB and assayed for β-galactosidase activity (as described in Materials and Methods). Values are the average and standard deviations of three independent experiments, each assayed in duplicate.

with another protein or protein fragment (5). When homodimerization of the fused moiety occurs it allows the chimeric LexA to repress expression from the *sulA* promoter. In this study, the DBD of LexA (residues 1 to 87) was fused in-frame to the full-length AraR protein ($AraR_{1–362}$) and to three truncated forms of the repressor ($AraR_{76–362}$, $AraR_{160–362}$, and $AraR_{258–362}$). The four LexA-AraR chimeras were independently used to transform the *E. coli* reporter strain SU101 harboring a *sulA-lacZ* fusion. Chloramphenicol acetyltransferase, a characterized homotrimer (5), fused to the LexA DBD was used as a positive control in these experiments. The $AraR_{1–362}$ and $AraR_{76–362}$ chimeras were able to repress efficiently the expression of the *sulA-lacZ* fusion, whereas $AraR_{160–362}$ and $AraR_{258–362}$ had an intermediate effect (Table 3). The differences observed in the levels of repression between $AraR_{160–362}$ and $AraR_{258–362}$ could be due to misfolding of the $AraR_{160–362}$ region when fused to LexA. These results indicate that AraR is able to form homodimers and suggest that amino acid residues 76 to 362 comprise determinants necessary for dimerization.

Mutagenesis of the *araR* allele and isolation of AraR defective (AraR$^-$) and superrepressor (AraR^s) mutants. To further identify regions in AraR involved in DNA and effector binding or dimerization, we performed nonbiased random mutagenesis of *araR*. We screened for mutations in *araR* leading to the expression of a protein unable to bind DNA, dimerize, or fold properly (constitutive phenotype [AraR$^-$]) and for mutations that affect the ability of the repressor to bind or respond to arabinose (superrepressor phenotype [AraR^s]). The shuttle plasmid pLS30 harboring the *araR* allele was the template for random mutagenesis by PCR in the presence of $MnCl_2$, according to the method of Leung et al. (19). To facilitate downstream mutation mapping and sequencing analysis, pLS30 carries two new restriction sites introduced in the *araR* allele by silent mutagenesis using pLS24 (carrying the wild-type *araR* allele) as a template (see Materials and Methods). The presence of these silent mutations did not interfere with the regulatory activity of AraR, as determined by comparing the β-galactosidase activity of strains IQB351 and IQB352 (Table 1) in the presence and absence of inducer (data not shown). This procedure allowed the division of the *araR* gene into three fragments flanked by unique restriction sites. The mutagenized fragments were cloned back into pLS30. Three different libraries of mutations were obtained after transforming *E. coli*, and clones from each group were pooled and their DNA was extracted and used to transform *B. subtilis* strain IQB350 (Table 1). This strain bears an *araAB′-lacZ* fusion and a deletion in the *araR* gene that leads to constitutive expression of β-galactosidase. To establish a screening strategy, we confirmed that when *B. subtilis* IQB350 was transformed with linearized plasmid (pLS30 or pLS24; see Materials and Methods) the *araR* wild-type allele was integrated at the nonessential *amyE* locus by a double-recombination event; consequently, the *araAB′-lacZ* expression was under the control of AraR. So the expression of the *araR* wild-type allele in these strains was responsible for the white color of the colonies in a complex medium (SFA) containing X-Gal and for the blue color in the same medium in the presence of arabinose. However, when the integrated allele expresses an AraR mutant form unable to bind DNA, dimerize, or fold properly, the strain produces blue colonies in the SFA medium without inducer, which is consistent with an AraR$^-$ phenotype. Deficiencies in response or binding to arabinose lead to the white color of the colonies in SFA medium with inducer, which correlates with an AraR^s phenotype. When this screening strategy was used, several clones showing defects with respect to the regulatory activity of AraR were isolated. DNA sequencing confirmed the presence of single missense mutations in the mutagenized fragments of 12 constitutive mutants and four superrepressors. These 16 fragments were recovered by subcloning into pLS30. The AraR$^-$ mutations were as follows: L-33→S, F-37→S, S-53→P, Q-61→R, I-82→T, E-102→Q, L-157→R, H-226→R, C-271→R, R-285→K, I-308→T, and P-319→L. Superrepressor AraR^s mutations were as follows: T-87→I, T-117→A, S-146→R, and D-211→G. The resulting mutant proteins were designated for the mutated position and amino acid substituted (e.g., AraRL33S).

In previous studies, *B. subtilis* mutants constitutive for arabinose utilization and mutants showing growth deficiencies in minimal medium with arabinose as the sole carbon source were isolated after mutagenesis with *N′*-methyl-*N′*-nitro-*N*-nitrosoguanidine and mapped in the *araR* locus (33, 44). This region of the chromosome was sequenced, and three novel single missense mutations conferring an AraR^s phenotype (E-142→K, G-215→V, and G-215→D) and one responsible for a constitutive phenotype AraR$^-$ (G-138→E) were detected. The DNA fragments carrying these mutations were also subcloned into pLS30.

