

Dissolution of Xylose Metabolism in *Lactococcus lactis*

KARN A. ERLANDSON,[†] JOO-HEON PARK, WISSAM EL KHAL,[‡] HSIN-HSIN KAO,[§]
PERVIN BASARAN,^{||} SUSANNAH BRYDGES,[#] AND CARL A. BATT^{*}

Department of Food Science, Cornell University, Ithaca, New York 14853

Received 3 January 2000/Accepted 9 June 2000

Xylose metabolism, a variable phenotype in strains of *Lactococcus lactis*, was studied and evidence was obtained for the accumulation of mutations that inactivate the *xyl* operon. The xylose metabolism operon (*xylRAB*) was sequenced from three strains of lactococci. Fragments of 4.2, 4.2, and 5.4 kb that included the *xyl* locus were sequenced from *L. lactis* subsp. *lactis* B-4449 (formerly *Lactobacillus xylosus*), *L. lactis* subsp. *lactis* IO-1, and *L. lactis* subsp. *lactis* 210, respectively. The two environmental isolates, *L. lactis* B-4449 and *L. lactis* IO-1, produce active xylose isomerases and xylulokinases and can metabolize xylose. *L. lactis* 210, a dairy starter culture strain, has neither xylose isomerase nor xylulokinase activity and is Xyl⁻. Xylose isomerase and xylulokinase activities are induced by xylose and repressed by glucose in the two Xyl⁺ strains. Sequence comparisons revealed a number of point mutations in the *xylA*, *xylB*, and *xylR* genes in *L. lactis* 210, IO-1, and B-4449. None of these mutations, with the exception of a premature stop codon in *xylB*, are obviously lethal, since they lie outside of regions recognized as critical for activity. Nevertheless, either cumulatively or because of indirect effects on the structures of catalytic sites, these mutations render some strains of *L. lactis* unable to metabolize xylose.

Xylose metabolism has been described for a wide array of microorganisms. Extensively characterized xylose-metabolizing (Xyl⁺) bacteria include *Escherichia coli* (24, 46), *Lactobacillus* spp. (3, 4, 26–28), *Bacillus* spp. (16, 33, 38), and *Staphylococcus xylosus* (44). Free xylose can be transported via low-affinity symporters (XylE or XylT) or high-affinity binding-protein dependent systems (XylFGH) (1, 6, 9, 41). Xylose isomerase (*xylA* gene) then converts the aldose sugar to xylulose, which is phosphorylated by xylulokinase (*xylB* gene). Further metabolism of xylulose-5-phosphate occurs via the pentose-phosphate or phosphoketolase pathways. Xylose metabolism is induced by xylose, mediated via XylR. In *Salmonella* and *E. coli* strains, XylR is an activator when xylose is present (42, 46). In gram-positive organisms, XylR is a repressor which is inactivated when xylose binds (13, 26, 27, 37, 44, 45). The catabolite repression of xylose metabolism by glucose is mediated primarily through the CcpA protein (40).

Lactococcus lactis subsp. *cremoris* and almost all *L. lactis* subsp. *lactis* strains cannot metabolize xylose. These organisms have undergone intense selection for use in dairy fermentations, but plants are thought to be their original ecological niche because *L. lactis* subsp. *lactis* isolates have been recovered from many different plants (22, 34, 35). Also, the plant isolate *Lactobacillus xylosus* has been reclassified as *L. lactis* subsp. *lactis* (39). However, *L. lactis* subsp. *cremoris* strains are almost exclusively dairy associated (22).

We discovered xylose metabolic genes from both plant (Xyl⁺) and dairy (Xyl⁻) isolates of *L. lactis*. The sequencing of *xylRAB* from *L. lactis* strains IO-1, 210, and B444-9 is described here. The induction of xylose isomerase and xylulokinase activities from these three strains grown on different carbon sources was measured. Activity measurements in conjunction with the discovery of a *xylR* gene suggest that a positive regulatory mechanism is present in lactococcal strains. Mutations which could account for the lack of xylose isomerase and xylulokinase activity were observed in *L. lactis* subsp. *lactis* 210. Further downstream of *xylRAB* are genes encoding a mutarotase, xylanase, and xyloside transporter (K. A. Erlandson, S. Delamarre, and C. A. Batt, unpublished results), whose presence collectively suggests that ancestors of these strains were fully able to degrade and metabolize xylan.

MATERIALS AND METHODS

Bacterial strains, plasmids, and cultivation. Strains and vectors are described in the Table 1. *L. xylosus* was reclassified as *L. lactis* ssp. *lactis* based on immunological characteristics, 16S sequences, and fatty acid composition (39) and is referred to as *L. lactis* B-4449 here. All strains were routinely cultivated at 30°C in M17 medium (Difco, Detroit, Mich.) with 0.5% glucose or xylose. *E. coli* strains were cultured with aeration at 37°C in Luria broth (LB; Sigma, St. Louis, Mo.), with 100 µg of ampicillin (Sigma) or 50 µg of carbenicillin (Sigma) per ml where appropriate.

Sequencing strategies. A combination of strategies was used to sequence the *xyl* operon genes (Table 2). First, the *L. lactis* B-4449 xylose isomerase gene was sequenced in two fragments. Degenerate *xylA* PCR primers (NH-1 and WG-1) were designed based on the published N-terminal sequence of the *L. xylosus* XylA (48) and an amino acid sequence (WGGREG, amino acids [aa] 188 to 193) conserved among bacteria including *E. coli*, *Salmonella enterica* serovar Typhimurium, *Lactobacillus pentosus*, and *Bacillus* spp. A cloned portion of the resulting 540-bp fragment was hybridized to a 2.4-kb *EcoRI/HindIII* insert in a chromosomal minilibrary. A portion of this fragment encoded the first 311 aa of XylA. The sequences of the 3' end of the *xylA* gene and the 5' end of *xylB* were obtained via inverse PCR with primers 7 and 5.

PCR primers (BFX and HRX) designed from the *L. lactis* B444-9 *xylA* and 5' *xylB* sequences allowed us to amplify the corresponding 1.5-kb product from *L. lactis* strains IO-1 and 210. The PCR product was sequenced after ligation into pUC19 and transformation into *E. coli* JM109. Three inverse PCR reactions with primer pairs BFX-rev-HRX-rev, 210ip820-rev-R210xb-rev, and ip412-rev-R210xb-rev (Table 2) were used to walk upstream of *xylA* to obtain the *xylR* nucleotide sequence and to walk downstream of *xylA* to obtain the *xylB* sequence from *L. lactis* 210. PCR primers were designed from the 210 sequences to determine the *xylR* and *xylB* sequences from *L. lactis* IO-1 and *L. xylosus*.

