The Lacl-Type Transcriptional Regulator AraR Acts as an L-Arabinose-Responsive Repressor of L-Arabinose Utilization Genes in Corynebacterium glutamicum ATCC 31831

Takayuki Kuge, Ab Haruhiko Teramoto, A Hideaki Yukawa, A Masayuki Inui A,b

Research Institute of Innovative Technology for the Earth, Kizugawadai, Kizugawa, Kyoto, Japan; Graduate School of Biological Sciences, Nara Institute of Science and Technology, Takayama, Ikoma, Nara, Japanb

The Corynebacterium glutamicum ATCC 31831 araBDA operon consists of three l-arabinose catabolic genes, upstream of which the galM, araR, and araE genes are located in opposite orientation. araR encodes a Lacl-type transcriptional regulator that negatively regulates the l-arabinose-inducible expression of araBDA and araE (encoding an l-arabinose transporter), through a mechanism that has yet to be identified. Here we show that the AraR protein binds in vitro to three sites: one upstream of araBDA and two upstream of araE. We verify that a 16-bp consensus palindromic sequence is essential for binding of AraR, using a series of mutations introduced upstream of araB in electrophoretic mobility shift assays. Moreover, the DNA-binding activity of AraR is reduced by l-arabinose. We employ quantitative reverse transcription-PCR (qRT-PCR) analyses using various mutant strains deficient in l-arabinose utilization genes to demonstrate that the prominent upregulation of araBDA and araE within 5 min of l-arabinose supplementation is dependent on the uptake but independent of the catabolism of l-arabinose. Similar expression patterns, together with the upregulation by araR disruption without l-arabinose, are evident with the apparent galM-araR operon, although attendant changes in expression levels are much smaller than those realized with the expression of araBDA and araE. The AraR-binding site upstream of araB overlaps the −10 region of the divergent galM promoter. These observations indicate that AraR acts as a transcriptional repressor of araBDA, araE, and galM-araR and that l-arabinose acts as an intracellular negative effector of the AraR-dependent regulation.

Corynebacterium glutamicum is a Gram-positive Actinobacteria species with a high G+C content in its genomic DNA. It is currently used in large-scale industrial production of amino acids such as l-glutamate and l-lysine (1, 2). Moreover, investigation into its application in the production of fuels and commodity chemicals has started to bear fruit (3–5). This is predicated on its potential to utilize cellulosic biomass, a renewable energy source that should play an important role in the global renewable energy of the future (6, 7). Cellulosic biomass contains significant amounts of pentose sugars, such as d-xylose and l-arabinose, that are often difficult for conventional bioprocesses to exploit. Recent advances in metabolic engineering of C. glutamicum have helped to open up new possibilities for efficient utilization of substrates containing mixtures of d-glucose, d-xylose, and l-arabinose (8–11). An understanding of the transcriptional regulation of sugar metabolism genes in C. glutamicum is an attractive avenue toward optimization of utilization of such sugar mixtures.

L-Arabinose is utilized by a Gram-negative bacterium, Escherichia coli, and a low-G+C Gram-positive bacterium, Bacillus subtilis, via a common pathway. L-Arabinose taken up by cells is sequentially converted to l-ribulose, l-ribulose 5-phosphate, and d-xylulose 5-phosphate by the actions of l-arabinose isomerase (encoded by araA), l-ribulokinase (encoded by araB), and l-ribulose 5-phosphate 4-epimerase (encoded by araD), respectively. D-Xylulose 5-phosphate is further catabolized through the pentose phosphate pathway. Despite the conservation of this pathway among various bacteria, two distinct regulatory systems of the l-arabinose utilization genes have been well established so far. In E. coli, AraC acts as an activator and as a repressor that regulates transcription of the genes required for the catabolism (araBAD) and uptake (araE and araFGH) of l-arabinose (12). AraC is composed of an N-terminal arabinose-binding and dimerization domain and a C-terminal DNA-binding domain. The expression of the araBAD operon and the AraC-encoding gene (araC), which are divergently transcribed, is repressed by binding of AraC to the promoter region forming a DNA loop in the absence of l-arabinose (13–15). In the presence of l-arabinose, three promoters, ParaBAD, ParaE, and ParaFGH, are activated, and ParaC is derepressed (16). In B. subtilis, AraR, which is structurally different from E. coli AraC, controls the transcription of several genes involved in the uptake and degradation of arabinose-containing polysaccharides, in addition to the genes corresponding to the E. coli AraC regulon (17–19). AraR comprises an N-terminal DNA-binding domain homologous to the GntR family proteins and a C-terminal effector-binding domain that shows similarity to members of the Lacl/GalR family (20, 21). B. subtilis AraR binds to five different promoter regions with distinct affinities (ParaAB-DLMNPQ-αβα, ParaE, Pαβα, Pxxα, and ParaR) (18, 19, 22). Moreover, an l-arabinose transporter encoded by araE under the control of AraR is also responsible for the uptake of d-xylose and d-galactose (23). Therefore, AraR is thought to play a central role in the regulation of utilization of the multiple sugars in B. subtilis. It is noted that the l-arabinose catabolic and uptake genes are also
subject to carbon catabolite repression, controlled by the cyclic AMP (cAMP)-dependent transcriptional regulator CRP in E. coli (24) and the LacI family transcriptional regulator CcpA in B. subtilis (25). These distinct regulatory systems are responsible for inhibition of l-arabinose utilization in the presence of d-glucose.