Mutations leading to inability to bind to the DNA (AraR$^-$) map across the entire length of the primary sequence of AraR, whereas mutations that affect the ability of the repressor to respond to arabinose (AraR^s) are clustered in the C-terminal region

of AraR (Fig. 2). In addition to this collection of mutants produced by random mutagenesis, an in-frame deletion of the HTH motif of AraR was generated in pLS30. The resulting *araR* allele (AraR Δaa13–65) encodes a mutant AraR protein missing 53 residues at the N-terminal region (from S-13 to T-65) (Fig. 2).

In vivo characterization of the AraR mutants. *B. subtilis* strain IQB350, carrying an *araR* deletion, was transformed with the plasmids encoding the mutated *araR* alleles as described in Materials and Methods (Table 1). The regulatory activity of the AraR mutated proteins was analyzed in the resulting strains by determination of the levels of accumulated β-galactosidase activity expressed from the *araAB′-lacZ* fusion under inducing (presence of arabinose) and noninducing (absence of arabinose) conditions. Strains IQB350 and IQB352 (*araR* wild-type allele) were used as controls. The results are summarized in Fig. 4. The presence of the wild-type AraR resulted in 98.6-fold repression, whereas in the *araR*-null mutant (Δ*araR*) regulation was completely abolished. All AraR$^-$ mutations displayed reduced regulatory activity. In mutants AraRL33S, G138E, C271R, and I308T, only residual repression or no repression was observed (Fig. 4A, top), whereas mutants Q61R, E102Q, and R285K showed a less than twofold reduction in repression. In the mutant AraR Δaa13–65, carrying an in-frame deletion comprising the HTH motif, no regulatory activity was detected (Fig. 4A). All AraR^s showed inability to respond to arabinose. The maximal value of increase in β-galactosidase activity in the presence of arabinose was 2.8-fold (D211G) compared to 98.6-fold of the wild type (Fig. 4B, top).

The constitutive phenotype, absence of regulatory activity, could be the result of deficient accumulation of mutant AraR$^-$ forms in the cell. Therefore, the abundance of each variant was tested to identify mutations that affected the production of stable AraR. The quantity of AraR was estimated by Western blotting using AraR antiserum and equivalent amounts of crude extracts of cells grown in inducing and noninducing conditions. Strains IQB350 and IQB352 were used as controls; the results of the Western blot analyses are shown in Fig. 5. The cellular level of the wild-type AraR increased in inducing conditions, which correlates with the negative regulation exerted by AraR with respect to its own expression (42). Mutants AraR Δaa13–65, AraRL33S, F37S, S53P, G138E, and C271R exhibited noticeable defects in accumulation of AraR (Fig. 5), suggesting a decrease of stability of these mutant forms of AraR. However, this decrease in stability does not correlate in all cases with the observed regulatory defect (see Discussion). The level of accumulation of the variants Q61R, I82T, E102Q, L157R, H226R, R285K, I308T, and P319L was comparable to that seen with wild-type AraR.

Site-directed mutagenesis of AraR residues. The structure of the modeled AraR sugar-binding pocket suggested that amino acids F94, D212, Q214, R218, D301, and H318 might contact arabinose (Fig. 3A). The construction of mutants F94A, D212A, Q214A, R218L, D301N, and H318A directly tested the prediction of the functional role of these residues. Substitutions of amino acids were designed to minimize structure disruption. The model depicted in Fig. 3B indicated as likely candidates for the dimerization interface a core of hydrophobic amino acid residues surrounded by polar and charged residues. These amino acids were subjected to mutational analysis and exchanged to I89A, Y92F, R99A, L114A, and F305A. Residues most likely to be interacting with DNA were also identified, and one of them, R45, located in the recognition helix (Fig. 3C), was changed to alanine.

