Activation of Silent gal Genes in the lac-gal Regulon of Streptococcus thermophilus

Vaughan, Elaine E.; Bogaard, Patrick T. C. van den; Catzeddu, Pasquale; Kuipers, Oscar P.; Vos, Willem M. de

Published in:
Journal of Bacteriology

DOI:
10.1128/JB.183.4.1184-1194.2001

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2001

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Activation of Silent gal Genes in the lac-gal Regulon of Streptococcus thermophilus

ELAINE E. VAUGHAN,1,2* PATRICK T. C. VAN DEN BOGAARD,1 PASQUALE CATZEDDU,1† OSCAR P. KUIPERS,1‡ and WILLEM M. DE VOS1,2

Wageningen Centre for Food Sciences, NIZO Food Research, 6718 ZB Ede,1 and Laboratory of Microbiology, Department of Agrotechnology and Food Sciences, Wageningen University, 6703 CT Wageningen,2 The Netherlands

Received 16 August 2000/Accepted 16 November 2000

Streptococcus thermophilus strain CNRZ 302 is unable to ferment galactose, neither that generated intracellularly by lactose hydrolysis nor the free sugar. Nevertheless, sequence analysis and complementation studies with Escherichia coli demonstrated that strain CNRZ 302 contained structurally intact genes for the Leloir pathway enzymes. These were organized into an operon in the order galKTE, which was preceded by a divergently transcribed regulator gene, galR, and followed by a galM gene and the lactose operon lacSZ. Results of Northern blot analysis showed that the structural gal genes were transcribed weakly, and only in medium containing lactose, by strain CNRZ 302. However, in a spontaneous galactose-fermenting mutant, designated NZ302G, the galKTE genes were well expressed in cells grown on lactose or galactose. In both CNRZ 302 and gal galR mutant NZ302G, the transcription of the galR gene was induced by growth on lactose. Disruption of galR indicated that it functioned as a transcriptional activator of both the gal and lac operons while negatively regulating its own expression. Sequence analysis of the gal promoter regions of NZ302G and nine other independently isolated Gal+ mutants of CNRZ 302 revealed mutations at three positions in the galK promoter region, which included substitutions at positions −9 and −15 as well as a single-base-pair insertion at position −37 with respect to the main transcription initiation point. Galactokinase activity measurements and analysis of gusA reporter gene fusions in strains containing the mutated promoters suggested that they were gal promoter-up mutations. We propose that poor expression of the gal genes in the galactose-negative S. thermophilus CNRZ 302 is caused by naturally occurring mutations in the galK promoter.

After its discovery almost 40 years ago, the Escherichia coli lactose operon, encoding enzymes of lactose metabolism, became the first model for gene regulation (reviewed in reference 4). The key component of this system is the lac repressor (LacI), the product of the lacI gene. The lac operon contains a primary operator (O1), which is the major element of repression by LacI, and two pseudo-operators, which enhance repressor binding to O1 by cooperativity. Control of the lac operon also involves activation by the cyclic AMP receptor protein. Many other operon systems of negative control have since been described, including GalR, one of the two repressors of the gal regulon encoding enzymes of galactose transport and metabolism in E. coli. Regulation of the gal regulon is mediated through GalR, GalS (Gal isorepressor), and the cyclic AMP receptor protein. GalR and GalS negatively regulate transcription of the two promoters of the gal operon, although GalS is not as efficient as GalR (57).

The bioconversion of lactose, which is the primary carbon and energy source in milk, into lactic acid is an essential process in industrial dairy fermentations carried out by lactic acid bacteria. Genetic studies of the metabolic pathways for lactose utilization in these gram-positive bacteria have revealed a variety of lac operons that differ from the paradigm known in E. coli (13). The thermophilic yogurt bacteria Streptococcus thermophilus and Lactobacillus bulgaricus contain a highly homologous lacSZ operon in which the β-galactosidase (lacZ) gene is located downstream from the lacS gene encoding a lactose permease (LacS), which belongs to the galactoside-pentose-hexulonide translocators (27, 41, 48, 49).

Although lactose is efficiently transported and hydrolyzed intracellularly, many strains of S. thermophilus and L. bulgaricus do not grow on galactose and ferment only the glucose portion of lactose, while the galactose is excreted into the medium in amounts stoichiometric with the uptake of lactose (20, 22). Kinetic studies indicated that LacS mediates both galactoside exchange (e.g., lactose-galactose) and movement of galactosides and protons (15). The exchange reaction is highly favored with excess galactosides on either side of the membrane and may account for the galactose-negative (Gal−) phenotype of S. thermophilus in milk which contains an excess of lactose (40). Another explanation for the Gal− phenotype may be the absence of functional Leloir pathway enzymes, including galactokinase (GalK), galactose-1-phosphate uridyltransferase (GalT), and UDPglucose 4-epimerase (GalE), products of the galK, galT, and galE genes, respectively. Remarkably, under appropriate selective conditions, such as limiting lactose and excess galactose, Gal− derivatives of S. thermophilus were obtained which fermented galactose and contained Leloir enzyme activities (21, 50). However, no
molecular explanation was given, and the genetics of the Leloir pathway has only been poorly investigated in *S. thermophilus*. The *lacSZ* operon of strain A147 was found to be preceded by *galE* and *galM* (42). The *galM* gene appeared to be constitutively expressed and could encode a mutarotase that, similar to the homologous enzyme of *E. coli*, forms the galactokinase substrate α-D-galactose from β-D-galactose pyranose generated from lactose by β-galactosidase (LacZ) (7). However, *S. thermophilus* A147 is not a Gal⁺ strain.