Most strains of C. glutamicum tested cannot grow using l-arabinose as the sole carbon source owing to the lack of the pentose sugar utilization pathway (26). However, C. glutamicum ATCC 31831 was recently shown to efficiently utilize l-arabinose due to the presence of araBDA and araE in its genome (26). The expression of these genes associated with l-arabinose utilization is induced in the presence of l-arabinose and is also upregulated by inactivation of araR, encoding a LacI family transcriptional regulator, in the absence of l-arabinose (26). Although the AraR protein is suggested to be involved in the regulation of the l-arabinose-inducible genes in this strain, whether or not AraR directly regulates the relevant genes remains to be verified. It should be noted that AraR of C. glutamicum ATCC 31831 and AraR of B. subtilis belong to distinct families, although the latter protein reveals a mosaic structure with a C-terminal effector-binding domain that has some similarity to the former, LacI family protein. Furthermore, C. glutamicum ATCC 31831 simultaneously utilizes l-arabinose and d-glucose (26), apparently in contrast to the glucose repression of l-arabinose utilization genes in E. coli and B. subtilis described earlier. Thus, elucidation of the regulatory mechanism of the l-arabinose utilization genes in C. glutamicum ATCC 31831 should extend our understanding of the regulatory systems of sugar metabolism diversified among bacterial species.

In this study, we show that C. glutamicum ATCC 31831 AraR binds in vitro to 16-bp consensus palindromic sequences found separately at one site upstream of araBDA and at two sites upstream of araE. The AraR-binding site upstream of araB overlaps the promoter of the apparent galM araR operon, divergently transcribed from araBDA. The binding activity of AraR is reduced by l-arabinose. In addition, we show effects of araR, araB, araD, araA, or araE deletion on l-arabinose-inducible gene expression in vivo that suggest that AraR acts as a transcriptional repressor of araBDA, araE, and galM-araR. The AraR-mediated expression of these multiple transcriptional units is coordinately derepressed in the presence of l-arabinose in response to uptake of the sugar, but independent of its catabolism.

MATERIALS AND METHODS

Bacterial strains. C. glutamicum ATCC 31831 was used as a wild-type strain in this study. The araB, araD, and araR single-deletion mutant strains were described previously (26). E. coli strains JM109 (Takara), JM110 (27), and BL21(DE3) (28) were used for cloning and/or expression of genes of interest.

Culture conditions. For genetic manipulations, E. coli strains were grown at 37°C in Luria-Bertani (LB) medium (27). C. glutamicum strains were grown at 33°C in nutrient-rich A medium [2 g/liter yeast extract, 7 g/liter Casamino Acids, 2 g/liter (NH₄)₂CO, 7 g/liter (NH₄)₂SO₄, 0.5 g/liter KH₂PO₄, 0.5 g/liter KHPO₄, 0.5 g/liter MgSO₄·7H₂O, 6 mg/liter FeSO₄·7H₂O, 4.2 mg/liter MnSO₄·H₂O, 0.2 mg/liter biotin, 0.2 mg/liter thiamine] (29) with 4% glucose. Where appropriate, the culture medium was supplemented with 10 mg/ml of kanamycin.

For analytical purposes, a C. glutamicum starter culture was grown aerobically in 10 ml of nutrient-rich A medium in a 100-ml test tube overnight. The cells were inoculated into fresh medium at a dilution of 100-fold or higher. The cells were cultivated in 100 ml of nutrient-rich A medium at 33°C in a 500-ml flask. To assess the response to sugars, the medium was supplemented with l-arabinose and/or d-glucose at the stated concentrations.

DNA technique. Chromosomal DNA was isolated from C. glutamicum cells by using a GenomicPrep cell and tissue DNA isolation kit (GE Healthcare, Buckinghamshire, England) according to the manufacturer’s instructions, but modified by using 10 mg/ml lysosome at 37°C for 1 h. Plasmid DNA was isolated using a QiAprep Spin miniprep kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions, but modified for extraction from C. glutamicum cells by using 10 mg/ml lysosome at 37°C for 1 h.

PCR was performed using a model 480 DNA thermal cycler (PerkinElmer, MA). After the reaction mixture was heated at 96°C for 2 min, the PCR proceeded with 30 cycles of 15 s at 96°C, 30 s at 55°C, and 30 s at 72°C by using DNA polymerase KOD Plus (Toyobo, Osaka, Japan).

C. glutamicum cells were transformed by electroporation as described previously (30). E. coli cells were transformed by the CaCl₂ procedure (27).

DNA sequencing was performed by the dyeoxy chain termination method (31), using an ABI Prism 3100xl genetic analyzer (Applied Biosystems, Foster City, CA) and a BigDye Terminator, version 3.1, cycle sequencing kit (Applied Biosystems). DNA sequence data were analyzed with the Genetyx program (Genetyx Corporation, Tokyo, Japan). Sequences were aligned by use of the ATGC program (Genetyx Corporation).

Mutant construction. The upstream and downstream regions of the target gene for deletion were amplified using sets of primers summarized in Table 1. The resultant amplicons were fused and cloned into pCR2067 (32), a suicide vector for markerless gene disruption. The resultant plasmids, pCRF500 and pCRF501, were used for in-frame deletion of araR and araE, respectively. C. glutamicum was subsequently transformed with the respective plasmid DNA, and screening for deletion mutants was performed as described previously (32). Deletion of the target genes was confirmed by PCR.

Quantitative reverse transcription-PCR (qRT-PCR) analysis and rapid amplification of cDNA ends (RACE). Total RNA was prepared from C. glutamicum cells by using an RNasey minikit (Qiagen) according to the manufacturer’s instructions. Single-stranded cDNA was synthesized from 0.2 μg of total RNA by using PrimeScript reverse transcriptase (Takara, Osaka, Japan) with a random hexamer mixture as a primer in 20 μl of reaction mixture, and then 2 μl of the cDNA mixture was added as a template to 18 μl of reaction mixture containing each primer (0.3 μM) and Power SYBR green PCR master mix ( Applied Biosystems). After the reaction mixture was heated at 95°C for 10 min, PCRs proceeded via 40 cycles of 15 s at 95°C and 40 s at 60°C. The amount of amplified DNA was monitored by fluorescence at the end of each cycle by using an Applied Biosystems 7500 Fast real-time PCR system. The primers used are listed in Table 1. The relative abundances of the target mRNAs were quantified based on the cycle threshold value. To standardize the results, the relative abundance of 16S rRNA was used as the internal standard.