All these AraR mutants generated by site-directed mutagen-

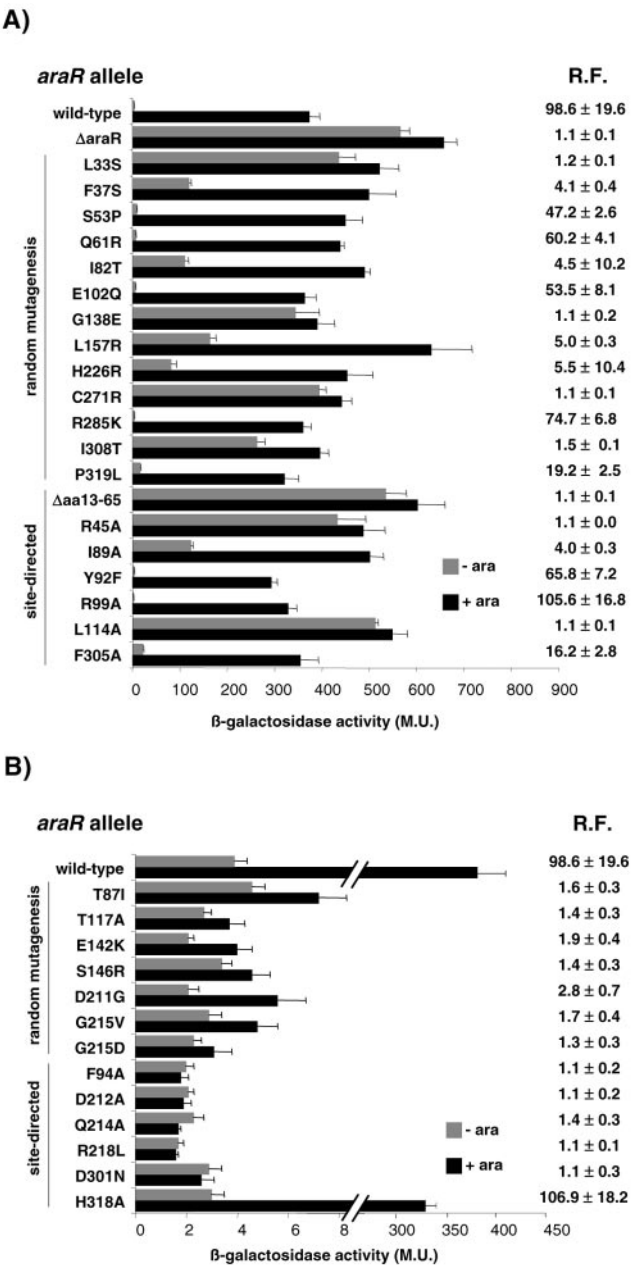


FIG. 4. In vivo characterization of mutant AraR proteins. β-Galactosidase activities of *B. subtilis* strains carrying an *araAB′-lacZ* fusion and an *araR* allele integrated at the *amyE* locus and grown in the absence or presence of arabinose (in gray and black bars, respectively) are shown. Amino acid substitutions (obtained by random or site-directed mutagenesis) leading to an AraR$^-$ phenotype (A) or AraR^S phenotype (B) are indicated for the mutated position and substituted amino acid by use of the standard one-letter designation. Values represent the averages of three independent experiments, each assayed in duplicate. Error bars represent the standard deviation. R.F. indicates the repression factor, calculated as the ratio between values obtained in the presence and in the absence of inducer. M.U., Miller units.

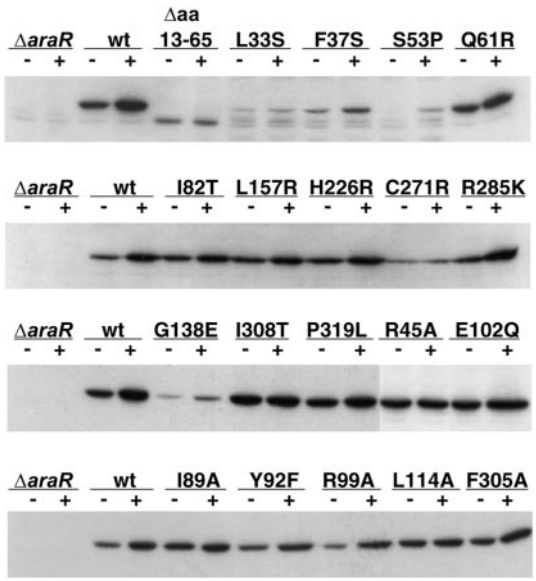


FIG. 5. AraR accumulation in the cell determined by Western immunoblot analysis. Equal amounts of the soluble fractions of cell extracts obtained from *B. subtilis* cultures harboring a wild-type, mutant, or no *araR* allele and grown in the absence (−) or presence (+) of inducer were prepared as described in Materials and Methods. Mutant proteins are designated for the mutated position and substituted amino acid by use of the standard one-letter code.

esis in plasmid pLS30 (Materials and Methods) were tested for their regulatory activity in vivo after transformation of the *B. subtilis* strain IQB350 (Δ*araR araAB′-lacZ*) as described above (see Table 1). As predicted by the model, all mutants designed to disrupt contact with the sugar, except H318A, displayed a superrepressor phenotype (AraR^s) by showing an inability to respond to arabinose (Fig. 4B, bottom). In the AraR mutant forms I89A, L114A, and F305A, constructed to prevent the assembly of monomers, a lower level of repression or no repression was observed (Fig. 4A, bottom). However, mutant Y92F resulted in a less than twofold reduction in repression and variant R99A displayed a level of regulation similar to that of wild-type AraR. The AraR R45A variant, designed to prevent binding to the DNA, exhibited a complete loss of regulatory activity (Fig. 4A, bottom). Additionally, the cellular abundance of each mutant variant leading to decreased regulatory activity, the constitutive (AraR$^-$) phenotype, was estimated by Western blot analysis in crude cellular extracts as described above. The results shown in Fig. 4 indicated that accumulation of the mutant forms of repressor was comparable to that seen with wild-type AraR.

The relevance of amino acids I89 and L114 in the assembly of AraR dimer was further analyzed by constructing LexA-AraR chimeras bearing mutations in these positions. An AraR full-length chimera carrying a single L114A mutation did not affect significantly the expression of the *sulA-lacZ* fusion compared to the results seen with the wild type (Table 3). However, a chimera bearing a double mutation, I89A and L114A, displayed a level of repression very similar to that constructed with a truncated form of AraR$_{258–362}$ (Table 3), indicating the importance of these amino acids in the formation of dimers.