The present study was undertaken to gain insight into the presence and regulation of the *gal* genes of *S. thermophilus* and the mechanism by which these genes, in particular the *galK* gene, are prevented from being expressed. Here we describe the characterization of the *gal* operon, consisting of the *galK*, *galT* and *galE* genes, and its promoter from *S. thermophilus* CNRZ 302, for which galactose-fermenting (Gal⁺) revertants have been reported (5). A regulatory gene, *galR*, was identified that is divergently transcribed from the *gal* operon. Analysis of mRNA for the *gal* metabolic genes from a Gal⁺ fermenting derivative of CNRZ 302 indicated that regulation occurred at the transcriptional level. In contrast, the *gal* metabolic genes of the original Gal⁻ strain were not sufficiently transcribed to allow galactose metabolism. Furthermore, we demonstrate that GalR acts as a transcriptional activator of both the *gal* and *lac* operons and negatively regulates its own expression. To the best of our knowledge, this is the first report describing the mechanisms regulating galactose utilization in *S. thermophilus* at the molecular level.

### MATERIALS AND METHODS

**Bacterial strains and growth conditions.** The bacterial strains and plasmids used in this study are described in Table 1. *S. thermophilus* strains were subcultured in M17 broth (Oxoid, Basingstoke, England), containing either 1% lactose, glucose, or galactose as necessary, at 42°C unless stated otherwise. The taxonomic position of strain CNRZ 302 was confirmed by 16S rRNA sequence analysis and corresponded to *Streptococcus thermophilus* (GenBank accession number X86419). *E. coli* strains HB101 or LE392 and TG1 were used for the isolation of pACYC184- and pUC19-derived plasmids and for the propagation of bacteriophage M13 chimeras, respectively. *E. coli* was routinely grown in TY medium (45) or brain heart infusion (Difco) broth with aeration at 37°C. MacConkey agar base (Difco Laboratories) supplemented with 1% galactose was used to detect galactose-positive (Gal⁺) *E. coli* strains. Agar media were prepared by adding 1.5% agar to broth. The antibiotics used for selection in media were chloramphenicol at 4 μg/ml and erythromycin at 2.5 μg/ml for *S. thermophilus* by the procedure of Hayes et al. (19). The Anderson and McKay (3) lysis procedure was used to detect plasmid DNA in *S. thermophilus*. Restriction enzymes, T4 DNA ligase, and other DNA-modifying enzymes were used as recommended by the supplier (Gibco-BRL, Life Technologies, Gaithersburg, Md.). DNA fragments were recovered from agarose gels with the GlassMatrix DNA isolation system (Gibco-BRL). Electroporation of *S. thermophilus* was performed by the procedure of Mollet et al. (32) with the modification that the harvested cells were incubated in the electroporation buffer at 4°C for at least 4 h prior to electroporation. PCR was performed under the conditions described previously (25) using Taq polymerase (Life Technologies) or Pwo polymerase (Boehringer Mannheim). Oligonucleotides were synthesized by Eurogentech (Gent, Belgium).
Cloning of gal genes. S. thermophilus CNZR 302 total genomic DNA was digested with EcoRI, and the DNA fragments were separated by agarose gel electrophoresis (0.7% agarose). The DNA was transferred to a GeneScreen Plus (Dupont, Boston, Mass.) membrane by established methods (46). The membrane was hybridized with a 700-bp probe containing partial galK gene kindly provided by B. Hutkins (34). The labeling of this fragment with horseradish peroxidase and the hybridization and detection methods were as described in the manufacturer's manual for the ECL system (Amersham, Little Chalfont, United Kingdom). Fragments of approximately 5 kb were recovered, ligated with EcoRI-linearized calf intestinal alkaline phosphatase-treated pUCYC184, and used to transform E. coli HB101 and LE392. Gal colonies were selected as red Tc’ colonies on McConkey galactose agar from the frequency of approximately 1%. Analysis of the plasmid content of all 10 Gal+ colonies indicated that they contained a recombinant plasmid with a 4.9-kb insert that showed an identical restriction pattern. One of the clones, designated HB101(pNZ680), was used in further experiments.

Nucleotide sequence analysis. DNA fragments were subcloned in the phage vectors M13mp18 and M13mp19 with T7G as a host by using standard techniques (46). Nucleotide sequences of both strands were determined by the dideoxy-chain termination method (47), adapted for Sequenase version 2.0 (US Biochemical Corp., Cleveland, Ohio) with either the M13 universal primer or specifically synthesized primers. The gal promoter regions of S. thermophilus CNZR 302 and its 10 Gal+ derivatives were isolated as 350-bp PCR fragments from agarose gels. The purified fragments and primers were annealed by boiling for 5 min and rapidly freezing in liquid nitrogen, and sequencing proceeded as described above. The sequence data were assembled and analyzed with PC/GENE version 6.6 (Genofit, Geneva, Switzerland). Amino acid sequence comparisons were performed with the EMBL (release 34.1), SwissProt (release 28.0), and NBRF/PIR (release 40.0) databases using the FASTA program (36), and sequence comparisons were performed with the EMBL (release 31.0), SwissProt (release 28.0), and NBRF/PIR (release 40.0) databases using the FASTA program (36).

Isolation of galK. S. thermophilus strains. S. thermophilus CNZR 302 cultures grown in M17 broth supplemented with 1% lactose were diluted 100-fold into M17 broth containing 1% galactose and 0.01% glucose and incubated for 24 h. Cultures that exhibited growth were transferred to M17 broth containing 1% galactose (Gal-M17) and incubated for 16 to 20 h. Ten Gal+ single-colony isolates were obtained by plating 10 cultures treated as described above on M17 agar with 1% galactose, and these were designated NZ302G and SS1 to SS9.