The 5’ end of mRNA was determined by RACE, using the primers summarized in Table 1. Using a 5’-Full RACE core set (Takara, Osaka, Japan), single-stranded cDNA synthesized from total RNA by using the 5’-phosphorylated primer was self-ligated with T4 RNA ligase. The first PCR proceeded using inverted primers, and then the second PCR proceeded using nested inverted primers. The amplified DNA was inserted into pUC118 Hinfl/BAP by using a Mighty Cloning Reagent set (TaKaRa, Osaka, Japan). More than 10 clones from E. coli transformed with the resulting plasmid were sequenced.

Purification of the AraR protein expressed in E. coli. A DNA fragment containing the araR gene was amplified by PCR using primers AraR-Ex-F and AraR-Ex-R (Table 1). The amplified DNA was digested with NdeI and EcoRI and was inserted into the corresponding site of the pET-28a expression vector (Merck KGaA, Darmstadt, Germany). The resulting plasmid, pCRF502, contains the araR gene fused to a His tag sequence at the N terminus. E. coli BL21(DE3) cells transformed with pCRF502 were
grown at 37°C in 100 ml of LB medium supplemented with kanamycin (50 μg ml⁻¹)\(^b\). The recombinant gene was expressed in exponentially growing cells (optical density at 610 nm [OD₆₁₀] = 0.6) by adding 1 mM isopropyl-β-D-thiogalactopyranoside. After 1 h of incubation, the cells were harvested by centrifugation. The His-tagged AraR protein was extracted and purified by affinity column chromatography, using a Ni-nitrilotriacetic acid (Ni-NTA) Fast Start kit (Qiagen). The AraR protein was loaded onto a gel filtration column (PD-10 column; GE Healthcare UK Ltd., Little Chalfont, United Kingdom) equilibrated with buffer containing 20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10 mM MgCl₂, and 1 mM dithiothreitol and eluted with the same buffer. The resultant AraR protein was used for an electrophoretic mobility shift assay (EMSA).

**RESULTS**

1-Arabinose-inducible expression of the 1-arabinose utilization cluster genes. On the chromosome of *C. glutamicum* ATCC 31831, six genes, i.e., *araA*, *araB*, *araD*, *araE*, *galM*, and *araR*, are located in a cluster (Fig. 1) (26). *araA*, *araB*, and *araD*, each of which is essential for 1-arabinose utilization in this strain (26), encode arabinose isomerase, ribulokinase, and ribulose-5-phosphate 4-epimerase, respectively. An H⁺ symporter encoded by *araE* is involved in 1-arabinose uptake (8). *galM* and *araR* encode a putative aldose 1-epimerase and a LacI-type transcriptional regulator, respectively. The *araB* gene is oriented in the same direction as its two downstream genes, *araD* and *araE*. The *araB-araD-araE* intergenic regions are 13 and 35 bp long, respectively. Upstream of the *araB* gene, *galM* is divergently oriented. *araD* and *araE* are located downstream of *galM*, in the same direction. The *galM-araR* and *araR-araE* intergenic regions are 26 and 234 bp long, respectively. We previously reported that the expression levels of *araBDA* and *araE* in this 1-arabinose utilization gene cluster are dramatically increased in exponentially growing cells in minimal medium supplemented with 1-arabinose compared with that in cells grown on d-glucose (26). Here we examined changes in the expression of

\(^a\) The restriction site overhangs used in the cloning procedure are underlined in the sequences.
\(^b\) 5'-Phosphorylated primer.
\(^c\) 5'-IRD700-labeled primer.

### Table 1 Primers used in this study

<table>
<thead>
<tr>
<th>Primer purpose and name</th>
<th>Sequence (5’-3’)</th>
<th>Overhanging restriction site(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid construction for gene deletion</td>
<td>XbaI</td>
<td>XbaI</td>
</tr>
<tr>
<td>araA-ML-U-F</td>
<td>GTTCTACGGTTCGAGTAGTGGGCGGA</td>
<td>XbaI</td>
</tr>
<tr>
<td>araA-ML-D-F</td>
<td>GTTCTACGGTTCGAGTAGTGGGCGGA</td>
<td>XbaI</td>
</tr>
<tr>
<td>araB-ML-U-F</td>
<td>GTTCTACGGTTCGAGTAGTGGGCGGA</td>
<td>XbaI</td>
</tr>
<tr>
<td>araB-ML-D-F</td>
<td>GTTCTACGGTTCGAGTAGTGGGCGGA</td>
<td>XbaI</td>
</tr>
<tr>
<td>araE-ML-U-F</td>
<td>GTTCTACGGTTCGAGTAGTGGGCGGA</td>
<td>XbaI</td>
</tr>
<tr>
<td>araE-ML-D-F</td>
<td>GTTCTACGGTTCGAGTAGTGGGCGGA</td>
<td>XbaI</td>
</tr>
</tbody>
</table>

### DNase I footprinting analysis

A labeled DNA fragment was prepared by PCR, using pUC-IR-Fw and a 5’-IRD700-labeled primer, pUC118 (TaKaRa) containing the region between positions +90 and +34 with respect to the transcription start site of *araB* or the region between positions +85 and +104 with respect to the transcription start site at *araE* was used as the template. The purified AraR protein at the stated concentrations was incubated with a DNA probe in 20 μl of binding buffer containing 20 mM Tris-HCl (pH 7.5), 50 mM NaCl, 2.5 mM MgCl₂, 0.45 mM EDTA, 0.5 mM dithiothreitol, and 10% (vol/vol) glycerol for 25 min at 25°C. The binding reaction mixture was subjected to electrophoresis on a 6% polyacrylamide gel containing 5% (vol/vol) glycerol in 0.5 mM EDTA and subsequent heating at 80°C for 10 min. The samples were mixed with IR2 stop solution (Li-Cor, Lincoln, NE), heated at 95°C for 3 min, and separated in a 5.5% KB Plus gel matrix (Li-Cor), using a Li-Cor 4300 DNA analyzer. DNA sequencing reaction mixtures using the same IRD700-labeled primer and a DYNEnamic direct cycle sequencing kit with 7-deaza-dGTP (GE Healthcare UK Ltd.) were subjected to the same gel analysis.