* Corresponding author. Mailing address: Cornell University, Department of Food Science, 312 Stocking Hall, Ithaca, NY 14853. Phone: (607) 255-2896. Fax: (607) 255-8741. E-mail: cab10@cornell.edu.

[†] Present address: Procter and Gamble Co., Cincinnati, Ohio.

[‡] Present address: Nestle S.A., Paris, France.

[§] Present address: Department of Environmental and Occupational Health, National Cheng Kung University Medical College, Tainan, Taiwan, Republic of China.

^{||} Present address: Department of Food Science and Technology, Cornell University, Geneva, N.Y.

[#] Present address: Department of Microbiology and Immunology, Johns Hopkins University, Baltimore, Md.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Genetic markers	Source
Strains		
<i>L. lactis</i> subsp. <i>lactis</i> 210	Xyl ⁻	J. Kondo (Marschall Products)
<i>L. lactis</i> subsp. <i>lactis</i> IO-1	Xyl ⁺	P. Stansbury (University of Hertfordshire, Hertfordshire, United Kingdom)
<i>L. lactis</i> subsp. <i>lactis</i> NRRL B-4449 (formerly <i>L. xylosum</i>)	Xyl ⁺	A. J. Sinsky (Massachusetts Institute of Technology)
<i>E. coli</i> TG1	<i>supE hsdD5 thi Δ(lac⁻ proAB)</i> (F' <i>traD36 proAB lacI^q lacZ ΔM15</i>)	
<i>E. coli</i> JM109	<i>el4⁻ recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 Δ(lac⁻ proAB)</i> (F' <i>traD36 proAB lacI^q lacZΔM15</i>)	
<i>E. coli</i> DH5αF'	F' <i>endA1 hsdR17 supE44 thi-1 recA1 gyrA relA1 Δ(lacIZYA-argF)</i> U169 <i>deoR</i> [80 <i>dlacΔ(lacZ)M15</i>]	
Plasmids		
pUC19	Ap ^r ; <i>lacZ'</i>	New England Biolabs (Beverly, Mass.)
pGEM-T	Ap ^r ; <i>lacZ'</i>	Promega (Madison, Wis.)
pET19d	Ap ^r	Novagen (Madison, Wisc.)

Inverse PCR. Chromosomal DNA was extracted from lactococci according to the method described by Kim and Batt (21). The chromosomal DNA was digested using a variety of restriction enzymes (New England Biolabs, Beverly, Mass.). The enzyme reactions were inactivated by incubation at 65°C for 15 min or by phenol extraction. Ligation reactions contained restricted DNA at a series of concentrations from 25 to 750 ng, together with 1 U of T4 DNA ligase (Gibco, Gaithersburg, Md.), 10 μl of 5× ligation buffer (Gibco), and sterile distilled H₂O, for a final reaction volume of 50 μl. Ligations were incubated for 1 h at 37°C and heat inactivated at 65°C for 20 min. The resulting self-ligated DNA was then precipitated with NaCl and ethanol and then resuspended in sterile distilled H₂O to give a final concentration of 5 ng/μl for use in PCR.

PCR amplification. PCR's were carried out in 1× PCR buffer (20 mM Tris-HCl [pH 8.3], 50 mM KCl) with a final volume of 50 μl. Reactions contained 1 μl of self-ligated chromosomal DNA or crude cell lysate prepared according to the method of Czajka and Batt (7), 50 pmol of each primer, 100 mM concentrations of each deoxynucleotide triphosphate, 1 U of AmpliTaq DNA Polymerase (Perkin-Elmer, Foster City, Calif.), and 1.5 mM MgCl₂. A Hybaid TR1 or Perkin-Elmer 2400 thermocycler was programmed with a 4-min hold at 94°C, followed by 30 cycles of 1 min at 94°C, 1 min at 55 to 65°C (depending on the primer *T_m*), and 1 min at 72°C, followed in turn by a 10-min hold at 72°C. For inverse PCR the extension time at 72°C was increased to 2 min.

Cloning. The *xyIA* genes from the three strains were cloned prior to sequencing. Standard (pUC19) and T-overhang (pGEM-T) vectors were used with a variety of host strains that allow blue-white screening, as indicated in Table 1. Ligations were performed according to the methods of Promega (Madison, Wis.) for pGEM-T or Sambrook et al. (36). Standard methods were used to prepare and transform competent cells (36).

Sequencing and sequence analysis. The *L. lactis* B444-9 *xyIA* gene was sequenced manually according to the double-stranded DNA method of Wang (49). Sequenase (v. 2.0; Amersham, Piscataway, N.J.), [α -³²P]dATP, and dideoxynucleotides were used, together with the universal M13 primers (-40 and reverse) or internal primers. The other genes were sequenced by the Cornell University BioResource Center using an ABI Prism 373A Stretch automated sequencer. Several *xyIA* PCR clones from each strain were analyzed to confirm the sequences. All other sequences were obtained directly from PCR products. Sequences were analyzed using several programs from the Lasergene software suite (EditSeq v. 3.12, SeqMan v. 3.53, and MegAlign v. 3.05; DNASTAR, Inc., Madison, Wis.). We also performed BLASTX searches (2) of the GenBank database with our nucleotide sequences to identify homologous genes.

Xylose isomerase assays. The cysteine carbazole method was used to measure xylose isomerase activity as indicated by the appearance of xylulose (10). Reagents were purchased from Sigma. Stationary phase M17 xylose- or glucose-grown cultures (100 ml) were harvested, resuspended in 4 ml of 50 mM Tris-HCl (pH 7.5) containing 1 mM dithiothreitol, and lysed by sonication. Sample absorbances were measured at 500 nm with an LKB Biochrom Ultraspec II spectrophotometer (Biochrom, Cambridge, United Kingdom) after 20 min at room temperature. One unit of enzyme activity is expressed as 1 nmol of D-xylulose produced per min.