RNA isolation, Northern blotting, and primer extension analysis. S. thermophilus strains. S. thermophilus CNZR 302 cultures grown in M17 broth supplemented with 1% lactose were diluted 100-fold into M17 broth containing 1% galactose and 0.01% glucose and incubated for 24 h. Cultures that exhibited growth were transferred to M17 broth containing 1% galactose (Gal-M17) and incubated for 16 to 20 h. Ten Gal+ single-colony isolates were obtained by plating 10 cultures treated as described above on M17 agar with 1% galactose, and these were designated NZ302G and SS1 to SS9.

RNA isolation, Northern blotting, and primer extension analysis. S. thermophilus strains. S. thermophilus CNZR 302 cultures grown in M17 broth supplemented with 1% lactose were diluted 100-fold into M17 broth containing 1% galactose and 0.01% glucose and incubated for 24 h. Cultures that exhibited growth were transferred to M17 broth containing 1% galactose (Gal-M17) and incubated for 16 to 20 h. Ten Gal+ single-colony isolates were obtained by plating 10 cultures treated as described above on M17 agar with 1% galactose, and these were designated NZ302G and SS1 to SS9.

Enzyme and protein assays and chromatography. The S. thermophilus strains were grown in M17 broth containing either 1% lactose, galactose, glucose, or sucrose with the appropriate antibiotics to an optical density at 600 nm (OD600) of 1.0. For the preparation of extracts, cells were disrupted with zirconium glass beads in a Bead Beater (Biospec Products, Bartlesville, Okla.) for 3-min treatments with intervals of 1 min on ice between treatments and cellular debris was removed by centrifugation. The extracts were kept on ice, and enzyme assays were performed with 4 h. Galactose 1-phosphate uridylyltransferase activity was measured in the resulting extracts with 30 to 350 μg of protein per assay by the spectrophotometric method of Isselbacher (23). β-Galactosidase was assayed at 37°C by the method of Miller (31) using 1 to 6 μg of protein per assay and galactosidase assays by the method of Avidov et al. (2). All enzyme activity measurements presented were the mean of at least two independent experiments. Proteins concentrations were estimated by a dye binding assay (8).

Nucleotide sequence accession number. The GenBank accession number assigned to the nucleotide sequence encoding S. thermophilus galk, galK, galT, and the partial galE gene is U61402.
A fragment was constructed in the chloramphenicol resistance (cat) gene of pACYC184, and the putative galK gene of *S. thermophilus* CNRZ 302 was isolated by functional complementation of the galK2 mutation of *E. coli* HB101. The complementing plasmid, designated pNZ680 (Fig. 1A), contained a 4.9-kb insert that allowed HB101 to form red colonies on McConkey galactose agar and utilize galactose as the sole carbon source in minimal M9 medium. Introduction of a frameshift mutation in the *Kpn*I site on pNZ680 resulted in a plasmid which could not restore a Gal+ phenotype to HB101, indicating that the galK gene was overlapping this site (data not shown). Moreover, pNZ680 also complemented both the galK2 and galT22 mutations of *E. coli* LE392, indicating that it also contained the galT gene of *S. thermophilus*. Although these gal genes are hardly expressed in *S. thermophilus*, the promoter upstream of the cat gene in pACYC184 is likely to be responsible for their expression in *E. coli*. The presence of a functional galT gene was further confirmed by assaying for GalT enzymatic activity. Cell extracts of LE392(pNZ680) contained 119 nmol of GalT activity per min per mg, whereas no activity was detected for the LE392 strain alone.

**Organization and similarity studies of the *S. thermophilus* gal region.** Commencing at the *Kpn*I site on pNZ680, nucleotide sequence analysis in both directions revealed the galK open reading frame (ORF), 1,164 bp in length (Fig. 1A). The deduced galK sequence had the strongest similarity to the GalK proteins of several gram-positive bacterial species including *Streptococcus mutans* (83%) and *Lactobacillus casei* (70%) (2, 6). The *S. thermophilus* CNRZ 302 GalK was also 79% similar to the GalK of the Gal+ *S. thermophilus* F410 strain (34). A potential ribosome binding site (5'-GAGA-3'), complementary to the 3' end of the 16S rRNA of lactic acid bacteria (12), was located 8 bp upstream from the first translational initiation codon at nucleotide (nt) 1483.

Upstream of galK located in a divergent orientation, a 1,014-bp ORF was designated galR on the basis of the similarity of its deduced amino acid sequence to proteins of the LacI-GalR family of transcriptional regulators (55) (Fig. 1A). The translational initiation site at nt 1340 is proposed on the basis of the position of the putative ribosome binding site (5'-AGGAGGA-3', nt 1351 to 1345) and the similarity between related proteins (see also below). The *S. thermophilus* GalR had the greatest similarity to the GalR repressor of *S. mutans* (75%; 57% identity) and the potential GalR repressor of *L. casei* (59%; 40% identity) (2, 6). There was also significant similarity, 53 and 48% (35 and 27% identity), to the evolved β-galactosidase (EbgR) and galactose (GalR) repressors, respectively, of *E. coli* (18, 54). An inverted-repeat structure and a stretch of five T nucleotides (nt 95 to 56) that could...
function as a rho-independent transcriptional terminator (38), followed the galR gene sequence.