Negative transdominance of AraR mutants. *B. subtilis* strain IQB101 (wild-type *araR*), harboring an *araAB′-lacZ* fusion, was transformed with the plasmids encoding selected AraR mutants obtained by site-directed mutagenesis. The mutated *araR* alleles, under the control of the IPTG-inducible *Pspac* promoter, were integrated at the nonessential *amyE* locus of the receptor strain by a double-recombination event. Since the position of this locus is near the origin of replication of the chromosome, more than one copy of the mutated allele is expected to be present in the cell; additionally, the expression of the mutated *araR* allele is under the control of the strong *Pspac* promoter (11). Together, these conditions should lead to an excess of the AraR mutant form over the wild type. A significant decrease of regulatory activity was observed with mutant R45A (affecting DNA binding), and loss of response to arabinose was detected with mutant F94A, predicted to be involved in binding of the sugar (Table 4). These results indicate titration of the wild-type AraR by assembly of heterodimers with inactive mutant proteins. In contrast, mutants I89A, L114A, and F305A were recessive, which is consistent with the assigned role in dimerization for the exchanged residues.

DISCUSSION

The *B. subtilis* AraR transcription factor plays an important role in carbohydrate utilization by controlling the transport and catabolism of arabinose and the uptake of xylose and galactose, three structurally different sugars that often occur associated in hemicelluloses. The mutational analysis presented here allowed the identification of the regions and specific amino acid residues involved in the different molecular events underlying the mechanism of transcriptional repression by AraR.

We used random mutagenesis of the *araR* allele and a genetic approach to identify missense mutations that resulted in a noninducible, superrepressor phenotype (AraR^s). Additionally, three other *araR* alleles conferring an AraR^s phenotype

TABLE 4. Transdominance of AraR mutants over the wild-type AraR in *B. subtilis*[a]

Strain	IPTG-inducible *araR* allele	β-Galactosidase activity		R.F.
		−ara	+ara	
IQB579		3.8 ± 0.4	400.8 ± 21.1	107.6 ± 16.4
IQB580	Wild type	3.8 ± 0.5	245.6 ± 13.7	65.5 ± 11.2
IQB580	Wild type (no IPTG)	3.3 ± 0.7	380.7 ± 42.8	119.4 ± 28.6
IQB584	F94A	3.1 ± 0.7	4.5 ± 0.6	1.5 ± 0.2
IQB584	F94A (no IPTG)	3.6 ± 0.7	360.1 ± 10.9	103.3 ± 18.5
IQB585	L114A	4.1 ± 0.6	364.1 ± 18.4	90.0 ± 14.9
IQB591	I89A	4.0 ± 0.7	311.4 ± 31.5	80.0 ± 13.1
IQB592	F305A	3.0 ± 0.6	281.4 ± 14.7	96.6 ± 15.4
IQB581	R45A	68.7 ± 2.9	435.9 ± 16.2	6.4 ± 0.2
IQB581	R45A (no IPTG)	4.4 ± 0.7	387.7 ± 21.3	90.2 ± 16.7

[a] β-Galactosidase activities (in Miller units) of B. subtilis strains carrying an *araAB′-lacZ* fusion and a *Pspac-araR* fusion integrated at the *amyE* locus and grown in the absence or presence of arabinose (−ara or +ara) and in the presence of IPTG, unless otherwise specified. The *araR* mutant alleles are indicated for the mutated position and amino acid substituted in AraR by use of the standard one-letter code. Values represent the averages and standard deviations of three independent experiments, each assayed in duplicate. R.F. indicates the repression factor, calculated as the ratio between values obtained in the presence and in the absence of inducer.

previously identified by Paveia and Archer (33) were characterized in this work. The AraR^s variants carrying single-amino-acid substitutions, T87I, T117A, E142K, S146R, D211G, G215V, and G215D, mapped in the carboxyl terminus of AraR and in positions highly conserved among the AraR-like proteins but not in other LacI/GalR family members (Fig. 2), suggesting their importance in the AraR-specific protein function. The AraR^s phenotype, quantified in vivo (Fig. 4B), could be due either to a decreased ability to bind arabinose or to an inability to undergo an allosteric transition that results in a conformational change of the repressor preventing binding to the cognate DNA. The three-dimensional structures of the effector-binding domain determined for some members of the LacI/GalR family show a common fold that is analogous to those of periplasmic-binding proteins (9, 20, 34, 50), being composed of two similar subdomains in whose interface is located the effector-binding cleft (Fig. 3D). The structure of the C-terminal domain of the *E. coli* PurR was used to model the C-terminal part of AraR, and the arabinose-binding pocket was modeled after the structure of the *E. coli* D-ribose binding protein (see above). Based on this model the AraR^s variants isolated displayed substitutions in residues located close to the arabinose cleft but not predicted to be in direct contact with the effector molecule (Fig. 3A). The residues more likely to be in contact with the sugar, F94, D212, Q214, R218, D301, and H318, were mutagenized, and all except H318 resulted in a noninducible AraR^s phenotype (Fig. 3A). Mutations of the corresponding residues in LacI (P76, S191, S193, R197, D274, and Q291) (Fig. 2) also give a superrepressor phenotype except mutations in Q291 and the precise change P76A (21, 31, 52). Mutations in residues R196 and D275 of PurR (R218 and D301 in AraR, respectively) (Fig. 2) also result in effector-binding defects (49). Since the mutations introduced in AraR were designed to minimize structure disruption and to probe loss of contacts, the results obtained suggest that amino acids F94, D212, Q214, R218, and D301 are directly involved in sugar binding rather than participating in the allosteric transition mechanism. In summary, the random and site-directed mutagenesis procedures and the subsequent in vivo analysis of the mutants obtained confirmed that, as predicted by the model, the effector-binding domain of AraR is comprised within its C terminal.