Xylulokinase assays. Xylulose formation was measured using a method modified from Shamana and Sanderson (43) that couples xylulokinase to pyruvate kinase and L-lactate dehydrogenase. Cell lysates (containing 80 mg of protein) were prepared by sonication as described above. The oxidation of NADH was monitored by measuring the absorbance at 340 nm with a Beckman DU 640

TABLE 2. Summary of primers

Primer	Sequence (5'-3') ^a	Location (nucleotides)	Comment
WG-1	CCYTCRCGRCCGCCCA	1913-1897 ^b	<i>L. lactis</i> B-4449 <i>xyIA</i> PCR
NH-1	ACNAARACNAARAAYATGTTY	1378-1399 ^b	<i>L. lactis</i> B-4449 <i>xyIA</i> PCR
5	TGATGCCAATCTTGG	2220-2234 ^b	<i>L. lactis</i> B-4449 <i>xyIA</i> inverse PCR
7	TCTGGTAGGTCTTCTCCTA	~1.6 kg upstream of <i>xyIA</i>	<i>L. lactis</i> B-4449 <i>xyIA</i> inverse PCR
BFX	CGGGATCCATGTTTCATTCTCTCTAATCAC	1171-1192 ^c	<i>xyIA</i> PCR and Northern primer
HRX	CCCAAGCTTGGTTCCTAAGT	2702-2691 ^c	<i>xyIA</i> PCR and Northern primer
BFX-rev	GTGATTAGGAGAGAATGAACAT	1192-1171 ^c	<i>xyIR-xyIB</i> inverse PCR
HRX-rev	TTAGGAATTGACTTAGGAACC	2681-2702 ^c	<i>xyIR-xyIB</i> inverse PCR
210ip820-rev	TGAGTAACAAGCACCCGGAACC	3544-3523 ^c	<i>xyIB</i> inverse PCR
R210xb-rev	TAGAAGCGGTCACTAAAGC	4011-4029 ^c	<i>xyIB</i> inverse PCR
Ip412-rev	AGAAATGCTGGATAAACTGG	2964-2945 ^c	<i>xyIB</i> inverse PCR
M13-40	GTTTTCCAGTCAAGAC	NA ^d	Vector sequencing primer
M13rev(ht)	GCTTTAGTGACCGCTTCTA	NA	Vector sequencing primer

^a R = A or G; Y = C or T; H = A or C or T; N = any nucleotide.

^b Numbering corresponds to *L. lactis* B-4449 *xyI-xyN* operon (including sequence upstream of *xyIR*); GenBank accession no. AF092042.

^c Numbering corresponds to *L. lactis* IO-1 *xyI-xyN* operon (including upstream sequence); GenBank accession no. AF092041.

^d NA, not available.

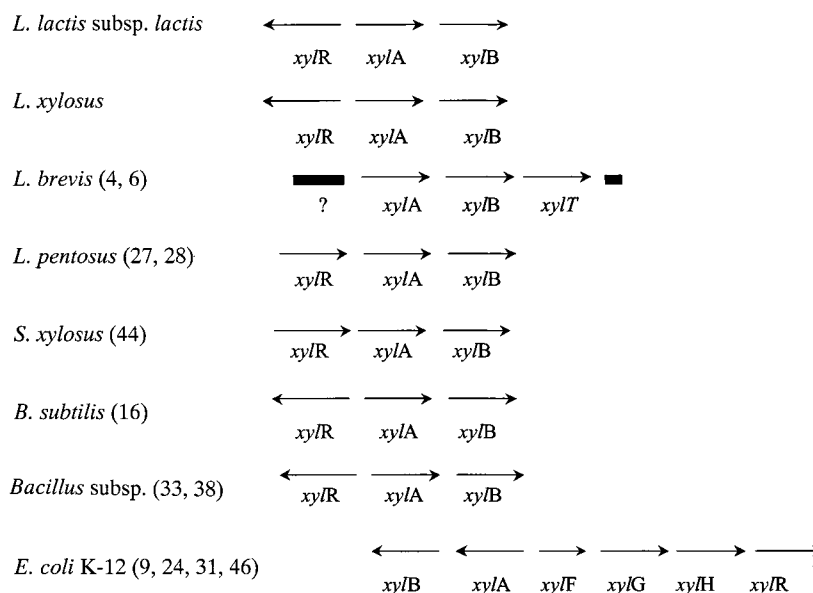


FIG. 1. Organization of the *xyl* locus in various bacteria. The *xyIR* gene encodes a regulatory protein, *xyIA* encodes a xylose isomerase, *xyIB* encodes a xylulose kinase; *xyIF* encodes a xylose-binding protein, *xyIT* encodes a xylose transporter, *xyIG* encodes an ATP-binding protein, and *xyIH* encodes a membrane transporter (ABC-type xylose transport). ?, Sequenced open reading frame of unknown function. References are given in parentheses.

spectrophotometer (Beckman Instruments, Fullerton, Calif.). One unit of enzyme activity was expressed as 1 nmol of D-xylulose consumed per min.

Nucleotide sequence accession numbers. The GenBank accession numbers for the *L. lactis* B444-9, *L. lactis* IO-1, and *L. lactis* 210 *xylRAB* genes are AF092042, AF092041, and AF092040, respectively.

RESULTS

Three xylose metabolism genes (*xyIRAB*) were sequenced from three lactic acid bacteria, including one Xyl⁻ strain. A combination of PCR products generated with degenerate conserved, homologous, and inverse primers was assembled to give a total of 4.2 kb of sequence from *L. lactis* subsp. *lactis* IO-1 and *L. lactis* subsp. *lactis* B444-9 and 5.4 kb of sequence from *L. lactis* subsp. *lactis* 210 (Fig. 1).

Xylose isomerase. Xylose isomerase is encoded by a 1,320-bp gene in *L. lactis* strains IO-1, 210, and B444-9. Figure 2 shows the IO-1 XylA sequence with conserved and similar residues indicated based on an alignment of 22 group II xylose isomerase sequences, including our three sequences. These xylose isomerases are grouped together because of their catalytic metal preference and size, which ranges from 438 to 465 aa (25, 48, 50). A total of 96 aa (21%) are identical for all 22 bacteria. An additional 118 residues (26%) are similar or identical for 90% of the strains. All of the group II XylAs conserved 17 active-site residues (Fig. 2) determined for the group I *Streptomyces rubiginosus* XylA (50–52). Although there are conserved residues throughout the group II protein, the most highly conserved region, aa 179 to 345, is in the middle of the protein. This region includes 12 *S. rubiginosus* active-site residues. The carboxyl terminus of XylA is the least conserved among the group II proteins. The XylA sequences of *L. lactis* B444-9, *L. lactis* IO-1, and *L. lactis* 210 are very similar (Fig. 2). Only 2 aa differ between the Xyl⁺ strains *L. lactis* IO-1 and *L. lactis* B444-9 (Lys407→Glu and His416→Tyr). These residues are not conserved among the other XylAs. Strain 210 differs from IO-1 at six residues: Met202→Arg, Asp218→Tyr, Ser247→Ala, Ala275→Val, Ser388→Thr, and Lys407→Glu. Three of the six residues (i.e., residues 202, 218, and 247) are

similar in the other group II XylAs, and residue 275 is conserved in 20 of the 22.

