university of groningen

University Medical Center Groningen

University of Groningen

Sugar Utilisation and Conservation of the gal-lac Gene Cluster in Streptococcus thermophilus

van den Boogaard, P.T.; Hols, P.; Kuipers, O.P.; Kleerebezem, M.; de Vos, W.M.

Published in:
Applied Microbiology

DOI:
10.1078/0723-2020-00258

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:
2004

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
van den Boogaard, P. T., Hols, P., Kuipers, O. P., Kleerebezem, M., & de Vos, W. M. (2004). Sugar Utilisation and Conservation of the gal-lac Gene Cluster in Streptococcus thermophilus. *Applied Microbiology*, *27*(1), 10 - 17. https://doi.org/10.1078/0723-2020-00258

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.

Download date: 03-06-2022

ELSEVIER

System. Appl. Microbiol. 27, 10–17 (2004)
http://www.elsevier-deutschland.de/syapm

SYSTEMATIC AND APPLIED MICROBIOLOGY

Sugar Utilisation and Conservation of the *gal-lac* Gene Cluster in *Streptococcus thermophilus*

Patrick T.C van den Bogaard[1], Pascal Hols[2], Oscar P. Kuipers[1], Michiel Kleerebezem[1], and Willem M. de Vos[1]

[1] Wageningen Centre for Food Sciences; NIZO food research, Department of Flavour and Natural Ingredients, Ede, The Netherlands
[2] Unité de Génétique, Université Catholique de Louvain, Louvain-La-Neuve, Belgium

Received: November 3, 2003

Summary

The adaptation to utilise lactose as primary carbon and energy source is a characteristic for *Streptococcus thermophilus*. These organisms, however only utilise the glucose moiety of lactose while the galactose moiety is excreted into the growth medium. In this study we evaluated the diversity of sugar utilisation and the conservation of the *gal-lac* gene cluster in a collection of 18 *S. thermophilus* strains isolated from a variety of sources. For this purpose analysis was performed on DNA from these isolates and the results were compared with those obtained with a strain from which the complete genome sequence has been determined. The sequence, organisation and flanking regions of the *S. thermophilus gal-lac* gene cluster were found to be highly conserved among all strains. The vast majority of the *S. thermophilus* strains were able to utilize only glucose, lactose, and sucrose as carbon sources, some strains could also utilize fructose and two of these were able to grow on galactose. Molecular characterisation of these naturally occurring Gal+ strains revealed up-mutations in the *galKTE* promoter that were absent in all other strains. These data support the hypothesis that the loss of the ability to ferment galactose can be attributed to the low activity of the *galKTE* promoter, probably as a consequence of the adaptation to milk in which the lactose levels are in excess.

Key words: *Streptococcus thermophilus* – lactic acid bacteria – lactose metabolism – *gal-lac* genes

Introduction

Streptococcus thermophilus is used in a variety of industrial dairy fermentations at elevated temperatures where it forms a major component in cheese starters, and, in symbiotic association with *Lactobacillus bulgaricus*, is involved in production of yogurt [1, 6, 20]. The primary function of *S. thermophilus* in these dairy fermentations is the rapid conversion of lactose into lactate but it also produces other components that may contribute to taste, aroma and texture of the fermented product. Usually this organism lacks the ability to ferment a large number of sugars that can be utilized by other lactic acid bacteria [14]. The phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS) catalyzes the transport and concomitant phosphorylation of carbohydrates and is the main sugar uptake system in lactic acid and other bacteria [4, 19]. For *S. thermophilus*, sucrose and fructose are the only sugars that are taken up by a PTS, albeit that the maximal growth rate on these sugars is lower than that on lactose. Remarkably, glucose is a non-PTS sugar for *S. thermophilus* and the maximal growth rate on glucose is several-fold lower than that on lactose [16, 25]. *S. thermophilus* is highly adapted to grow on lactose although this sugar is not transported by PTS but by the efficient dedicated permease LacS [5]. The *S. thermophilus lacSZ* operon codes for this lactose permease (*lacS*) as well as the β-galactosidase (*lacZ*) that is involved in the hydrolysis of the internalised lactose. The transcription of the *lacSZ* operon is induced during growth on lactose and subject to catabolite control by CcpA, resulting in fine-tuning of lactose uptake and hydrolysis to subsequent catabolism via glycolysis [18, 25, 26]. Most *S. thermophilus* strains used as industrial starters do not utilise galactose and only ferment the glucose moiety of lactose, while the galactose moiety is excreted into the medium in stoichiometric amounts relative to the lactose taken up [10]. Recently, the genes for galactose utilisation and its control were found to be organised in the *galR-galKTEM* gene cluster that was identified upstream of the *S. thermophilus lac* operon in strain CNRZ302 [26]. The same study also revealed that the promoters of the *galKTE* as well as the *lacSZ* operons were induced by the operon-specific regulator GalR upon growth on lactose. Howev-