A group of missense mutations in the *araR* allele leading to a constitutive phenotype (AraR$^-$) were also isolated and/or characterized. These mutations localized across the entire length of the primary sequence of AraR (Fig. 2). The AraR$^-$ phenotype could be generated from a defect in specific functions of the repressor, such as DNA binding or dimerization, or by protein misfolding. The substitutions I82T, E102Q, G138E, L157R, H226R, C271R, R285K, I308T, and P319L mapped in the C-terminal region of AraR (Fig. 2). All mutant forms showed reduced regulatory activity, which is completely abolished in mutants G138E and C271R (Fig. 4A). The intracellular accumulation level of these mutant proteins was similar to that observed in the wild-type AraR. The exceptions were G138E and C271R, which accumulate at lower levels, presumably due to the increase of proteolytic degradation (Fig. 5). In these two mutants the defect in accumulation and the lack of regulatory activity were comparable, suggesting that the former caused the latter. Based on the model the exchanges I82T, E102Q, H226R, C271R, I308T, and P319L are localized in the hydrophobic core of the protein and most probably affect the fold (data not shown). In variants L157R, R285K, and I308T, substitutions are solvent exposed; consequently, the reason for the phenotype displayed is unclear. However, it is noteworthy that R285K and I308T are positioned near the dimer interface.

The members of the group of constitutive mutants L33S, F37S, S53P, and Q61R, isolated after random mutagenesis, are discussed together because the substitutions mapped at the N terminus of AraR (Fig. 2). This region comprises a DNA-binding motif representative of the GntR family of bacterial regulatory proteins (10). The exchanges L33S and F37S resulted in complete loss of regulatory activity, while a twofold reduction was observed with S53P (Fig. 4A). These results do not correlate with the intracellular accumulation of the proteins. In fact, whereas a drastic reduction was observed with mutants L33S and S53P, only a light decrease in the amount of F37S was detected (Fig. 5). The regulatory activity of Q61R, which accumulated at levels similar to those seen with the wild type, decreased less than twofold. In addition, a mutant lacking 53 residues comprising the helix-turn-helix motif was generated and characterized in vivo. Although this variant showed only a small decrease in accumulation in vivo, the regulatory activity was completely abolished (Fig. 4A and Fig. 5). The N-terminal region of AraR was modeled using the structure of the *E. coli* transcription factor FadR, the only member of the GntR family with the structure determined (54, 58). The DNA-binding domain contains an HTH motif of the "winged" type (4, 8). Based on the model only Q61 is predicted to contact the DNA in the minor groove (Fig. 3C), and exchanges of the corresponding residue in FadR (H75) and GntR (R75), the gluconate repressor from *B. subtilis*, were shown to have a transdominant negative phenotype (35, 59). In FadR this residue is at the tip of the wing and buried deep in the minor groove (58). The exchanges L33S and F37S, located in the second helix, may influence binding of other amino acids to the DNA, and S53P makes the third helix, the recognition helix, shorter. The specific role of residue R45 in AraR, predicted to be in contact with the DNA, was tested by exchange to an alanine (Fig. 3C). This mutation led to complete loss of regulatory activity, although it accumulated at wild-type levels (Fig. 4A and Fig. 5). The involvement of this particular residue in AraR-DNA binding is further supported by the transdominant negative phenotype of the *araR* R45A allele in *B. subtilis* (Table 4) and validates the model presented in this work. The corresponding residue in FadR (R49) (Fig. 2) was shown to be in contact with the DNA by crystal data and mutational analysis (35, 58).

On the basis of studies performed with *E. coli* by construction and expression of LexA-AraR chimeric fusions, we show that AraR is able to dimerize in vivo (Table 3). Furthermore, the analysis of different chimeras with truncated versions of AraR suggests that amino acid residues 76 to 362 include determinants necessary for dimerization. The model derived for the C-terminal region of the protein predicts candidate residues for the dimer interface (Fig. 3B) that were mutagenized, and the variants were analyzed according to both regulatory activity and in vivo accumulation. All accumulated at levels comparable to that of the wild type (Fig. 5), and mutants I89A, L114A, and F305A showed severe defects in