Xylose isomerase activity. Xylose isomerase activities for crude extracts of *L. lactis* strains IO-1, B444-9, and 210 were measured as described in Materials and Methods. The specific activity for IO-1 cells grown on xylose was 464 U/mg of protein, a value almost fourfold greater than that for B444-9 (127 U/mg of protein). Glucose-grown IO-1 and B444-9 xylose isomerase activities averaged 12.8 U/mg of protein. *L. lactis* 210 had little xylose isomerase activity whether grown on xylose (<6 U/mg of protein) or glucose (<12 U/mg of protein).

Xylulokinase. *L. lactis* strains IO-1 and B444-9 have a 1,506-bp *xyIB* gene located 66 bp downstream of the *xyIA* translation stop codon. The *L. lactis* 210 *xyIB* gene is 67 bp downstream of *xyIA*. Xylulokinase sequences, which range in size from 475 to 511 aa, are available from 14 of the 22 XylA group II organisms. Figure 3 displays the IO-1 XylB sequence with conserved and similar residues indicated, based on an alignment of the 14 sequences. The XylB sequences are less conserved than the XylA sequences, because only 45 residues (10%) are identical for all 14 strains. An additional 62 aa are conserved and 51 aa are similar for 85% of the strains. Similar and identical residues are found throughout the XylB protein, but homologous regions with clusters of conserved residues are small and concentrated in the N-terminal half of the protein.

The *L. lactis* strain IO-1 and B-4449 XylB sequences, like their XylA sequences, are very similar. However, the 210 *xyIB* sequence includes a premature translation stop codon corresponding to residue 154 (indicated in Fig. 3). The stop codon is immediately after an apparent -1 frameshift in a run of six (instead of seven) adenine residues. Resequencing this region confirmed the frameshift. Premature termination of translation is the likely explanation for the lack of significant xylulokinase activity of strain 210 (see below). If the stop codon were ignored, only seven residues (aa 200 to 202, 230, 290, 448, and 496) differ between *L. lactis* strains IO-1, B444-9, and 210 (Fig. 3). None of these differences correspond to conserved XylB residues. Aside from the stop codon, strain 210 differs

MAYFN[↓]DIAP[↓]IKYEGTK[↓]KNMFAFRHYNPEEVVAGKTMEEQLHFALAF[↓]WHT - 50
 ITMDGSD[↓]PF[↓]GGATMER[↓]PWDLEGGSELDRARRVD[↓]AFFEIAEKLGVKYYCF - 100
 HD[↓]IDIAP[↓]TGNSLKEFYANLDEITDHLLEKQKATG[↓]IKLLWNTANMFSNPRY - 150
 MNGVSTSNRAEVFA[↓]YGA[↓]AQVKKGLELSK[↓]LLGGENYV[↓]VFGGREGYESLLNT - 200
 DMGLEMDHMAK^{*}FFHLAIDYAKS^{*}INHLPIFLIE[↓]PK[↓]PE[↓]PMTHQYDFDSATA - 250
 LAF[↓]LQKYDL[↓]DKYFK[↓]LNLE[↓]TN[↓]HAWLAGHTFEHELNTARTFPNAL[↓]GSIDANQG - 300
 NYLL[↓]GWD[↓]TDEFPTLVIDITLAM[↓]HQILLNGGLGK[↓]GGINFD[↓]AKVRR[↓]TSFKAE - 350
 DLILA^{*}HIAGMDTYARALKGAAAI^{*}IEDKFLSDIVDERYSSYKNTEV^{*}QGSIE - 400
 NGTAT^{*}FKSLAAFA^{*}FEHGDDIELD^{*}SNHLEYIKSVLNDYL^{*}V. - 440

FIG. 2. *L. lactis* IO-1 XylA sequence with residues conserved among 22 group II XylA sequences indicated. The group II xylose isomerase sequences examined were from *L. lactis* IO-1 (this work), *L. lactis* 210 (this work), *L. lactis* B-4449 (this work), *L. brevis* (GenBank accession no. AF045552), *L. pentosus* (M57384), *S. xylosus* (X57599), *B. licheniformis* (Z80222), *B. megaterium* (Z71474), *B. subtilis* (U66480), *Bacillus* sp. strain LW2 (L12967), *Thermotoga neapolitana* (L38994), *Thermoanaerobacter* spp. strain (U21678), *Thermoanaerobacter* *thermosulfurogenus* (J05650), *Thermoanaerobacter* *saccharolyticum* (L09699), *Thermoanaerobacter* *thermosaccharolyticum* (M91248), *Thermoanaerobacter* *thermohydrosulfuricus* (D00756), *Thermoanaerobacter* *ethanolicus* (AF001974), *E. coli* (X04691), *Klebsiella pneumoniae* (X61059), *Klebsiella aerogenes* (Swiss-Prot accession no. P29442), and *H. influenzae* (GenBank accession no. U32791). Conserved residues are indicated as follows: bold underlined residues, identical for all 22 strains; underlined residues, identical or similar for 20/22 strains; * (above sequence), difference between *L. lactis* IO-1, 210, and/or B-4449; ↓, conserved *S. rubiginosus* active site residue, according to Whitaker (50).

from the Xyl⁺ strains at residue 290 (cysteine instead of tyrosine).

Xylulokinase activity. Xylulokinase activities for *L. lactis* strains IO-1, 210, and B444-9 were determined using a coupled enzyme assay as described in Materials and Methods. As with the xylose isomerase activities, the xylulokinase activity of IO-1 surpassed that of B444-9. The magnitude of the difference was much less, however, with 53.8 U/mg of protein for IO-1 compared to 41.4 U/mg of protein for B444-9. The enzyme activities for glucose-grown strains IO-1 or B-4449 were two to three times lower than when grown on xylose (24.9 and 12.4 U/mg, respectively). *L. lactis* 210 had a very low level of xylulokinase activity whether grown on xylose (7.8 U/mg) or on glucose (6.8 U/mg).