DNA sequence analysis downstream of galK revealed the galT gene (1,482 bp), whose deduced sequence was similar (67 to 74%) to the GalT proteins from several gram-positive bacteria including S. mutans, L. casei, and Lactococcus lactis (Fig. 1A) (2, 6, 53). The stop codon for the S. thermophilus galK gene and the start codon of the galT gene were separated by just 19 bp. The translational initiation site for galT was preceded by a putative ribosome binding site (nt 2654 to 2660). Finally, a fourth ORF was present immediately downstream of the galT gene and reading beyond the pNZ680 clone (Fig. 1A). The nucleotide and predicted amino acid sequence were identical to the aminoterminus of the previously characterized galE gene from S. thermophilus A147 (42).

The organization of lac genes in relation to the galE gene, which have been characterized for S. thermophilus A147 (41, 42), was demonstrated to be identical in CNRZ 302. Long-range PCR was performed using primers based on the sequences of the galK gene of CNRZ 302 and the lacS gene (41) of A147, which resulted in the expected 5.7-kb product. Restriction enzyme analysis of the PCR product showed an identical pattern to that of A147, confirming the presence of the galE, galM, and lacS genes downstream of galK-galT in strain CNRZ 302 (Fig. 1B).

Transcriptional analysis of the gal genes. The transcription of the gal genes was analyzed in the wild-type S. thermophilus strain CNRZ 302, and in strain NZ302G, an isogenic spontaneous Gal" mutant strain. Strain NZ302G has a doubling time of 58 min in M17 medium containing 1% galactose at 42°C, in contrast to the wild-type CNRZ 302 parental strain which does not grow at all on galactose. The Gal" phenotype of NZ302G was stably maintained even after several subcultures in M17 containing lactose.

Northern analysis failed to detect hybridization signals for galK or galTE from the glucose-grown wild-type and Gal" strains (Fig. 2). Only after prolonged exposure, were weak signals obtained for lactose-grown wild-type cells (data not shown). However, mRNA was detected for the lactose- and galactose-grown NZ302G cells (Fig. 2), in accordance with the Gal" phenotype. The weaker signal obtained for the galactose-grown cells may be due to the poorer-quality RNA obtained as a result of their slower growth. The transcripts were stronger when hybridized with the larger galTE probe than when hybridized with the galK probe. The size of the predominant mRNA hybridizing to the galK and galTE probes was approximately 3.7 kb, indicating that the galK, galT, and galE genes, which are 1.2, 1.4, and 1.0 kb, respectively, are transcribed together as a single mRNA. In conclusion, sufficient induction of galKTE mRNA for galactose metabolism occurred only in the Gal" NZ302G strain when it was cultured in lactose- or galactose-containing M17 medium.

A major transcript of approximately 1.2 kb was identified for galR from both the wild-type strain grown in lactose medium and the mutant strain grown in lactose or galactose medium (Fig. 2). However, only a weak signal for this galR mRNA was detected in glucose-grown cells for both strains. The size of the transcript for the 1.0-kb galR gene suggests that it is transcribed alone and supports the functional role of the terminator following galR. These data indicate that the expression of the galR gene is also regulated at the level of transcription.

Mapping and characterization of the galR and galK promoters. The transcriptional start points of the galK and galR genes were determined by primer extension analysis. Total RNA was isolated from lactose- and glucose-grown S. thermophilus CNRZ 302 cells and from lactose-, glucose-, and galactose-grown NZ302G cells. A major transcriptional start site was observed for the galK gene of strain NZ302G that mapped at an A residue (nt 1452), 6 bp downstream of the inferred −10 sequence (TACGAT) (Fig. 3A). The latter was separated by 17

FIG. 2. Northern blot analysis of RNA isolated from S. thermophilus strains CNRZ 302 (Gal') and NZ302G (Gal') grown on glucose (Gl), lactose (L), or galactose (Ga). The probes used for the hybridization are indicated below the blot. The 23S and 16S rRNAs and the mRNA transcript sizes are indicated by arrows.
bp from a −35 sequence (TTGATT) that conforms well to the *E. coli* and *S. thermophilus* promoter consensus sequences (Fig. 4A). Essentially no clear signal for the *galK* gene of CNRZ 302 was detected, in accordance with its Gal− phenotype (Fig. 3A).

For *galR*, a strongly labeled extension product was detected that initiated at a G residue (nt 1352) 7 bp upstream of a putative −10 sequence (TATACT) (Fig. 3B) for both CNRZ 302 and NZ302G. A putative −35 sequence (TGGTGA) could be found 18 bp upstream of the −10 box (Fig. 4A). The presence of the primer extended products for *galK* and *galR* matched exactly the results obtained in the Northern experiments and confirmed control at the transcriptional level.

**Effect of the *galR* gene disruption on galactose utilization.**

To determine the function of the *galR* gene in the *Gal*+ *S. thermophilus* NZ302G, the gene was disrupted using pNZ684, consisting of the temperature-sensitive pG+ host p9 vector that carried an internal fragment of the *galR* gene (Fig. 1C). The disruption caused by pNZ684 resulted in two partially deleted copies of the *galR* gene, one of which lacks the DNA for the 17 N-terminal amino acids including most of the conserved DNA binding motif while the other suffers from a 200-bp deletion.

The −10 region is denoted by a vertical bar, and the transcription initiation sites are indicated by +1. The arrows point to the major primer-extended products.

A high-performance liquid chromatography analysis of the spent medium indicated that the galactose moiety of lactose was completely metabolized by strain NZ302G while, in contrast, a substantial amount of galactose (72% of the amount that could be hydrolyzed from lactose) was not utilized by the NZ302GΔ*galR* integrant. Moreover, the doubling time of NZ302GΔ*galR* on lactose increased to approximately 1 h in comparison to that of NZ302G, which is 25 min, and the final OD600 was less than half that of NZ302G (Table 2).