their repression capacity whereas Y92F displayed only a twofold reduction in the regulatory activity (Fig. 4A). In addition, the relevance of residues I89 and L114 in the assembly of monomers was confirmed by constructing LexA-AraR chimeras bearing mutations in these positions. An AraR full-length chimera bearing a double mutation, I89A and L114A, affected the expression of the *sulA-lacZ* fusion similarly to the chimera constructed with the truncated version $AraR_{258–362}$ (Table 3). The role of residues I89, L114, and F305 in dimerization is further supported by the classical transdominance experiments (12). In this work we show that alleles containing either mutations involved in sugar binding and response or in DNA binding are dominant over the wild-type allele in *B. subtilis* whereas alleles bearing the exchanges I89A, L114A, and F305A are recessive (Table 4). Interestingly, in CcpA, the master regulator of carbon catabolite repression in *B. subtilis* and a member of the LacI/GalR family, the corresponding residues (I71, L96, and L280; not shown) are predicted on the basis of the crystal structure determined for the C-terminal region to be involved in dimerization (48). The homology between the C terminus of AraR and proteins of the LacI/GalR family allows predicting the location of the arabinose-binding site of AraR between the two C subdomains in the effector-binding cleft (Fig. 3D). Accordingly, the allosteric transition triggered by arabinose can be expected to occur by a mechanism showing similarities to that described for LacI and PurR, the only two full-length members of this family bound to the DNA with structures available. In LacI, the binding of IPTG causes a small motion of the C subdomains, which alters the dimer interface within the N-terminal subdomain (1, 2, 20), leading to the disruption of the interaction between the hinge helices that make important contacts with the minor groove. This frees the DNA binding HTH domain, which becomes disordered (1, 2, 20). PurR-specific binding to the DNA is mediated by either guanine or hypoxanthine. Similarly to the effector-bound form of LacI, the corepressor-free PurR shows a rotation of the C subdomains that results in the destabilization of the hinge helices and disruption of specific DNA binding (49, 50). Since the DNA-binding domain of AraR is not homologous to the members of the LacI/GalR family, the hinge helix that plays a fundamental role in DNA recognition and in the allosteric transition may be absent in AraR, anticipating differences in the allosteric mechanism triggered by arabinose relative to that described for the LacI/GalR proteins. Thus, the rare modular structure of AraR that combines functional domains from different origins (GntR family and LacI/GalR family) allows hypothesizing a novel mechanism of effector-regulated specific DNA binding. To elucidate the mechanisms of transcriptional regulation by AraR and extend the mutation analysis presented in this report, attempts to determine the crystal structure of free AraR, AraR-DNA complexes, and AraR bound to arabinose are currently in progress.

ACKNOWLEDGMENTS

We thank Dayle Daines for providing the *E. coli* strains and vectors used in the construction and analysis of the LexA-AraR chimeras, Leonor Sarmento, Rodrigo Saraiva, Sara Cunha, and Vanessa Barroso for the construction of some plasmids and strains, and Adriano O. Henriques for critical reading of the manuscript.

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.D.S.-N. I.S.F. is the holder of Ph.D. fellowship SFRH/BD/5233/01 from Fundação para a Ciência e Tecnologia (FCT).