Xylose regulator. In all three strains, the 1,280 bp of sequence upstream of *xylA* included a 950-bp open reading frame, divergently oriented and 104 bp upstream from *xylA*, which we have termed *xylR*. The XylR protein is not homologous to XylRs from other gram-positive bacteria such as *Bacillus megaterium* and *L. pentosus*. Instead, it is moderately homologous to proteins such as RhaS (21% identical and 40% similar over 250 aa), the *E. coli* rhamnose metabolism regulator in the AraC/XylS family. This family includes other carbohydrate operon regulators such as *melR* (from *E. coli*), *rafR* (*Pediococcus pentosaceus*), and *xylR* (*E. coli*) which regulate melibiose, raffinose, and xylose metabolism, respectively. Gallegos et al. (12) have recently redefined this family to include over 100 proteins, with a C-terminal 99-aa stretch of homology constituting the DNA-binding domain of the regulator pro-

MTYV[↓]VLGIDLGTSS[↓]LK[↓]GILMDEVGNLITTKSAEYQIDTPKQGYSEQRPEY[↓]W - 50
 IVALESVLTGLSVEISDFGQQLAGIS[↓]FSGQM[↓]HSLV[↓]LVDDNNKPVY[↓]PAILW - 100
 NDVRTSKQCQEITDRLGQRLLLEITKNIALE[↓]GF[↓]TL[↓]PK[↓]LILW[↓]LQENE[↓]EPEVWSR - 150
 VKKIML[↓]PK[↓]DYLSLWLTGNIY[↓]TEFSDAAG[↓]TLLLDIEKKQWSEEITDAFNIE - 200
 RKIL[↓]PELIE[↓]STDR[↓]TGFV[↓]KABIAERYKLTNKVKV[↓]FAGGADNAAALGVGLI - 250
 NEEVGLIS[↓]MGTS[↓]GVVSAYEPK[↓]IA[↓]DYK[↓]GK[↓]LHFFN[↓]H[↓]TVFGAYYS[↓]MGV[↓]T[↓]LAAG - 300
 NSLN[↓]WYK[↓]ETFGKLSFNELLSEVYTVSPGSEGLLFT[↓]PYI[↓]V[↓]GERT[↓]PH[↓]FDSK - 350
 IRG[↓]SFIG[↓]ISAH[↓]HEQKHFS[↓]RAVL[↓]HGIT[↓]FSLR[↓]DSK[↓]DIMEKTKNKKFKRLISV - 400
 GGAQNP[↓]DIMQ[↓]QADIFNSEMIRLTVEQGPGLGAC[↓]MI[↓]AAFGCGLFDSFEA - 450
 VTKAFVHYKEAS[↓]FIP[↓]PNPK[↓]NVARYEQIYQIWKQY[↓]YKNTSE[↓]ISHQLVEFNDE - 500

G.

FIG. 3. *L. lactis* IO-1 XylB sequence with residues conserved among 14 XylB sequences indicated. The xylulokinase sequences examined were from *L. lactis* IO-1 (this work), *L. lactis* 210 (this work), *L. lactis* B-4449 (this work), *L. brevis* (GenBank accession no. AF045552), *L. pentosus* (M57384), *B. licheniformis* (Z80222), *B. subtilis* (U66480), *B. megaterium* (Z71474), *S. xylosus* (X57599), *T. ethanolicus* (AF001974), *E. coli* (X04691), *K. pneumoniae* (X61059), *K. aerogenes* (Swiss-Prot accession no. P29444), and *H. influenzae* (GenBank accession no. U32791). Conserved residues are indicated as follows: bold underlined residues, identical for all 14 strains; underlined residues, identical or similar for 12/14 strains; * (above sequence), difference between *L. lactis* IO-1, 210, and/or B-4449; X (underneath sequence), premature 210 stop codon. The FGGY kinase signature sequences (PROSITE patterns PS00933 and PS00445) are boxed.

teins. This region contains two helix-turn-helix (HTH) motifs. Although our XylR is only moderately homologous to individual AraC family proteins, a PFSCAN search (15, 29) of the PROSITE profile database revealed that it was significantly homologous to the 99-aa AraC family consensus region (Fig. 4). The IO-1 XylR sequence matched 37 of the 99 residues and was similar to 17 residues. Gallegos et al. identified 17 residues in the 99-aa stretch that are highly conserved (shared by more than 60% of the family member sequences). Thirteen of these residues are conserved by the IO-1 XylR.

Eight residues differ among the three lactococcus or lactobacillus strains. *L. lactis* 210 has a Asn106→Ser substitution outside of the AraC family consensus region and an Phe269→Ile substitution very near the C terminus of the consensus region. *L. lactis* B-4449 has an Asp147→His substitution. Also, due to an additional adenosine at the end of a run of six adenosines, the 5 aa at the C terminus of XylR are not conserved in the *L. lactis* 210. The strain 210 *xylR* gene encodes 2 aa after the frameshift before a premature stop codon (indicated in Fig. 4) results in a protein that is 3 aa shorter than that of *L. lactis* strains IO-1 and B-4449.

DISCUSSION

Xylose metabolism has been described and characterized in lactic acid bacteria, but it is rare in *L. lactis*. We have discovered a metabolic pathway for the metabolism of xylose to xylulose-5-phosphate in *L. lactis*. Xylose metabolism requires

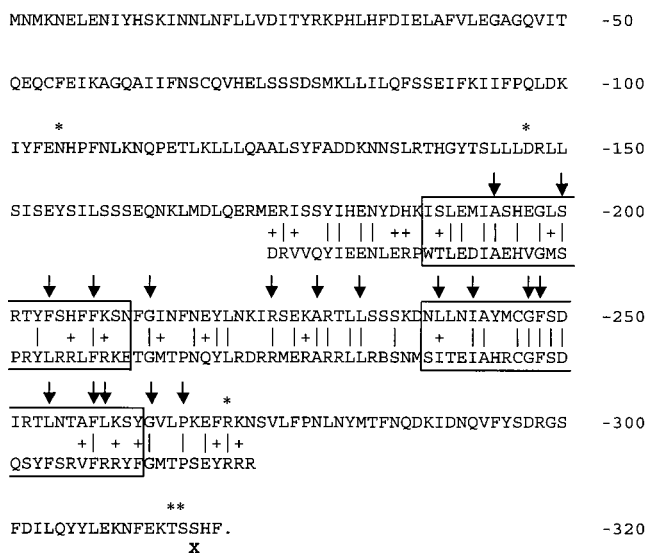


FIG. 4. *L. lactis* IO-1 XylR sequence comparison with HTH-AraC family 2 consensus (PROSITE profile PS01124). Conserved residues are indicated as follows: |, identical to HTH-AraC family 2 consensus; +, similar to HTH-AraC Family 2 consensus; ↓, residues conserved in >60% of AraC/XylS family (see reference 12 for details); X, (underneath sequence), premature 210 stop codon. Boxed residues are HTH motifs 1 and 2 (see the text). Residues which differ between *L. lactis* strains IO-1, 210, or B-4449 are indicated by an asterisk above the sequence.

the products of the *xylR*, *xylA*, and *xylB* genes. Although the three strains we studied possess sequences for these *xyl* genes, xylose metabolism in *L. lactis* is a function of the source of the strains. *L. lactis* subsp. *lactis* 210 is a dairy starter culture strain and cannot metabolize xylose, whereas *L. lactis* subsp. *lactis* B-4449 and *L. lactis* subsp. *lactis* IO-1 (which were isolated from a vegetation sample and a kitchen drain in Japan, respectively) can metabolize xylose. Recently, other Xyl⁺ *L. lactis* strains have been isolated from minimally processed fruit and vegetable products, which further supports the link between functional xylose metabolism and habitat (20).