To exclude any possible polar effects of the integration of pNZ684 in the NZ302GΔ*galR* integrant, complementation of its *galR* mutation was studied. Plasmid pNZ6811 contains the *galR* gene under its own promoter on a high-copy-number vector based on pNZ123 (12). Three transformants of NZ302GΔ*galR* harboring pNZ6811 grew to an OD600 of approximately 1.4 in M17 medium containing 0.4% lactose broth, and no residual galactose was detected in the cell-free supernatant (Table 2). Furthermore, the transformants could again utilize galactose as the sole carbon source (data not shown). Thus, GalR is necessary for the ability to utilize galactose.

**Effect of *galR* disruption on transcription of the *gal* and *lac* genes.**

Northern hybridizations were performed to determine the effect of the *galR* disruption on the transcription of the *galR* gene and of the *gal* and *lac* operons. The primer GALRAS, which is complementary to the 5′ end of the sense strand of *galR*, was chosen since it can hybridize to the single copy of the 3′-deleted *galR* (*galR′*) in the chromosome that is under control of the *galR* promoter (Fig. 1D). The 1.2-kb *galR* transcript, which was only weakly visible for NZ302G growing on glucose, gave a signal of much greater intensity for strain NZ302GΔ*galR* (Fig. 5A). The 5′-deleted copy of *galR* did not appear to be transcribed since the same result was obtained by probing with a 600-bp fragment internal to the *galR* gene (Fig. 5B). This constitutive overexpression of the 3′-truncated *galR′* was also

![Image](image-url)
observed for lactose-grown cells (data not shown) and suggests that the product of the galR gene is a negative regulator of its own expression. It should be noted that termination of transcription of galR in NZ302G is located at an unknown point within the integration plasmid pNZ684, and therefore it is coincidental that the galR and galR transcripts appear to have similar sizes.

A galK probe was used to monitor the transcription of the galKTE operon in NZ302G and the mutant NZ302G D (Fig. 5C). While the 3.7-kb transcript of the galKTE operon was observed when NZ302G was grown on lactose, no transcript was detected for NZ302G D. This indicates that GalR is an activator of transcription for the galKTE operon and explains why the galactose moiety of lactose is not effectively metabolized by strain NZ302G D.

LacS is the sole galactoside transporting activity in S. thermophilus and therefore is essential for both lactose and galactose transport (15). We speculated that GalR may also play a role in the regulation of the lac operon (lacS-lacZ). Northern hybridization of NZ302G probed with an internal fragment of lacS showed that a 5.2-kb transcript, corresponding to the sizes of the lacS and lacZ genes, was expressed weakly when the strain was grown on glucose (barely visible on the blot) and induced when the strain was grown on lactose or galactose (Fig. 5D). The lac operon was still transcribed in NZ302G D when the strain was grown on glucose, but there was no longer induction of transcription when it was grown on lactose. This indicates that in addition to being the positive regulator of the gal operon, GalR functions as an inducer of transcription of the lac operon.

**Effect of galR disruption on GalT and LacZ enzyme activities.** To further substantiate the effect of the galR gene disruption on the expression of the gal and the lac operons, enzyme assays were performed. Since the galT gene is the central one of the three genes in the gal operon, GalT activity was used as a measure of the activity of the Leloir pathway enzymes (Table 3). In strain NZ302G, GalT activity is induced three- to four-fold in lactose- and galactose-grown cultures in comparison to growth in glucose-containing medium. In contrast, very low activity was detected in the galR disruption strain NZ302G D when grown on glucose and no significant increase was observed following growth on lactose, indicating a lack of induction of the enzyme. However, the levels of activity of the NZ302G D mutant complemented by pNZ6811 were similar.

<table>
<thead>
<tr>
<th>Microorganism</th>
<th>Gene</th>
<th>Promoter regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. thermophilus</td>
<td>galK</td>
<td>TTTTACTAAAGAAAATAGTAAAATATGGATT</td>
</tr>
<tr>
<td>S. thermophilus</td>
<td>lacS</td>
<td>TATTACTAAATTTTTACTAAAACACTGAAAATTGACT</td>
</tr>
<tr>
<td>S. thermophilus</td>
<td>galR</td>
<td>TAAAAGTTATGTATACTGAAA</td>
</tr>
<tr>
<td>S. mutans</td>
<td>galK</td>
<td>ATTTACTAAATTTTTACTAAAATATGGACC</td>
</tr>
</tbody>
</table>

**FIG. 4.** (A) Intergenic region containing the promoter sequences for the galK and galR genes. The −10 and −35 regions are underlined, and the transcriptional start sites are indicated as +1. The point mutations for the three classes (I, II, and III) of S. thermophilus Gal1 mutants are as described in the text. Inverted repeats are indicated by arrows. (B) Alignment of similar regions in the promoters of lac and/or gal genes of S. thermophilus and S. mutans. Inverted repeats in the S. thermophilus galK promoter are indicated by arrows. Nucleotides that are complementary in each arm of the inverted repeats are in boldface type.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Final OD 600</th>
<th>% Galactosea</th>
</tr>
</thead>
<tbody>
<tr>
<td>NZ302G</td>
<td>1.84</td>
<td>NDb</td>
</tr>
<tr>
<td>NZ302G ΔR</td>
<td>0.76</td>
<td>0.143</td>
</tr>
<tr>
<td>NZ302G ΔR1(pNZ6811)</td>
<td>1.37</td>
<td>ND</td>
</tr>
<tr>
<td>NZ302G ΔR2(pNZ6811)</td>
<td>1.30</td>
<td>ND</td>
</tr>
<tr>
<td>NZ302G ΔR3(pNZ6811)</td>
<td>1.43</td>
<td>ND</td>
</tr>
</tbody>
</table>

a Percentage of galactose remaining in medium following growth.
b ND, not detectable (<0.01%)
to those of NZ302G on all the sugars tested and thus were also induced on lactose and galactose.