REFERENCES

1. **Bell, C. E., and M. Lewis.** 2000. A closer view of the conformation of the Lac repressor bound to operator. Nat. Struct. Biol. **7:**209–214.
2. **Bell, C. E., and M. Lewis.** 2001. The Lac repressor: a second generation of structural and functional studies. Curr. Opin. Struct. Biol. **11:**19–25.
3. **Bjorkman, A. J., R. A. Binnie, H. Zhang, L. B. Cole, M. A. Hermodson, and S. L. Mowbray.** 1994. Probing protein-protein interactions. The ribose-binding protein in bacterial transport and chemotaxis. J. Biol. Chem. **269:**30206–30211.
4. **Brennan, R. G.** 1993. The winged-helix DNA-binding motif: another helix-turn-helix takeoff. Cell **74:**773–776.
5. **Daines, D. A., M. Granger-Schnarr, M. Dimitrova, and R. P. Silver.** 2002. Use of LexA-based system to identify protein-protein interactions in vivo. Methods Enzymol. **358:**153–161.
6. **Delano, W.** 2003. The Pymol Molecular Graphics System, version 0.90. Delano Scientific LLC, San Carlos, Calif.
7. **Englesberg, E., R. L. Anderson, R. Weinberg, N. Lee, P. Hoffee, G. Huttenhauer, and H. Boyer.** 1962. L-Arabinose-sensitive, L-ribulose 5-phosphate 4-epimerase-deficient mutants of *Escherichia coli*. J. Bacteriol. **84:**137–146.
8. **Gajiwala, K. S., and S. K. Burley.** 2000. Winged helix proteins. Curr. Opin. Struct. Biol. **10:**110–116.
9. **Hars, U., R. Horlacher, W. Boos, W. Welte, and K. Diederichs.** 1998. Crystal structure of the effector-binding domain of the trehalose-repressor of *Escherichia coli*, a member of the LacI family, in its complexes with inducer trehalose-6-phosphate and noninducer trehalose. Protein Sci. **7:**2511–2521.
10. **Haydon, D. J., and J. R. Guest.** 1991. A new family of bacterial regulatory proteins. FEMS Microbiol. Lett. **63:**291–295.
11. **Henner, D. J.** 1990. Inducible expression of regulatory genes in *Bacillus subtilis*. Methods Enzymol. **185:**223–228.
12. **Herskowitz, I.** 1987. Functional inactivation of genes by dominant negative mutations. Nature **329:**219–222.
13. **Kraus, A., E. Kuster, A. Wagner, K. Hoffmann, and W. Hillen.** 1998. Identification of a co-repressor binding site in catabolite control protein CcpA. Mol. Microbiol. **30:**955–963.
14. **Krispin, O., and R. Allmansberger.** 1998. The *Bacillus subtilis* AraE protein displays a broad substrate specificity for several different sugars. J. Bacteriol. **180:**3250–3252.
15. **Laskowski, A., M. MacArthur, D. Moss, and J. Thorton.** 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26:**283–291.
16. **Leal, T. F., and I. de Sá-Nogueira.** 2004. Purification, characterization and functional analysis of an endo-arabinanase (AbnA) from *Bacillus subtilis*. FEMS Microbiol. Lett. **241:**41–48.
17. **Lee, M. H., M. Scherer, S. Rigali, and J. W. Golden.** 2003. PlmA, a new member of the GntR family, has plasmid maintenance functions in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **185:**4315–4325.
18. **Lepesant, J. A., and R. Dedonder.** 1967. Metabolism of L-arabinose in *Bacillus subtilis* Marburg Ind-168. C. R. Acad. Sci. Ser. D **264:**2683–2686. (In French.)
19. **Leung, D. W., E. Chen, and D. V. Goeddel.** 1989. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. Technique **1:**11–15.
20. **Lewis, M., G. Chang, N. C. Horton, M. A. Kercher, H. C. Pace, M. A. Schumacher, R. G. Brennan, and P. Lu.** 1996. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. Science **271:**1247–1254.
21. **Markiewicz, P., L. G. Kleina, C. Cruz, S. Ehret, and J. H. Miller.** 1994. Genetic studies of the *lac* repressor. XIV. Analysis of 4,000 altered *Escherichia coli lac* repressors reveals essential and non-essential residues, as well as "spacers" which do not require a specific sequence. J. Mol. Biol. **240:**421–433.
22. **Mendes, J., A. M. Baptista, M. A. Carrondo, and C. M. Soares.** 1999. Improved modelling of side chains in proteins with rotamer-based methods: a flexible rotamer model. Proteins **37:**530–543.
23. **Mendes, J., H. A. Nagarajaram, C. M. Soares, T. L. Blundell, and M. A. Carrondo.** 2001. Incorporating knowledge-based biases into an energy-based side-chain modeling method: application to comparative modeling of protein structure. Biopolymers **59:**72–86.
24. **Mendes, J., C. M. Soares, and M. A. Carrondo.** 1999. Improvement of side-chain modelling in proteins with the self-consistent mean field theory method based on an analysis of the factors influencing prediction. Biopolymers **50:**111–131.
25. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Mota, L. J., L. M. Sarmento, and I. de Sá-Nogueira.** 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.
27. **Mota, L. J., P. Tavares, and I. de Sá-Nogueira.** 1999. Mode of action of