0723-2020/04/27/01-10 $ 30.00/0

er, the induced activity of the *galKTE* promoter was very low which was also reflected in the activities of the enzymes encoded by these genes. Galactose-fermenting mutants were isolated from *S. thermophilus* CNRZ302 that all showed strong *galKTE* expression upon induction due to insertions or substitutions in the promoter region [26].

This study was undertaken to investigate the diversity of sugar utilisation and the genetic conservation of the *gal-lac* gene cluster in a collection of *S. thermophilus* strains isolated from various sources. Sugar utilisation was found to be restricted to glucose, lactose, sucrose and fructose in most strains, while only two strains were able to ferment galactose. The genomic sequence of *S. thermophilus* LMG18311 [8] allowed detailed genetic analysis of the *gal-lac* locus and its position in the chromosome. The genetic organisation of the *gal-lac* gene cluster was conserved in all strains tested as were the surrounding regions, suggesting conservation of its chromosomal location. Genetic analysis of the divergent *galR-galKTE* promoter region showed that the natural galactose-fermenting *S. thermophilus* isolates contained *galKTE* promoter-up mutations that are to likely mediate the galactose-fermenting phenotype of *S. thermophilus*.

Materials and Methods

Bacterial strains, plasmids, media and culture conditions

The strains and plasmids used in this study are listed in Table 1. *S. thermophilus* was routinely grown at 42 °C in M17 broth (Difco, Surrey, U.K.) containing 1% (w/v) of the chosen carbon source unless stated otherwise. *Escherichia coli* HB101 [21] was grown in TY broth with aeration at 37 °C, and for plasmid selection ampicillin (Ap) 50 µg/ml was used.

Sequencing of the 16S rRNA variable regions

Species verification of the *S. thermophilus* strains used in this study was carried out by sequencing the 16S rRNA variable regions as described previously [23].

Table 1. Strains used in this study.

S. thermophilus strain	Relevant features	Reference or source
ATCC 19258	Type-strain, pasteurized milk	ATCC [7]
LMG18311	Yogurt starter	LMG [8]
CNRZ302	Yogurt starter	CNRZ [2]
NZ302G	Gal$^+$ derivative of CNRZ302	[26]
SS2	Gal$^+$ derivative of CNRZ302	[26]
SS7	Gal$^+$ derivative of CNRZ302	[26]
A054	Yogurt starter	[22]
ST8	Proteolytic starter	LODI, Italy
ST11	Commercial starter	[15]
NIZO B110	Yogurt	NIZO
NIZO B112	Pasteurized milk	NIZO
NIZO B113	Pasteurized milk	NIZO
NIZO B126	Milk	NIZO
NIZO B132	Milk	NIZO
NIZO B1122	Danish yogurt starter	NIZO
NCIMB 701968	Bulgarian cheese starter	NCIMB
NCIMB 702533	Italian cheese	NCIMB
NCIMB 702560	Emmental starter	NCIMB
NCIMB 702561	Calf rennet	NCIMB
NCIMB 702562	Gorgonzola cheese	NCIMB
NCIMB 702637	Gruyère starter	NCIMB