Other than their differences in habitat, *L. lactis* B-4449, *L. lactis* IO-1, and *L. lactis* 210 are closely related. The ribotype patterns of *L. lactis* strains IO-1 and 210 have similarity values to the *L. lactis* B-4449 pattern of 0.69 and 0.72, respectively (data not shown). The L-lactate dehydrogenase sequences of the three strains also differ by less than 1% compared to 12 and 17% dissimilarities to *Streptococcus thermophilus* and *Lactobacillus casei*, respectively (11).

Although the xylose isomerases of our two Xyl⁺ strains differed by only 2 aa, the specific activity of IO-1 was almost fourfold greater than that of B-4449 when both strains were grown on xylose. The two residues are not conserved among the other XylAs nor are they in obvious proximity to any active-site residues. Northern analysis indicates that the *xylA* gene is transcribed at similar levels in both IO-1 and B-4449 (data not shown). Also, although the *xylA* gene is transcribed, *L. lactis* 210 did not have significant xylose isomerase activity. Its sequence differs by only 6 aa from IO-1. Three of these residues, R202, Y218, and V275, are located in highly conserved regions and differ from the group II XylA consensus (L/M202, D/E218, and A/S275; Fig. 2). The nucleotide sequence of the *xylR-xylA* intergenic region is identical in strains IO-1, 210, and B-4449, so it is more likely that differences in activity among the three strains are due to differences in protein structure rather than transcription levels. However, all of

the conserved xylose isomerase active-site residues identified (50) are intact in *L. lactis* IO-1, *L. lactis* 210, and *L. lactis* B-4449. Site-directed mutagenesis studies are in progress in our laboratory to characterize the functional importance of the differences between the *L. lactis* strain IO-1, 210, and B-4449 XylAs.

Differences in xylulokinase activities were also measured among *L. lactis* strains IO-1, 210, and B-4449. The xylulokinase activity of *L. lactis* 210 is very low regardless of the carbon source (7.8 U/mg of protein when xylose grown versus 6.8 U/mg when glucose grown). Although the strain 210 *xylB* gene is transcribed (Erlandson et al., unpublished), a premature translation stop codon (corresponding to residue 154) would almost certainly prevent formation of active xylulokinase.

As with xylose isomerase, xylose-grown IO-1 had more xylulokinase activity than B-4449. Their XylB sequences differed at six residues, none of which are conserved among the compared XylBs. XylB is part of a large, extensively characterized ATPase superfamily of proteins which includes other sugar kinases, eukaryotic heat shock and stress-70 proteins, and actin (17, 18). There are 127 structurally equivalent residues shared among the family members which fold into two symmetrical domains (each with a $\beta\beta\beta\alpha\beta\alpha$ topology) to form an ATP- and substrate-binding cleft (17). Using the algorithm of Rost and Sander (32), the secondary structure of XylB was predicted (data not shown). The sequence differences of IO-1 versus B-4449 did not map to a β -sheet or α -helix. However, residue 230 (lysine in IO-1, glutamate in B-4449) is 10 residues from Asp239 (Asp245 of *E. coli* glycerol kinase), a key residue conserved throughout the superfamily that is thought to act as a general catalytic base (30). Residue 230 is also 3 aa from a sequence block thought to be involved in hinge motion during protein conformational change (5). In addition, residue 448 (phenylalanine in IO-1, leucine in *L. xylosus*) is located six residues after a second sequence block putatively involved in hinge motion (5). Therefore, differences in the predicted three-dimensional structure of the IO-1 XylB could account for its greater activity relative to B-4449.

Xylulokinase and xylose isomerase activities decreased substantially when *L. lactis* strains IO-1 and B-4449 were grown on glucose. Catabolite repression by glucose is a phenomenon common to all xylose metabolizers (13, 14, 31, 33). We could not find *cis*-acting *cre* sites upstream of *xylA* or *xylB*. However, it is likely that lactococci have a homologue to the *cre* sequences identified in *Bacillus* spp., *L. pentosus*, and *S. xylosus* that bound the catabolite regulator CcpA (19, 26, 40, 45).

The XylR sequences of *L. lactis* strains IO-1, 210, and B-4449 are moderately homologous to AraC/XylS family proteins, including the *E. coli* and *Haemophilus influenzae* XylR activator proteins. The AraC/XylS family profile (PROSITE entry PS01124) consists of 99 aa, usually toward the C-terminal end of the protein, encompassing two HTH motifs. The region is thought to constitute the DNA-binding domain of the proteins. Our XylR is homologous over the entire AraC family profile, corresponding to aa 174 to 272 of the IO-1 sequence (Fig. 4). The *xylR* genes from all other gram-positive bacteria (*B. subtilis*, *B. megaterium*, *B. licheniformis*, *L. pentosus*, and *S. xylosus*) fall into a different family of transcriptional regulators and sugar kinases, the ROK family (47). The repressor proteins in the ROK family (PROSITE entry PS01125) have an HTH motif in the N-terminal portion of the protein and a 300-residue domain with a central glycine-rich region, although the rest of the consensus pattern is not highly conserved. The ROK XylR protein has been shown to bind a tandem repeat, designated *xylO*, near the *xylA* transcription start site (8, 13, 23, 27, 37, 45). This *xylO* site does not appear to be present up-

stream of *xylA* in our three strains. It is unusual that our XylR is homologous to the AraC family proteins because all of these proteins are activators, except the *E. coli* cellbiose repressor (CelD) and the bifunctional *E. coli* AraC arabinose and *Yersinia pestis* YbtA pesticin and yersiniabactin regulators (which activate or repress depending on the effectors). In general, xylose regulation is positive in the case of gram-negative organisms but negative in gram-positive organisms (16, 23, 42, 46). Our sequence homology suggests a positive regulatory mechanism different from that previously described for gram-positive xylose metabolism.

What will be the fate of the *xyl* genes in dairy *L. lactis* strains? We speculate that the xylose metabolic genes in dairy *L. lactis* strains are becoming pseudogenes because there is no selective pressure to maintain xylose metabolism in the dairy environment. Deleterious mutations in the *xyl* genes have accumulated in *L. lactis* 210, and dairy starter *L. lactis* subsp. *cremoris* strains have even more extensive mutations to the *xyl* genes than *L. lactis* subsp. *lactis* 210 (data not shown). Interestingly, *L. lactis* subsp. *cremoris* strains have never been isolated from plant or other environmental samples, despite thorough searching (22, 34, 35). It is also possible that the *xyl* genes are evolving new catalytic functions in dairy lactococcal strains. To examine this, the activity of the isomerase and kinase enzymes would need to be tested against alternative substrates and/or mutants isolated with increased activity against these substrates.