### Electrophoretic mobility shift assay

**RT-PCR analysis**

165-F | CAGGTCTTCGGGAGATACCTAG |
165-R | CGTTAAGCGGATGACTACCA |
araA-F | GCCTGACGAAGACATTGATCT |
araA-R | TTGTCCTGAGGGATCTGGAAGA |
araB-F | CAGCAGAGAGAAGTTGTAGTG |
araB-R | AGCTTCGATTGGTGTTATCGCT |
araC-F | GTGACATCTGGCGAGGATAG |
araC-R | GGATCGAGCTCTTCAGTGATG |
araD-F | CTTCTGGGGCGATGTGTCAT |
araD-R | AGCTTCGATTGGTGTTATCGCT |
araE-F | CCGCGGGAACAGTTCGCGCA |
araE-R | CTTCTGGGGCGATGTGTCAT |
araF-F | CGAGCGACAAACACGACAA |
araF-R | CTTCTGGGGCGATGTGTCAT |
araG-F | CTCTGTCGCTGCGTCTGCTC |
araG-R | GGTGCTGCGCTGCGTCTGCT |
araH-F | GTCGTTCTGGGCGATGTGTCAT |
araH-R | GTTCTACGGTTCGAGTAGTGGGCGGA |
araI-F | GTTCGGTTCCTGGATGAC |
araI-R | GGATCGAGCTCTTCAGTGATG |
araJ-F | CGAGGGACACGACGAC |
araJ-R | CTTCTGGGGCGATGTGTCAT |
araK-F | GCTCTACGGTTCGAGTAGTGGGCGGA |
araK-R | TTGTCCTGAGGGATCTGGAAGA |
araL-F | TGAGAACTAGGTCG |
araL-R | GCGGGAACAGCTATGACCATGA |
araM-F | TGTTCTGATGATGATGATGAT |
araM-R | CTTCTGGGGCGATGTGTCAT |
araN-F | TTGTCCTGAGGGATCTGGAAGA |
araN-R | GTTCTACGGTTCGAGTAGTGGGCGGA |
araO-F | GGTCGAATGAGTGTGCACGAT |
araO-R | GCGGGAACAGCTATGACCATGA |
araP-F | GCGATGGTGGTGCTGAT |
araP-R | GGTGCTGCGCTGCGTCTGCTC |
araQ-F | TCATTAGCCTGGGTGTGAAGGT |
araQ-R | GCGGGAACAGCTATGACCATGA |
araR-F | GCCGGATCCTGCTGCTGCT |
araR-R | TGTGTCGCGCTGCGTCTGCT |
araS-F | GCCGGATCCTGCTGCTGCT |
araS-R | TGTGTCGCGCTGCGTCTGCT |
araT-F | GCCGGATCCTGCTGCTGCT |
araT-R | TGTGTCGCGCTGCGTCTGCT |
araU-F | GCCGGATCCTGCTGCTGCT |
araU-R | TGTGTCGCGCTGCGTCTGCT |
araV-F | GCCGGATCCTGCTGCTGCT |
araV-R | TGTGTCGCGCTGCGTCTGCT |
araW-F | GCCGGATCCTGCTGCTGCT |
araW-R | TGTGTCGCGCTGCGTCTGCT |
araX-F | GCCGGATCCTGCTGCTGCT |
araX-R | TGTGTCGCGCTGCGTCTGCT |
araY-F | GCCGGATCCTGCTGCTGCT |
araY-R | TGTGTCGCGCTGCGTCTGCT |
araZ-F | GCCGGATCCTGCTGCTGCT |
araZ-R | TGTGTCGCGCTGCGTCTGCT |

**DNA footprinting analysis**

1-Arabinose-induced expression of the 1-arabinose utilization cluster genes. On the chromosome of *C. glutamicum* ATCC 31831, six genes, i.e., *araA*, *araB*, *araD*, *araE*, *galM*, and *araR*, are located in a cluster (Fig. 1) (26). *araA*, *araB*, and *araD*, each of which is essential for 1-arabinose utilization in this strain (26), encode arabinose isomerase, ribulokinase, and ribulose-5-phosphate 4-epimerase, respectively. An H⁺ symporter encoded by *araE* is involved in 1-arabinose uptake (8). *galM* and *araR* encode a putative aldose 1-epimerase and a LacI-type transcriptional regulator, respectively. The *araB* gene is oriented in the same direction as its two downstream genes, *araD* and *araE*. The *araB-araD-araE* intergenic regions are 13 and 35 bp long, respectively. Upstream of the *araB* gene, *galM* is divergently oriented. *araD* and *araE* are located downstream of *galM*, in the same direction. The *galM-araR* and *araR-araE* intergenic regions are 26 and 234 bp long, respectively. We previously reported that the expression levels of *araBDA* and *araE* in this 1-arabinose utilization gene cluster are dramatically increased in exponentially growing cells in minimal medium supplemented with 1-arabinose compared with that in cells grown on d-glucose (26). Here we examined changes in the expression of

\[^a\] The restriction site overhangs used in the cloning procedure are underlined in the sequences.
\[^b\] 5'-Phosphorylated primer.
\[^c\] 5'-IRD700-labeled primer.