LacZ activity was measured to determine the expression of the lac operon genes. In NZ302G, LacZ activity was induced about two- and fourfold during growth on lactose and galactose, respectively, in comparison to growth on glucose (Table 3). In contrast, LacZ levels in NZ302GR were not induced during growth on lactose. LacZ was again inducible in NZ302GR carrying the galR-expressing plasmid pNZ6811, although the activity was not as high as that of the original strain. Thus, the enzyme activity measurements confirmed the transcriptional analysis results, demonstrating the role of GalR as an activator of the gal and lac operons.

To analyze whether the activating role of GalR was specific for the galactose-utilizing strain NZ302G, the galR gene of S. thermophilus ST11, a well-characterized LacGal strain (32), was also disrupted using the pNZ684 construct. PCR analysis indicated an identical organization of the gal genes in this strain and in CNRZ 302 (data not shown). Since strain ST11 grows very poorly on glucose, LacZ activities were compared using medium containing sucrose (Table 3). Both ST11 and ST11ΔR showed an induction of LacZ of approximately 2.5- to 3-fold after growth on lactose in comparison to growth on sucrose, but the levels of the LacZ activity of ST11ΔR were at least half those produced by ST11 on both carbon sources. The failure to fully induce LacZ in ST11ΔR strongly suggests that GalR also plays a role in activating the lac operon of this strain.

Characterization of the galK promoters for the Gal+ mutants. To gain an understanding of the ability of NZ302G to transcribe the gal metabolic genes, in contrast to the parent strain CNRZ 302, a series of hybridization experiments using specific galR and galK probes on the gal region and PCR

<table>
<thead>
<tr>
<th>Carbohydrate</th>
<th>Strain</th>
<th>Gal-1-phosphate uridylyltransferase activity (nmol/min/mg of protein)</th>
<th>β-Galactosidase activity (μmol/min/mg of protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>NZ302G</td>
<td>59 ± 6</td>
<td>4.2 ± 0.9</td>
</tr>
<tr>
<td></td>
<td>NZ302GΔR</td>
<td>19 ± 3</td>
<td>3.1 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>NZ302GΔR(pNZ6811)</td>
<td>70 ± 12</td>
<td>2.5 ± 0.4</td>
</tr>
<tr>
<td>Lactose</td>
<td>NZ302G</td>
<td>178 ± 73</td>
<td>7.7 ± 0.8</td>
</tr>
<tr>
<td></td>
<td>NZ302GΔR</td>
<td>23 ± 9</td>
<td>3.1 ± 0.4</td>
</tr>
<tr>
<td></td>
<td>NZ302GΔR(pNZ6811)</td>
<td>149 ± 35</td>
<td>6.7 ± 1.8</td>
</tr>
<tr>
<td>Galactose</td>
<td>NZ302G</td>
<td>230 ± 70</td>
<td>15.6 ± 2.8</td>
</tr>
<tr>
<td></td>
<td>NZ302GΔR</td>
<td>393 ± 21</td>
<td>9.9 ± 0.8</td>
</tr>
<tr>
<td></td>
<td>NZ302GΔR(pNZ6811)</td>
<td>NDa</td>
<td>ND</td>
</tr>
<tr>
<td>Sucrose</td>
<td>ST11</td>
<td>NDb</td>
<td>4.8 ± 1.0</td>
</tr>
<tr>
<td></td>
<td>ST11ΔR</td>
<td>ND</td>
<td>2.2 ± 0.3</td>
</tr>
<tr>
<td>Lactose</td>
<td>ST11</td>
<td>ND</td>
<td>13.3 ± 0.9</td>
</tr>
<tr>
<td></td>
<td>ST11ΔR</td>
<td>ND</td>
<td>5.5 ± 0.6</td>
</tr>
</tbody>
</table>

a Results are expressed as mean ± standard deviation. The enzyme assays were performed on the same extracts within 4 h.

b ND, not determined.
amplication of the galKTE gene cluster were performed on the Gal\textsuperscript{+} mutant. However, no DNA structural rearrangements were detected within galR or the metabolic gal gene cluster compared to the parent CNRZ 302 strain (data not shown). The galR and galK promoter regions of CNRZ 302 and NZ302G were amplified by PCR in duplicate with the same primers as those used in the primer extension experiments, and both strands of each product were sequenced. The analysis revealed that an extra A residue was inserted in a stretch of adenines preceding the −35 region (nt 1410 to 1414 [Fig. 4A]) of the galK promoter in the NZ302G DNA sequence, resulting in 6 A residues in the mutant in comparison to 5 in the Gal\textsuperscript{−} parent. To determine whether other Gal\textsuperscript{+} mutants of CNRZ 302 contained similar mutations in the promoter region, nine more independently isolated Gal\textsuperscript{+} mutants of CNRZ 302 were investigated. DNA sequence analysis of these nine promoters showed that the galK promoter of each Gal\textsuperscript{+} mutant also contained a point mutation. The mutants could be divided into three classes based on their mutations: class I consisted of five mutants with an A insertion, as described above for NZ302G; in class II, the three mutants contained a G-to-T substitution 3 bp preceding the −10 box; and the one mutant in the third class had a G-to-A substitution in the −10 box (Fig. 4A).