In conclusion, we have discovered the genetic potential in *L. lactis* for xylose metabolism encoded by the *xylRAB* genes. These genes are organized in what could be a single *xylAB* operon, potentially regulated by *xylR*. The Xyl⁻ dairy starter culture strain *L. lactis* 210 has experienced a loss of function in an environment no longer selective for xylose metabolism. Plant environmental isolates such as *L. lactis* B-4449 and *L. lactis* IO-1 retain the ability to metabolize xylose.

REFERENCES

- Ahlem, C., W. Huisman, G. Neslund, and A. S. Dahms. 1982. Purification and properties of a periplasmic D-xylose-binding protein from *Escherichia coli* K-12. *J. Biol. Chem.* **257**:2926-2931.
- Altschul, S. F., W. Gish, W. Miller, E. W. Meyers, and D. J. Lipman. 1990. Basic Local Alignment Search Tool. *J. Mol. Biol.* **215**:403-410.
- Bor, Y.-C. 1993. Ph.D. thesis. Cornell University, Ithaca, N.Y.
- Bor, Y.-C., C. Moraes, S.-P. Lee, W. Crosby, A. J. Sinskey, and C. A. Batt. 1992. Cloning and sequencing of the *Lactobacillus brevis* gene encoding xylose isomerase. *Gene* **114**:127-131.
- Bork, P., C. Sander, and A. Valencia. 1992. Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase and galactokinase families. *Prot. Sci.* **2**:31-40.
- Chaillou, S., Y.-C. Bor, C. A. Batt, P. W. Postma, and P. H. Pouwels. 1998. Molecular cloning and functional expression in *Lactobacillus plantarum* 80 of *xylT*, encoding the D-xylose-H⁺ symporter of *Lactobacillus brevis*. *Appl. Environ. Microbiol.* **64**:4720-4728.
- Czajka, J., and C. A. Batt. 1994. Verification of causal relationships between *Listeria monocytogenes* isolates implicated in food-borne outbreaks of listeriosis by randomly amplified polymorphic DNA patterns. *J. Clin. Microbiol.* **32**:1280-1287.
- Dahl, M. K., J. Degenkolb, and W. Hillen. 1994. Transcription of the *xyl* operon is controlled in *Bacillus subtilis* by tandem overlapping operators spaced by four base-pairs. *J. Mol. Biol.* **243**:413-424.
- Davis, E. O., and P. J. F. Henderson. 1987. The cloning and DNA sequence of the gene *xylE* for xylose-proton symport in *Escherichia coli* K12. *J. Biol. Chem.* **262**:13928-13932.
- Dische, Z., and E. Borenfreund. 1951. A new spectrophotometric method for the detection and determination of keto sugars and trioses. *J. Biol. Chem.* **192**:583-587.
- El Khal, W. 1997. M.S. thesis. Cornell University, Ithaca, N.Y.
- Gallegos, M., R. Schlieff, A. Bairoch, K. Hofman, and J. L. Ramos. 1997. AraC/XylS family of transcriptional regulators. *Microbiol. Mol. Biol. Rev.* **61**:393-410.
- Gartner, D., J. Degenkolb, A. E. Ripperger, R. Allmansberger, and W. Hillen. 1992. Regulation of the *Bacillus subtilis* W23 xylose utilization operon: interaction of the Xyl repressor with the *xyl* operator and the inducer xylose. *Mol. Gen. Genet.* **232**:415-422.
- Ghangas, G. S., and D. B. Wilson. 1984. Isolation and characterization of the *Salmonella typhimurium* xylose regulon. *J. Bacteriol.* **157**:158-164.
- Gribskov, M., A. D. McLachlan, and D. Eisenberg. 1987. Profile analysis: detection of distantly related proteins. *Proc. Natl. Acad. Sci. USA* **84**:4355-4359.
- Hastrup, S. 1988. Analysis of the *Bacillus subtilis* xylose regulon, p. 79-83. In A. T. Ganesan and J. A. Hoch (ed.), *Genetics and biotechnology of bacilli*, vol. 2. Academic Press, Inc., New York, N.Y.
- Hurley, J. H. 1996. The sugar kinase/heat shock protein 70/actin superfamily: implications of conserved structure for mechanism. *Annu. Rev. Biophys. Biomol. Struct.* **25**:137-162.
- Hurley, J. H., H. R. Faber, D. Worthylake, N. D. Meadow, S. Roseman, D. W. Pettigrew, and S. J. Remington. 1993. Structure of the regulatory complex of *Escherichia coli* III^{Glc} with glycerol kinase. *Science* **259**:673-677.
- Jacob, S., R. Allmansberger, D. Gartner, and W. Hillen. 1991. Catabolite repression of the operon for xylose utilization from *Bacillus subtilis* W23 is mediated at the level of transcription and depends on a *cis* site in the *xylA* reading frame. *Mol. Gen. Genet.* **229**:189-196.
- Kelly, W. J., G. P. Davey, and L. J. Ward. 1998. Characterization of lactococci isolated from minimally processed fresh fruit and vegetables. *Int. J. Food Microbiol.* **45**:85-92.
- Kim, S. G., and C. A. Batt. 1988. Heterologous expression and stability of the *Escherichia coli* β -galactosidase gene in *Streptococcus lactis* by translation fusion. *Food Microbiol.* **5**:59-73.
- Klijn, N., A. H. Weerkamp, and W. M. De Vos. 1995. Detection and characterization of lactose-utilizing *Lactococcus* spp. in natural ecosystems. *Appl. Environ. Microbiol.* **61**:788-792.
- Kreuzer, P., D. Gartner, R. Allmansberger, and W. Hillen. 1989. Identification and sequence analysis of the *Bacillus subtilis* W23 *xylR* gene and *xyl* operator. *J. Bacteriol.* **171**:3840-3845.
- Lawlis, V. B., M. S. Dennis, E. Y. Chen, D. H. Smith, and D. J. Henner. 1984. Cloning and sequencing of the xylose isomerase and xylulokinase genes of *Escherichia coli*. *Appl. Environ. Microbiol.* **47**:15-21.
- Lee, C., L. Bhatnagar, B. C. Saha, Y. E. Lee, M. Takagi, T. Imanaka, M. Bagdasarjan, and J. G. Zeikus. 1990. Cloning and expression of the *Clostridium thermosulfurogenes* glucose isomerase gene in *Escherichia coli* and *Bacillus subtilis*. *Appl. Environ. Microbiol.* **56**:2638-2643.
- Lokman, B. C., M. Heerikhuisen, R. J. Leer, A. van den Broek, Y. Borsboom, S. Chaillou, P. Postma, and P. H. Pouwels. 1997. Regulation of expression of the *Lactobacillus pentosus* *xylAB* operon. *J. Bacteriol.* **179**:5391-5397.
- Lokman, B. C., R. J. Leer, and R. van Sorge. 1994. Promoter analysis and transcriptional regulation of *Lactobacillus pentosus* genes involved in xylose catabolism. *Mol. Gen. Genet.* **245**:117-125.
- Lokman, B. C., P. Van Santen, J. C. Verdoes, J. Kruse, R. J. Leer, M. Posno, and P. H. Pouwels. 1991. Organization and characterization of three genes involved in D-xylose catabolism in *Lactobacillus pentosus*. *Mol. Gen. Genet.* **230**:161-169.
- Lüthy, R., I. Xenarios, and P. Bucher. 1994. Improving the sensitivity of the sequence profile method. *Prot. Sci.* **3**:139-146.
- Pettigrew, D. W., G. B. Smith, K. P. Thomas, and D. C. Dodds. 1998. Conserved active site aspartates and domain-domain interactions in regulatory properties of the sugar kinase superfamily. *Arch. Biochem. Biophys.* **349**:236-245.
- Rosenfeld, S. A., P. E. Stevis, and W. Y. Ho. 1984. Cloning and characterization of the *xyl* genes from *Escherichia coli*. *Mol. Gen. Genet.* **194**:410-415.
- Rost, B., and C. Sander. 1993. Prediction of protein structure at better than 70% accuracy. *J. Mol. Biol.* **232**:584-599.
- Rygu, T., A. Scheler, R. Allmansberger, and W. Hillen. 1991. Molecular cloning, structure, promoters and regulatory elements for transcription of the *Bacillus megaterium* encoded regulon for xylose utilization. *Arch. Microbiol.* **155**:535-542.
- Salama, M. S., T. Musafija-Jeking, W. E. Sandine, and S. J. Giovannoni. 1995. An ecological study of lactic acid bacteria: isolation of new strains of *Lactococcus* including *Lactococcus lactis* subspecies *cremoris*. *J. Dairy Sci.* **78**:1004-1017.
- Salama, M. S., W. E. Sandine, and S. J. Giovannoni. 1993. Isolation of *Lactococcus lactis* subsp. *cremoris* from nature by colony hybridization with rRNA probes. *Appl. Environ. Microbiol.* **59**:3941-3945.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed., vol. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Scheler, A., and W. Hillen. 1994. Regulation of xylose utilization in *Bacillus licheniformis*: Xyl repressor-*xyl*-operator interaction studied by DNA modification, protection and interference. *Mol. Microbiol.* **13**:505-512.
- Scheler, A., T. Rygu, R. Allmansberger, and W. Hillen. 1991. Molecular cloning, structure, promoters and regulatory elements for transcription of the *Bacillus licheniformis*-encoded regulon for xylose utilization. *Arch. Microbiol.* **155**:526-534.
- Schleifer, K. H., J. Kraus, C. Dvorak, R. Klipper-Balz, M. D. Collins, and W. Fischer. 1985. Transfer of *Streptococcus lactis* and related streptococci to the genus *Lactococcus* gen. nov. *Syst. Appl. Microbiol.* **6**:183-195.