---

**Journal of Bacteriology**

2244 | jb.asm.org | March 2021, 2021 by guest http://jb.asm.org/ Downloaded from
these genes within 15 min of addition of l-arabinose by using qRT-PCR. *C. glutamicum* ATCC 31831 wild-type cells were cultured in nutrient-rich A medium for 4 h, and exponentially growing cells were then supplemented with l-arabinose or d-glucose to a final concentration of 2% (wt/vol). The levels of *araB*, *araD*, and *araA* mRNAs increased ~20-fold within 5 min and then decreased to some extent in the subsequent 10 min (Fig. 2A to C). The expression levels of these mRNAs after 15 min of incubation with l-arabinose were nearly 10-fold higher than their levels before l-arabinose addition. Thus, it is likely that *araB*, *araD*, and *araA* are transcribed as a tricistronic mRNA. A very similar expression pattern in response to l-arabinose was observed for *araE* (Fig. 2D). The expression levels of *galM* and *araR* mRNAs increased ~4-fold within 5 min and then decreased to some extent in the subsequent 10 min (Fig. 2A to C). The levels of these mRNAs 15 min after l-arabinose addition were 2-fold higher than their levels before the addition of l-arabinose. These results indicate that *galM* and *araR* are upregulated in response to l-arabinose, similar to the other genes in this cluster, but to a much smaller extent. Therefore, it is likely that *araE* is transcribed under the control of its own promoter, independent of the probable upstream *galM-araR* operon. Addition of d-glucose in place of l-arabinose resulted in a decrease in the expression levels of these five genes during 15 min of cultivation (Fig. 2A to F), and the effects of glucose on l-arabinose-inducible gene expression were examined further as described below.

Exponentially growing cells cultured in nutrient-rich A medium were supplemented with l-arabinose and/or d-glucose at 2% (wt/vol) (each) and then incubated for 60 min. Addition of d-glucose resulted in at least a 2-fold decrease in the levels of *araA* and *araE* mRNAs within 30 min, but after 60 min of incubation, these mRNAs reverted to their levels before the addition of d-glucose (Fig. 3). The prominent upregulation of *araA* and *araE* observed within 30 min of l-arabinose supplementation was strongly suppressed by supplementation of d-glucose simultaneously. However, the levels of these mRNAs were increased within 60 min of incubation in the presence of both l-arabinose and d-glucose, to the same levels as those in the presence of only l-arabinose. Mean values and standard deviations for at least three independent cultures are shown.

Next, to confirm the involvement of AraR in l-arabinose-inducible gene expression, we compared the levels of *araB*, *araE*, and *galM* mRNAs in an *araR* deletion mutant strain (ΔaraR) with those in the wild-type strain after 5 min of incubation in the presence or absence of l-arabinose. In the absence of l-arabinose, the expression level of *araB* in the ΔaraR strain was 50-fold higher...
than that in the wild-type strain (Fig. 4A). The level of araB mRNA in the ΔaraR strain was even higher than the L-arabinose-induced level in the wild-type strain but decreased significantly in the presence of L-arabinose. The same results were observed for araE (Fig. 4B). Deletion of araR also resulted in an increase in the level of galM mRNA in the absence of L-arabinose, and the AraR-independent downregulation in response to L-arabinose was observed (Fig. 4C), but these changes in galM expression were much smaller than those in the case of araB and araE.

The AraR protein binds to the upstream regions of araB, galM, and araE. To confirm the ability of the AraR protein to bind to the upstream regions of the L-arabinose-inducible genes, EMSA

FIG 3 Effects of D-glucose on L-arabinose-dependent upregulation of araA (A) and araE (B). The C. glutamicum ATCC 31831 wild-type strain was grown in nutrient-rich A medium for 4 h, and exponentially growing cells were then supplemented with D-glucose (white bars), L-arabinose (black bars), or L-arabinose plus D-glucose (gray bars), to a final concentration of 2% (wt/vol) for each sugar. The mRNA levels in cells incubated with sugars for 0, 30, and 60 min were determined by qRT-PCR and are presented relative to the values obtained for the cells before addition of L-arabinose. Mean values and standard deviations for at least three independent cultures are shown.

FIG 4 Effects of araR disruption on expression of araB (A), araE (B), and galM (C). The C. glutamicum ATCC 31831 wild-type strain (WT) and the araR deletion mutant strain (ΔaraR) were grown in nutrient-rich A medium for 4 h, and exponentially growing cells were then supplemented with L-arabinose to a final concentration of 2% (wt/vol). The mRNA levels in cells incubated with L-arabinose for 0 min (−) or 5 min (+) were determined by qRT-PCR and are presented relative to the values obtained for the wild-type cells before addition of L-arabinose. Mean values and standard deviations for at least three independent cultures are shown.
was performed using the C. glutamicum ATCC 31831 AraR protein, which was expressed in E. coli and purified. DNA fragments of 381 bp (P_A; between positions −353 and +28 with respect to the araB translation start site), 400 bp (P_B; between positions −380 and +20 with respect to the galM translation start site), 394 bp (P_C; between positions −364 and +30 with respect to the araE translation start site), and 399 bp (P_D; between positions −388 and +11 with respect to the araR translation start site) were used as DNA probes (Fig. 1A). The AraR protein reduced the electrophoretic mobility of the araB upstream DNA fragment; the amount of the AraR-DNA complex increased as the concentration of the AraR protein increased (Fig. 5A). The EMSAs also revealed that the AraR protein bound to the galM and araE upstream DNA fragments (Fig. 5B and C). In contrast, no complex of AraR with the araR upstream DNA fragment was detected (Fig. 5D). These results indicated that AraR binds to the upstream regions of araB, galM, and araE in a sequence-specific manner.