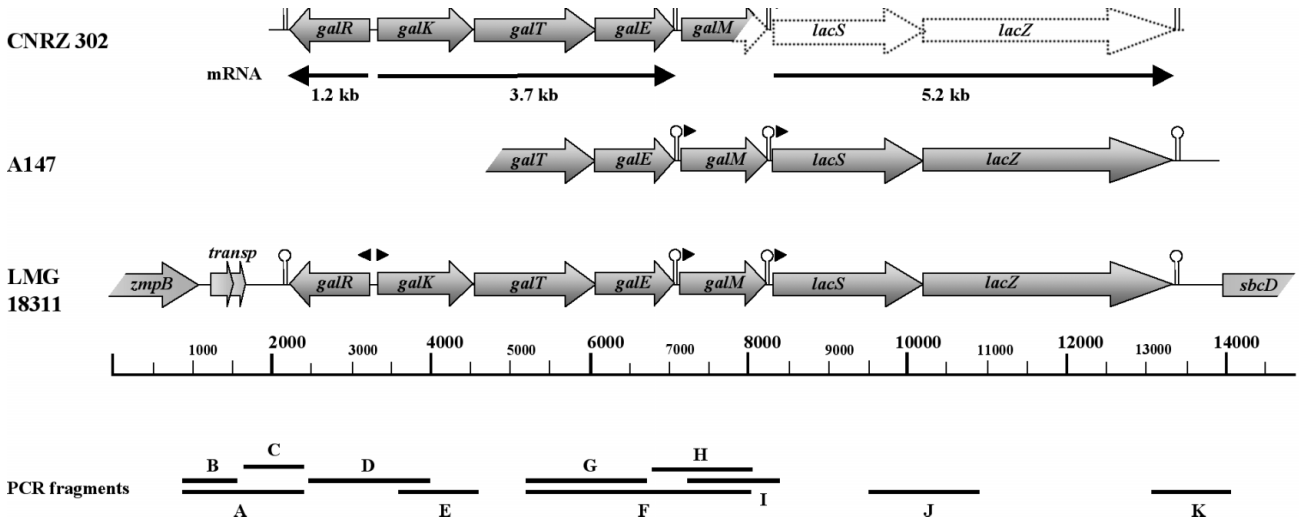

Fig. 1. Organization of (partialy) sequenced *S. thermophilus gal-lac* gene clusters. A: Open arrows indicate genes, promoters are denoted by black arrowheads and the terminators are indicated by stem-loop structures. For strain CNRZ302 the presence of the *galM-lacSZ* genes downstream of the complete *galE* was confirmed by long-range PCR and partial sequence analysis. The transcripts of this *gal-lac* gene cluster are denoted underneath as arrows (26). Underneath the LMG18311 *gal-lac* gene cluster (and surrounding regions) the nucleotide scale is presented (bp) together with the PCR fragments generated in this study with the oligonucleotides from Table 2.

DNA manipulations and transformations

Transfer to and isolation of plasmid DNA from *E. coli* was performed using established protocols [21]. Plasmid and chromosomal DNA from *S. thermophilus* was isolated as described previously for *L. lactis* [27]. All DNA-modifying enzymes were used as recommended by the suppliers (Gibco/BRL, Boehringer Mannheim or Roche). DNA fragments were recovered from agarose gels using the Glass Matrix DNA Isolation System (Gibco/BRL).

Organisational analysis of the *S. thermophilus gal-lac* gene cluster and surrounding regions

A PCR-based approach was used in to determine the organisation of *gal-lac* gene cluster and its surrounding regions in the different laboratory and starter strains of *S. thermophilus*. Primer combinations were selected from the oligonucleotides (Table 2) that allowed the organisational analysis of the *gal-lac* gene cluster in such a manner that the position of each gene and the length of the intergenic regions could be determined from the generated PCR products using Taq polymerase (Gibco/BRL) (Fig. 1). The oligonucleotides were designed based on the *gal-lac* operon sequence from strain LMG18311 [8].

Cloning and sequence analysis of the *galK* promoter regions

Oligonucleotides galR1 and 81 (Table 2) were used in a PCR reaction using high fidelity Pwo polymerase (Roche) to amplify the *galR-galK* promoter regions with chromosomal DNA isolated from the different *S. thermophilus* strains as template. The 0.3-kb PCR fragments obtained were cloned in the pGEM-T vector (Promega) and subsequently both strands were sequenced using an ALF DNA sequencer (Pharmacia Biotech) according to protocol recommended by the supplier. These *galR-galK* promoter regions were compared in an DNA sequence alignment to the *galR-galK* promoter region of CNRZ302 and its galactose-fermenting mutants [26].

Growth and sugar utilization

Overnight cultures grown in glucose M17 medium were diluted to an OD600 of 0.05 in fresh medium supplemented with 1% (w/v) of the appropriate sugar, and growth was followed spectrophotometrically. All strains were subjected to an API-50 test to evaluate the sugar fermentation ability according to the protocol recommended by the supplier (Biomerieux, France).