**galK promoter activity.** The GalK activity of the mutant strains was used as a reporter for comparing the expression of the mutated promoters with that of the wild-type. A low level of GalK activity, 10 nmol/min/mg, was detected for the Gal\textsuperscript{−} S. thermophilus CNRZ 302 grown on glucose-containing medium, and it increased to 37 nmol/min/mg on lactose. For the Gal\textsuperscript{−} NZ302G and the other five class I mutants, the GalK activity increased from an average 46 nmol/min/mg on glucose to 264 and 305 nmol/min/mg on lactose and galactose, respectively. The activity increased from 53 nmol/min/mg on glucose to 193 and 323 nmol/min/mg on lactose and galactose, respectively, for the class III mutant. The highest activity was observed with the three class II mutants, an average of 68 nmol/min/mg on glucose to 400 and 458 nmol/min/mg on lactose and galactose, respectively. Thus, it is likely that the basal GalK activity has increased in the mutant strains due to promoter-up mutations, which allows sufficient expression of the galK gene on induction for galactose utilization.

To support our hypothesis that point mutations in the promoters were largely responsible for the Gal\textsuperscript{+} phenotype of the mutants, the galK promoters of CNRZ 302 and SS2 (class II mutant) were cloned in front of the gusA reporter gene in plasmids pNZ6871 and pNZ6872, respectively. Both plasmids were stable in *E. coli*, but the frequency of transformation for pNZ6872 in *S. thermophilus* ST11 was consistently 10-fold lower than for pNZ6871. Since ST11 grows very poorly on glucose, the strength of the promoters was examined in medium containing sucrose or lactose. Expression of the gusA gene results in β-glucuronidase activity, which is indicated by the development of a blue color in colonies on plates containing the substrate X-Gluc. Surprisingly, growth of *S. thermophilus* ST11 transformants harboring pNZ6871 resulted in blue colonies on X-Gluc plates containing sucrose, while on lactose plates very small and blue colonies developed, suggesting increased activity from the galK promoter. Colonies of ST11 harboring pNZ6872 developed as very small blue colonies on both sucrose and lactose plates, indicating strong activity of the class II mutant promoter on both carbon sources. However, ST11 transformed with pNZ6872 resulted in colonies that were smaller than those obtained with pNZ6871, and some white colonies also appeared. Analysis of the plasmid content of these colonies demonstrated that the pNZ6871 construct remained intact while deletion derivatives of the pNZ6872 construct were present in ST11 (data not shown). Thus, pNZ6872 is unstable, probably due to toxic effects of high β-glucuronidase activity.

**DISCUSSION**

This study demonstrates conclusively that *S. thermophilus* CNRZ 302 possesses the full complement of genes necessary for galactose metabolism despite its Gal\textsuperscript{+} phenotype. In many organisms, the *galK, galT*, and *galE* genes, which constitute the Leloir pathway of galactose metabolism, may be clustered or organized in a single operon, and the order of these genes within the operon may be highly variable. While the similarity between the deduced primary sequences of the enzymes is very high within the lactic acid bacteria, the genomic organization of the gal clusters and gene order is species specific (17). The gene order for the *S. thermophilus* galKTE operon and the divergent *galR* gene is identical to that in *S. mutans* (2), which reflects the evolutionary relationship between these bacteria. A potential transcriptional regulatory gene, *galR*, has also been identified in *L. casei* and is transcribed in the *L. casei* gal operon galKETRM (6). In contrast, the putative homologues in *E. coli*, *galR* and *galS*, are not linked to the *gal* operon (56), and regulatory genes have as yet not been identified for other *gal* operons in lactic acid bacteria such as *L. lactis*, *Leuconostoc lactis*, and *Lactobacillus helvicus*.

The effect of glucose, lactose, and galactose on the *galR* mRNA levels strongly suggested that GalR was a transcriptional regulator of the *S. thermophilus* gal operon. Transcriptional analysis of the two *galR* copies generated following disruption of *galR* in the Gal\textsuperscript{−} NZ302G indicated that the copy resulting in a C-terminally truncated GalR protein is driven by the *galR* promoter. The second copy, which lacks most of the DNA binding motif, was not transcribed. This was expected based on the orientation of the *galR* gene fragment in pNZ684. Since the substantial deletion from the C terminus includes residues contributing to inducer binding and dimerization (55), the truncated *S. thermophilus* GalR proteins are no longer functional. Northern analysis of NZ302GΔR confirmed that GalR functions as an activator of transcription for the *gal* operon and explained the inability of the disruption mutant to use galactose or the galactose moiety of lactose as a carbon source. The very low activity still detected for the GalT enzyme in NZ302GΔR might be due to a basal transcription level in the absence of GalR. Alternatively, it might reflect a low level of reversion to wild type. When the *galR* gene was provided in trans, the Gal\textsuperscript{+} phenotype was restored.

The constitutive transcription of the nonfunctional *galR* gene in the absence of GalR indicates that the latter normally functions as a negative regulator of its own expression. Auto-regulation is a common feature in prokaryotic gene regulation strategies. Negative autoregulation has been reported for other members of the LacI-GalR family such as GalS, PurR, and CytR (16, 30, 56). It is noteworthy that the majority of LacI-GalR members function as negative regulators while some
positive regulators also belong to this family and some (like CcpA) perform both functions (43, 45, 51).

The expression of the lac operon genes of both *S. thermophilus* NZ302G and ST11 is induced by growth on lactose- or galactose-containing medium. In both strains, β-galactosidase activity could no longer be induced to the usual level when the galR gene was disrupted, confirming that GalR also functions as a transcriptional activator of the lac operon. β-Galactosidase is more strongly induced in NZ302G in galactose-containing medium than in lactose-containing medium. This differential gene expression of the lac genes may be due to catabolite repression by the glucose moiety of lactose on the lac operon promoter (51), which will result in repression of the lac operon by lactose but not by galactose. Furthermore, the excretion of galactose by LacS in lactose medium would effectively reduce the availability of inducer in the cell.