There was a 124-bp overlap between the upstream regions of araB and galM used for the DNA probes, each of which formed a complex with AraR as detected as a single shifted band (Fig. 5A and B). We performed EMSAs using a deletion series of the galM upstream region as DNA probes (Fig. 6A). The AraR protein was incubated with the respective DNA fragments and subjected to gel electrophoresis. The results of these EMSAs are shown in Fig. 6B;
positive or no binding of AraR is indicated by $+$ or $-$, respectively, in Fig. 6A. The EMSAs revealed that the AraR protein bound to all these 5'-deletion fragments (P2, P3, P4, and P5 in Fig. 6). A shifted band corresponding to a complex with AraR was observed for two of the 3'-deletion fragments (P6 and P7 in Fig. 6), but not for the other two (P8 and P9 in Fig. 6). These results indicate that the AraR-binding site upstream of galM is located within the 50-bp region between positions −59 and −109 with respect to the translation start site of galM. This AraR-binding site is also included in the araB upstream DNA fragment (PΔ) used for the EMSAs described above.

AraR-binding activity is reduced by l-arabinose. The effects of l-arabinose on the DNA-binding activity of the AraR protein were examined by EMSA (Fig. 7). At a concentration of 5 mM, l-arabinose clearly reduced the formation of the complex of AraR with the araB upstream DNA fragment. In contrast, D-xylose did not affect the DNA-binding activity of AraR even at 15 mM. The binding of AraR to the araB upstream DNA fragment was inhibited as a function of the concentration of l-arabinose, from 1 mM to 15 mM. These results indicate that l-arabinose acts as a negative effector of the binding activity of AraR.

Identification of AraR-binding sites in the region upstream of araB and araE. The transcription start sites of the araB, araE, and galM genes determined by 5'-RACE were located 7, 104, and 69 bp upstream of the respective translation start sites (Fig. 1B and C). The consensus sequence of the −10 region of C. glutamicum SigA-dependent promoters (33) is found in their 5'-upstream regions. The consensus sequence of the −35 region is found in the araB and araE promoters but not in the galM promoter.

Next, we performed DNase I footprinting analysis to determine the AraR-binding site within the intergenic region between araB and galM (Fig. 8A). The DNA region protected against DNase I digestion by the binding of AraR was identified between positions +4 and +28 with respect to the transcription start site of araB. This protected region also corresponds to the region between positions −5 and −29 with respect to the transcription start site of galM, which is located divergently from that of araB (Fig. 1B and 8A). Therefore, the single AraR-binding site overlaps not only the translation start site of araB but also the −10 region of the divergent galM promoter. This AraR-binding site is included in the overlap between the araB and galM upstream DNA fragments used in the EMSAs described earlier (Fig. 5A and B).

DNase I footprinting analysis using the araE promoter region revealed that the AraR protein protects two regions: one between positions +5 and +28 and one between positions −28 and −46 with respect to the transcription start site of araE (Fig. 1C and 8B). This is consistent with the results of EMSA using the araE promoter region, which showed two shifted bands corresponding to complexes with AraR (Fig. 5C). These results indicate that one of the two AraR-binding sites corresponds to the 5'-untranslated region and the other overlaps the −35 region of the araE promoter (Fig. 1C).

We found the consensus sequence ATGTtAGCGlnTaAcat based on alignment of the three AraR-binding sites: one in the araB upstream region and two in the araE upstream region (Fig. 9A). This 16-bp consensus sequence has an imperfect inverted repeat of the 8-bp element. The palindromic sequence has some similarity to the operator sequences of the LacI/GalR family transcriptional regulators (34, 35). Effects of mutations in the putative recognition sequence in the araB upstream region on the binding of AraR were examined by EMSAs (Fig. 9B to H). Three consecutive base pairs at various positions within the 16-bp consensus sequence were mutated, and the resulting araB upstream fragments carrying the mutations (mut1 to mut6 in Fig. 9B) were used as DNA probes in EMSAs. The EMSAs revealed that the AraR protein bound to the mut1 and mut6 fragments (Fig. 9C and H) with the same affinity as that for the wild-type araB upstream fragment, as shown earlier (Fig. 5A). In contrast, lower binding affinities were observed for the mut2, mut3, and mut5 fragments (Fig. 9D, E, and G), and no binding of AraR to the mut4 fragment was detected at all (Fig. 9F). These results indicate that the CG nucleotides located in the center of the palindromic sites are essential for the binding of AraR and that the nucleotides located in the 5' half of the palindromes are more important for AraR binding than those of the 3' half, consistent with the conservation between the target sequences identified here.

l-Arabinose-dependent induction of AraR-mediated gene expression is independent of catabolism of l-arabinose. The role of l-arabinose catabolism in the l-arabinose-dependent induction of gene expression was examined by qRT-PCR, using the araA, araB, and araD deletion mutants (ΔaraA, ΔaraB, and ΔaraD strains, respectively). Previously, we reported that these deletion mutants cannot grow on l-arabinose as the sole carbon source (26). For qRT-PCR analysis, these strains were grown in nutrient-rich A medium, and then exponentially growing cells were supplemented with 2% (wt/vol) l-arabinose, as described earlier. In the absence of l-arabinose, the level of araB mRNA in the ΔaraA and ΔaraD strains was the same as that in the wild-type strain. Within 5 min of l-arabinose supplementation, araB mRNA was markedly upregulated in the ΔaraA and ΔaraD strains, to an extent comparable to that in the case of the wild-type strain (Fig. 10A). However, in the subsequent 10 min, the expression level of araB remained high in both of the deletion mutants, in contrast to the decreased level in the wild-type strain described earlier. Similar results were observed for araA expression in either the ΔaraB (Fig. 10A) or ΔaraD (data not shown) strain, and the same was observed for araD expression in either the ΔaraA or ΔaraB strain (data not shown). Furthermore, the expression of araE and araR was also upregulated within 5 min of l-arabinose supplementation in each of the ΔaraA, ΔaraB, and ΔaraD strains, as in the
wild-type strain (Fig. 10B and C). After induction, the expression levels of these genes in the mutant strains hardly decreased in the subsequent 10 min, in contrast to the case in the wild-type strain. These results indicate that catabolism of L-arabinose taken up by cells is not required for the L-arabinose-dependent induction of these genes under the control of AraR but is required for the subsequent downregulation observed for the wild-type strain.