AraR, the key regulator of L-arabinose metabolism in *Bacillus subtilis*. Mol. Microbiol. **33:**476–489.
28. **Mowbray, S. L., and L. B. Cole.** 1992. 1.7 A X-ray structure of the periplasmic ribose receptor from *Escherichia coli*. J. Mol. Biol. **225:**155–175.
29. **Nagadoi, A., S. Morikawa, H. Nakamura, M. Enari, K. Kobayashi, H. Yamamoto, G. Sampei, K. Mizobuchi, M. A. Schumacher, and R. G. Brennan.** 1995. Structural comparison of the free and DNA-bound forms of the purine repressor DNA-binding domain. Structure **3:**1217–1224.
30. **Pabo, C. O., and R. T. Sauer.** 1992. Transcription factors: structural families and principles of DNA recognition. Annu. Rev. Biochem. **61:**1053–1095.
31. **Pace, H. C., M. A. Kercher, P. Lu, P. Markiewicz, J. H. Miller, G. Chang, and M. Lewis.** 1997. Lac repressor genetic map in real space. Trends Biochem. Sci. **22:**334–339.
32. **Pascal, M., F. Kunst, J. A. Lepesant, and R. Dedonder.** 1971. Characterization of two sucrase activities in *Bacillus subtilis* Marburg. Biochimie **53:**1059–1066.
33. **Paveia, H., and L. Archer.** 1992. Genes for L-arabinose utilization in *Bacillus subtilis*. Brotéria Genét. **13:**149–159.
34. **Quiocho, F. A., and P. S. Ledvina.** 1996. Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. Mol. Microbiol. **20:**17–25.
35. **Raman, N., P. N. Black, and C. C. DiRusso.** 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.
36. **Raposo, M. P., J. M. Inácio, L. J. Mota, and I. de Sá-Nogueira.** 2004. Transcriptional regulation of genes encoding arabinan-degrading enzymes in *Bacillus subtilis*. J. Bacteriol. **186:**1287–1296.
37. **Real, G., S. Autret, E. J. Harry, J. Errington, and A. O. Henriques.** 2005. Cell division protein DivIB influences the Spo0J/Soj system of chromosome segregation in *Bacillus subtilis*. Mol. Microbiol. **55:**349–367.
38. **Rigali, S., A. Derouaux, F. Giannotta, and J. Dusart.** 2002. Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MocR, and YtrA subfamilies. J. Biol. Chem. **277:**12507–12515.
39. **Sali, A., and J. P. Overington.** 1994. Derivation of rules for comparative protein modeling from a database of protein structure alignments. Protein Sci. **3:**1582–1596.
40. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
41. **Sá-Nogueira, I., and H. de Lencastre.** 1989. Cloning and characterization of *araA*, *araB*, and *araD*, the structural genes for L-arabinose utilization in *Bacillus subtilis*. J. Bacteriol. **171:**4088–4091.
42. **Sá-Nogueira, I., and L. J. Mota.** 1997. Negative regulation of L-arabinose metabolism in *Bacillus subtilis*: characterization of the *araR* (*araC*) gene. J. Bacteriol. **179:**1598–1608.
43. **Sá-Nogueira, I., T. V. Nogueira, S. Soares, and H. de Lencastre.** 1997. The *Bacillus subtilis* L-arabinose (*ara*) operon: nucleotide sequence, genetic organization and expression. Microbiology **143:**957–969.
44. **Sá-Nogueira, I., H. Paveia, and H. de Lencastre.** 1988. Isolation of constitutive mutants for L-arabinose utilization in *Bacillus subtilis*. J. Bacteriol. **170:**2855–2857.
45. **Sá-Nogueira, I., and S. S. Ramos.** 1997. Cloning, functional analysis, and transcriptional regulation of the *Bacillus subtilis araE* gene involved in L-arabinose utilization. J. Bacteriol. **179:**7705–7711.
46. **Schleif, R.** 2000. Regulation of the L-arabinose operon of *Escherichia coli*. Trends Genet. **16:**559–565.
47. **Schleif, R.** 2003. AraC protein: a love-hate relationship. Bioessays **25:**274–282.
48. **Schumacher, M. A., G. S. Allen, M. Diel, G. Seidel, W. Hillen, and R. G. Brennan.** 2004. Structural basis for allosteric control of the transcription regulator CcpA by the phosphoprotein HPr-Ser46-P. Cell **118:**731–741.
49. **Schumacher, M. A., K. Y. Choi, F. Lu, H. Zalkin, and R. G. Brennan.** 1995. Mechanism of corepressor-mediated specific DNA binding by the purine repressor. Cell **83:**147–155.
50. **Schumacher, M. A., K. Y. Choi, H. Zalkin, and R. G. Brennan.** 1994. Crystallization and preliminary X-ray analysis of an *Escherichia coli* purine repressor-hypoxanthine-DNA complex. J. Mol. Biol. **242:**302–305.
51. **Scott, W. R. P., P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, and W. F. van Gunsteren.** 1999. The GROMOS biomolecular simulation program package. J. Phys. Chem. **103:**3596–3607.
52. **Suckow, J., P. Markiewicz, L. G. Kleina, J. Miller, B. Kisters-Woike, and B. Muller-Hill.** 1996. Genetic studies of the Lac repressor. XV: 4,000 single amino acid substitutions and analysis of the resulting phenotypes on the basis of the protein structure. J. Mol. Biol. **261:**509–523.
53. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
54. **van Aalten, D. M., C. C. DiRusso, J. Knudsen, and R. K. Wierenga.** 2000. Crystal structure of FadR, a fatty acid-responsive transcription factor with a novel acyl coenzyme A-binding fold. EMBO J. **19:**5167–5177.
55. **van Gunsteren, W. F., S. R. Billeter, A. A. Eising, P. H. Hunenberger, P. Kruger, A. E. Mark, W. R. P. Scott, and I. G. Tironi.** 1996. Biomolecular simulation: the GROMOS96 manual and user guide. Groninger, Zurich, Switzerland.
56. **Weickert, M. J., and S. Adhya.** 1992. A family of bacterial regulators homologous to Gal and Lac repressors. J. Biol. Chem. **267:**15869–15874.
57. **Wipat, A., N. Carter, S. C. Brignell, B. J. Guy, K. Piper, J. Sanders, P. T. Emmerson, and C. R. Harwood.** 1996. The *dnaB-pheA* (256 degrees–240 degrees) region of the *Bacillus subtilis* chromosome containing genes responsible for stress responses, the utilization of plant cell walls and primary metabolism. Microbiology **142:**3067–3078.
58. **Xu, Y., R. J. Heath, Z. Li, C. O. Rock, and S. W. White.** 2001. The FadR.DNA complex. Transcriptional control of fatty acid metabolism in *Escherichia coli*. J. Biol. Chem. **276:**17373–17379.
59. **Yoshida, K., Y. Fujita, and A. Sarai.** 1993. Missense mutations in the *Bacillus subtilis gnt* repressor that diminish operator binding ability. J. Mol. Biol. **231:**167–174.