40. **Schmiedel, D., and W. Hillen.** 1996. Contributions of XylR, CcpA and *cre* to diauxic growth of *Bacillus megaterium* and to xylose isomerase expression in the presence of glucose and xylose. *Mol. Gen. Genet.* **250**:259–266.
41. **Schmiedel, D., M. Kintrup, E. Kuster, and W. Hillen.** 1997. Regulation of expression, genetic organization and substrate specificity of xylose uptake in *Bacillus megaterium*. *Mol. Microbiol.* **23**:1053–1062.
42. **Shamanna, D. K., and K. E. Sanderson.** 1979. Genetics and regulation of D-xylose in *Salmonella typhimurium* LT2. *J. Bacteriol.* **139**:71–79.
43. **Shamanna, D. K., and K. E. Sanderson.** 1979. Uptake and catabolism of D-xylose in *Salmonella typhimurium* LT2. *J. Bacteriol.* **139**:64–70.
44. **Sizemore, C., E. Buchnere, T. Rygus, C. Witke, F. Goetz, and W. Hillen.** 1991. Organization, promoter analysis and transcriptional regulation of the *Staphylococcus xylosus* xylose utilization operon. *Mol. Gen. Genet.* **227**:377–384.
45. **Sizemore, C., B. Wieland, F. Gotz, and W. Hillen.** 1992. Regulation of *Staphylococcus xylosus* xylose utilization genes at the molecular level. *J. Bacteriol.* **174**:3042–3048.
46. **Song, S., and C. Park.** 1997. Organization and regulation of the D-xylose operons in *Escherichia coli* K-12: XylR acts as a transcriptional activator. *J. Bacteriol.* **179**:7025–7032.
47. **Titgemeyer, F., J. Reizer, A. Reizer, and M. H. Saier.** 1994. Evolutionary relationships between sugar kinases and transcription repressors in bacteria. *Microbiology* **140**:2349–2354.
48. **Vangrype, W., J. VanDamme, J. Vanderkerckhove, C. K. DeBruyne, R. Cornelis, and H. Kerster-Hilderson.** 1990. Localization of the essential histidine and carboxylate group in D-xylose isomerases. *Biochem. J.* **265**:699–706.
49. **Wang, Y.** 1988. Double-stranded DNA sequencing with T7 polymerase. *BioTechniques* **6**:843–845.
50. **Whitaker, R. D.** 1995. Ph.D. thesis. Cornell University, Ithaca, N.Y.
51. **Whitaker, R. D., Y. Cho, J. Cha, H. L. Carrell, J. P. Glusker, P. A. Karplus, and C. A. Batt.** 1995. Probing the roles of active site residues in D-xylose isomerase. *J. Biol. Chem.* **270**:22895–22906.
52. **Whitlow, M., A. J. Howard, B. C. Finzel, T. L. Poulos, E. Winborne, and G. L. Gilliland.** 1991. A metal-mediated hydride shift mechanism for xylose isomerase based on the 1.6 Å *Streptomyces rubiginosus* structures with xylitol and D-xylose. *Proteins Struct. Funct. Genet.* **9**:153–173.