**Upregulation of L-arabinose utilization genes is dependent on uptake of L-arabinose.** Previously, we reported that deletion of **araE**, encoding an L-arabinose transporter, causes a low growth rate and reduces the final cell density in the presence of a low concentration of L-arabinose (26). Effects of different concentrations of L-arabinose on **araB**, **araE**, and **araR** expression were examined using the **araE** deletion mutant strain (**H9004 araE**) and the wild-type strain. Exponentially growing cells cultured in nutrient-rich A medium were supplemented with L-arabinose to final concentrations of 0.5 mM, 5 mM, 50 mM, and 200 mM and subsequently incubated for 5 min (Fig. 11). qRT-PCR analysis revealed that before the supplementation of L-arabinose, there was no apparent difference in the expression levels of each of the **araB**, **araE**, and **araR** genes between the **H9004 araE** strain and the wild-type strain.

The expression of **araE** was detected using a pair of qRT-PCR primers corresponding to a region upstream of the deleted region in the **H9004 araE** strain. With the wild-type strain, lower concentrations of L-arabinose tended to be more effective for induction of the **araB** and **araE** genes. The expression of these genes was also markedly induced by L-arabinose at concentrations of 5 mM, 50 mM, and 200 mM in the **DeltaaraE** strain, to an extent comparable to that in the case of the wild-type strain. However, addition of 0.5 mM L-arabinose was much less effective for induction of the expression of **araB** and **araE** in the **H9004 araE** strain than in the wild-type strain. Similar effects of **araE** inactivation on L-arabinose-dependent induction were observed for **araR**, although its dose response to L-arabinose in the wild-type strain was minimal, in contrast to the case of **araB** and **araE** as described above (Fig. 11). These results indicate that the uptake of L-arabinose is required for the induction of these genes under the control of AraR.

**DISCUSSION**

In this study, we showed that the *C. glutamicum* ATCC 31831 AraR protein, belonging to the LacI family of transcriptional regulators, binds to the **araB** and **araE** upstream regions and that the *in vitro* DNA-binding activity of AraR is reduced by L-arabinose. We previously reported that the expression of **araBDA** and **araE** is upregulated in the presence of L-arabinose compared to that in the presence of D-glucose and by inactivation of **araR** in the absence of L-arabinose (26). These results indicate that AraR acts as a transcriptional repressor of the L-arabinose catabolic and uptake genes.
and araE) and that AraR-mediated expression is derepressed in the presence of l-arabinose. This is consistent with the rapid induction of these genes after supplementation with l-arabinose, as shown in this study. It should be noted that l-arabinose induced these genes under the control of AraR in mutant strains deficient in l-arabinose catabolism, as well as in the wild-type strain (Fig. 10). Furthermore, in the strain deficient in araE, encoding an l-arabinose transporter, the induction of these genes by a low concentration of l-arabinose was markedly suppressed (Fig. 11), suggesting that the uptake of l-arabinose is required for the derepression of AraR-mediated gene expression. It is noted that the inactivation of araE results in slow growth with a low concentration of l-arabinose as the sole carbon source (26). However, this mutant strain grows as well as the wild-type strain in the presence of a high concentration of l-arabinose (26), which is consistent with the high induction levels of the l-arabinose utilization genes shown in this study. Taken together, our data establish that l-arabinose acts as an efficient inducer of AraR-mediated gene expression in C. glutamicum ATCC 31831 in vivo.

As in C. glutamicum ATCC 31831, l-arabinose plays a key role in the regulation of l-arabinose utilization genes as a primary inducer in E. coli and B. subtilis (36–39). However, their l-arabinose-responsive transcriptional regulators belong to distinct families. The AraC protein in E. coli functions as an activator in the presence of l-arabinose but as a repressor in the absence of l-arabinose (13, 14, 16). In B. subtilis, the oligomerization and effector-binding domains, located in the C terminus of the l-arabinose-responsive transcriptional regulator AraR, have sequence similarity to the LacI/GalR family proteins, but the DNA-binding domain, located in the N terminus, has sequence similarity to members of the GntR family (21, 40). As shown in this study, the LacI-type transcriptional regulator AraR in C. glutamicum ATCC 31831 recognizes a 16-bp palindromic sequence which has common features of the recognition sequences described so far for the other members of this family (35). It has been reported that the LacI/GalR family members coordinate available nutrients with expression of catabolic genes, but some regulate processes as diverse as nucleotide biosynthesis and toxin expression (41–43). As established for this family of transcriptional regulators, it is likely that binding of l-arabinose to the C-terminal domain in the C. glutamicum ATCC 31831 AraR protein results in a conformational change in the N-terminal DNA-binding domain. Thereby, the DNA-binding activity is inhibited in the presence of l-arabinose. This is the first report of the l-arabinose-dependent changes in
DNA-binding activity of the LacI/GalR family regulators. The gene cluster associated with L-arabinose utilization has never been found in other strains of \textit{C. glutamicum} so far. However, \textit{Mycobacterium smegmatis} and \textit{Streptomyces venezuelae} have homologs of the L-arabinose catabolic genes of \textit{C. glutamicum} ATCC 31831, and LacI-type transcriptional regulators encoded near these genes on their chromosomes have 44% and 41% amino acid sequence identity, respectively, to \textit{C. glutamicum} ATCC 31831 AraR. In \textit{M. smegmatis}, which is phylogenetically related to \textit{C. glutamicum}, the L-arabinose catabolic genes are upregulated by L-arabinose (44, 45), but the involvement of the AraR homolog in this regulation remains to be investigated. Many other strains of actinobacteria identified to date lack the L-arabinose catabolic genes, but the
AraR-dependent regulatory system of L-arabinose utilization reported here might be conserved among other strains that have yet to be unidentified.