The mutation to a Gal− phenotype does not result in constitutive expression of the gal genes that are induced in the presence of lactose and galactose, which strongly suggests that *S. thermophilus* was Gal− but became Gal+ in the recent past. While the advantages of the exchange reaction of the lactose transport protein offer a rationale for the observed excretion of galactose (40), the precise mechanism by which the enzymes of the Leloir pathway are suppressed has not been determined. Characterization of the 10 Gal+ mutants of CNRZ 302 revealed that a point mutation had occurred in the galK promoter region of every isolate, the majority of which were single-base insertions in a homopolymeric run of adenine residues. Interestingly, a study of the molecular basis for the adaptive response of *E. coli* populations to conditions of non-essential selection such as nutrient deprivation also identified single-base variations mainly in short mononucleotide repeats (14, 44), and slipped-strand mispairing was proposed as the responsible mechanism. In the *S. thermophilus* galK promoter, the G-to-A substitution in the class III mutant results in a −10 box (TACAAT) with greater homology to the −10 consensus (TATAAT) sequence (29). In class II mutants, the G-to-T substitution gives a TG doublet 1 bp upstream of the −10 sequence, which is a feature present in the promoters of gram-positive bacteria (12). This may correspond to the “extended −10” sequence that functions as a −35-independent promoter and requires the TG motif for efficient initiation at such promoters (24). Thus, these substitutions that resemble promoter-up mutations may increase the level of transcription of the gal genes and allow metabolism of galactose. The A insertion in class I mutants may also be a promoter-up mutation, although the reason for the enhanced activity in this case is not so apparent. The extra A increases the size of the inverted repeat preceding the −35 box from 11 to 15 nt (see below). In particular, the intrinsic DNA curvature that is predicted in this region is enhanced by the A insertion (data not shown), and this may result in increased promoter strength. The presence of curved DNA upstream of promoters, of which A tracts appear to be a major determinant, is associated with increased transcription (37).

The CNRZ 302 and SS2 galK promoter fusions to the gusA gene support the hypothesis that mutations in the galK promoter of *S. thermophilus* CNRZ 302 suppress the expression of the gal genes. Although β-glucuronidase activity was not expected from the ST11(pNZ6871) strain since galactokinase activity is barely detectable in CNRZ 302, factors such as the high copy number of this plasmid (12) and the gene dosage effect of the GalR activator are likely to be responsible. The only difference between the pNZ6871 and pNZ6872 plasmids was the G-to-T mutation in the galK promoter of the latter. Very high β-glucuronidase activity as a consequence of this promoter-up mutation would result in lethal effects on the host (P. G. G. A. de Ruyter, personal communication). This would explain the reduced frequency of transformation for pNZ6872, the small size of the blue colonies on medium containing X-Gluc, and the instability observed for this plasmid in the *S. thermophilus* host.

The intergenic region between galR and galK of *S. thermophilus* consists of 142 bp, which contains the promoter sequences of both genes in a back-to-back configuration (Fig. 4A). An 11-bp inverted-repeat sequence (IR) (nt 1392 to 1431; 5′-TTTTACTA-3′, 8 out of 11 matching nt) was detected in this region that could be an operator for GalR. The potentially global regulation by GalR prompted us to search for homologous sequences in the promoters of the lac operon and galR gene of CNRZ 302. Similar 11-bp IRs were found 13 bp upstream of the −35 box in the lac promoter and also overlapping the −10 box of the galR promoter (Fig. 4B). A consensus sequence could be deduced, with the 11-bp half of the IR consisting of a central 3 bp highly conserved portion, A(C/G)T, flanked on either side by four predominantly adenine and thymine bases. It is usually the −40 to −35 region of a promoter that is approached by an activator site; exceptions include the MerR family regulators (10, 35). In contrast, sites for repressors may be located from the −35 to −10 region. When activator proteins are used for repression, which generally occurs in cases of autoregulation, the operators often appear in positions for repression rather than activation. The potential operator sites for GalR conform to these general rules.

The gal genes of *S. mutans* are homologous to those of *S. thermophilus* and are organized in a similar divergent orientation (2). Alignment of the nucleotide sequences of the *S. thermophilus* and *S. mutans* promoter regions revealed homologous palindromic sequences, as described above (Fig. 4B). It is noteworthy, however, that a single-crossover disruption of the *S. mutans* galR gene resulted in constitutive expression of galactokinase, indicating that GalR functions as a repressor of the gal operon in this species. In contrast, GalR of *S. thermophilus* activates the gal and lac operons while repressing its own expression. The presence of these potential operators gives further credence to the hypothesis that repression of the gal operon was caused by recent mutations in the galK promoter.

**ACKNOWLEDGMENTS**

We thank R. Hutkins for the gift of the galK gene fragment from *S. thermophilus* F410, and we thank B. Poolman for communicating the unpublished galT gene sequence and for helpful discussions. We thank R. Holleman for performing the high-performance liquid chromatography assays and E. Maguin of INRA, Jouy en Josas, France, for the kind gift of pG′− host9. We are grateful to M. Kleerebezem and R. Siezen for their interest in this work and for critical reading of the manuscript.

Part of this work was funded by the BIOTECH Programmes of the European Union (contracts ERBBI02-CT92-5123 and ERBBIO04-CT96-0439) and by a grant from Ancona University for a specialization abroad to Pasquale Catzeddu.
REFERENCES