On the chromosome of C. glutamicum ATCC 31831, the araR gene is located immediately downstream of galM, encoding a putative aldose 1-epimerase. We showed here that araR and galM are similarly upregulated by L-arabinose. These results suggest that araR is cotranscribed with galM under the control of the L-arabinose-inducible galM promoter. Since the AraR-binding site is located within the intergenic region between arab and galM (Fig. 1), AraR may regulate the expression of the divergently transcribed araBDA and galM-araR genes through the same mechanism. However, L-arabinose supplementation and/or inactivation of araR showed much smaller effects on the expression of galM and araR than the effects on the expression of araBDA (Fig. 4). The AraR-binding site in the galM promoter overlaps its −10 region (Fig. 1) but is located downstream of the transcription start site of araB. The difference in the location of the AraR-binding site may explain the different degree of transcriptional repression. Furthermore, our previous study showed that the expression of araR, but not that of galM, is upregulated in cells grown in minimal medium with L-arabinose as the sole carbon source compared to that in cells grown on D-glucose (26). Thus, it is possible that the araR gene is under the control of its own promoter, although no binding of AraR to the araR upstream region was observed (Fig. 5). Further studies are needed to elucidate the role of AraR in galM promoter activity. The other regulatory mechanism independent of AraR should also be considered.

We found the 16-bp consensus palindromic sequence ATGTGCTAgTAc based on the alignment of the three binding sites of C. glutamicum ATCC 31831 AraR identified here. This motif is present at positions +8 and +23 bp with respect to the araB transcription start site and in two positions, one between positions +10 and +25 bp and the other between positions −29 and −44 bp, with respect to the araE transcription start site (Fig. 1). Tremendous efforts have been devoted to characterizing the LacI/ GalR transcriptional regulators so far, and the results indicate that a few bases in the center of the palindromic structure are strictly conserved among the members of this family (34, 35). However, the peripheral nucleotides in the palindrome are less conserved, which allows these regulators to specifically recognize their own target sequences. The consensus sequence of the multiple C. glutamicum ATCC 31831 AraR-binding sites has CG in the center of the palindrome, which is conserved among the LacI/GalR family members. As expected, the introduction of mutations in the center of the palindromic sequence resulted in no formation of a complex with AraR (Fig. 9F). The AraR-binding site is located downstream of the transcription start sites of arabinose and D-glucose, suggesting that binding of AraR prevents the elongation of transcription by RNA polymerase. It is noted that a loop structure formed by the cooperative binding of a transcriptional repressor to separate sites in its target promoter is one of the most efficient mechanisms to repress transcription, as described for the LacI family (e.g., LacI and GalR in E. coli) (46, 47). This was also reported for the L-arabinose utilization genes under the control of different types of L-arabinose-responsive transcriptional regulators, i.e., AraC in E. coli (12) and AraR in B. subtilis (19). It is interesting that the second distal binding site of C. glutamicum ATCC 31831 AraR overlaps the −35 region of the L-arabinose transporter gene araE promoter, with a 37-bp space from its proximal binding site (Fig. 1C). This distal site may act cooperatively with the proximal one, although binding of AraR to the former site may prevent RNA polymerase from the interaction with the promoter, irrespective of the latter site. In this context, it should be noted that no similar distal AraR-binding site was found in the upstream region of the L-arabinose catabolic genes araBDA.

We found that in the araR deletion mutant strain, the expression levels of araB and araE markedly decreased within 5 min of the addition of L-arabinose, although in the absence of L-arabinose these mRNAs in this mutant were significantly elevated compared with those in the wild-type strain (Fig. 4). This AraR-independent regulatory mechanism may be involved in the decline after the rapid induction of the araBDA operon and araE within 15 min of the addition of L-arabinose to the wild-type cells (Fig. 2), although fluctuation of the intracellular level of L-arabinose may also affect the AraR-dependent expression of these genes in the beginning of L-arabinose catabolism. In E. coli, the induction of the L-arabinose uptake and catabolic genes was followed by repression during 1 h of incubation with L-arabinose (16). It has been suggested that the cAMP-dependent transcriptional regulator CRP is responsible for the L-arabinose-dependent repression. The expression of the L-arabinose-inducible genes in E. coli is severely inhibited in the presence of D-glucose, through a carbon catabolite repression mechanism mediated by CRP (48–50). We found that the L-arabinose catabolic and uptake genes in C. glutamicum ATCC 31831 were transiently downregulated upon the addition of D-glucose (Fig. 3). Moreover, the addition of L-arabinose plus D-glucose resulted in a delay of the upregulation of L-arabinose-inducible gene expression (Fig. 3). However, the steady-state expression level upregulated in the presence of both L-arabinose and D-glucose was comparable to that in the presence of only L-arabinose. This is consistent with our previous study showing the simultaneous utilization of L-arabinose and D-glucose, at the same rates (26). It should be noted that the global carbon catabolite repression system, mediated by CRP and CcpA in E. coli and B. subtilis, respectively, has never been found in C. glutamicum. Moreover, C. glutamicum has the ability to consume various carbon sources simultaneously (51–54). At present, a search for another transcriptional regulator involved in the probable carbon source-responsive multilayered control of the L-arabinose utilization genes is under way in our laboratory. An understanding of this regulatory system of L-arabinose utilization genes in C. glutamicum ATCC 31831 should provide critical insights into the unique control of genes involved in sugar metabolism in this industrially useful microorganism and a basis for the development of efficient bioprocesses.

ACKNOWLEDGMENTS

We thank Crispinus A. Omumasaba (RITE) for critically reading the manuscript.

This work was financially supported in part by the New Energy and Industrial Technology Development Organization (NEDO), Japan.

REFERENCES