Nucleotide sequence accession numbers

The *gal/lac* sequence data have been submitted to the GenBank database under accession number AF503446.

Results

S. thermophilus gal-lac operon sequence comparison

The organisation and regulation of the *gal* and *lac* genes has been studied in considerable detail in the *S. thermophilus* strains CNRZ302 and A147, respectively [17, 18, 26]. However, the availability of the genome sequence of *S. thermophilus* LMG18311 allowed for the complete annotation of the *gal-lac* gene cluster [8]. The *gal-lac* gene cluster is located at approximately 1.24 Mb relative to the predicted origin of the 2.0 Mb genome of this yoghurt starter and resides on the minus strand. This cluster is flanked at the upstream side by two small incomplete and partially overlapping ORF's, predicted to code for products with high degree of homology (82%) to a putative IS*861* transposase from *S. agalacticae* (Fig. 1). Further upstream, an ORF was identified that could encode a product with strong similarity (77%) with the zinc-metalloprotease encoded by the *zmpB* gene of *S. pneumoniae*. Downstream of the *gal-lac* cluster an ORF was located, predicted to code for a product with high similarity (70%) to an exonuclease, encoded by the *sbcD* gene of *Lactococcus lactis*.

The comparison of the complete or partial *gal-lac* sequence information of the strains LMG18311,

Table 2. Oligonucleotides used in this study.

Primer	Sequence (5′- 3′)	Strand	Position	Used for PCR products
MetalF	TGACGACAGTTGCAATTGACGG	C	839–860	A, B
TransR	CATGCCATTGTCTGGACGATTCC	NC	1564–1542	B
TransF	CGTCCAGACAATGGCATGATGG	C	1547–1568	C
GalRR	CTCTTTTAACGATACAACTCTAGC	NC	2392–2369	A, C
41	GTCGTCAGGGACCTTG	C	2404–2419	D
GalR1	GTTGAAATAGATACACCTGC	C	3154–3174	*galRK* promoter region
81	TAGTCCGTATGCTCACCAATCA	NC	3456–3435	*galRK* promoter region
42	CTTGCAAGAAGCTGGGC	C	3649–3665	E
12	CGACATTATCCTTGAGGTC	NC	3968–3950	D
88	GATACGGTCCAATTCCTC	NC	4607–4590	E
GalT4	CTCCTAACAATTGTTGAAGC	C	5241–5260	G, F
GalER	TACGAAGATATTCAACAGCAAGG	NC	6773–6751	G
GalEF	GCCTTAGACTTAGGTTCATCAACAGG	C	6791–6816	H
GalMF	CAATGGAAACTGGTGCTC	C	7245–7262	I
GalMR	TCCATAGCAATAGCTTCACG	NC	8046–8027	H, F
LacS2R	GATAAAGTATGTTGACAAGG	NC	8362–8343	I
LacS2F	CATTTGCCGTAGCTGCCGG	C	9471–9489	J
LacZR	CAAAGAGATCTTGAACGTGAAC	NC	10850–10829	J
ExoF	GCAAGAATCTGATGCTACATGG	C	13102–13123	K
ExoR	CGTCCTACATGCCAATCTGACG	NC	14143–14122	K

CNRZ302 and A147 revealed that their overall sequence similarity was very high, while the organisation of the *gal-lac* genes was found to be identical (Fig. 1). This conserved organisation was also found in the recently reported sequence of *S. thermophilus* strain SMQ-301 [24]. Moreover, the G+C content of the LMG18311 *gal-lac* gene cluster matches that of the whole genome, suggesting that these important catabolic genes were not recently acquired via horizontal gene transfer. However, detailed comparisons showed considerable difference in degree of sequence similarities between the various individual genes of the *gal-lac* gene cluster (Table 3). While the regulatory gene *galR* was highly conserved and showed no substitutions at all, the enzyme-coding genes showed larger diversity. This diversity was found to be largest in the *galE* (22/1010 nucleotides) and *lacS* genes (16/1904 nucleotides). Remarkably, the *galE* sequence of strain A147 contained 3 nucleotides less than that of the other strains, resulting in two frame-shifts. This generated a new sequence of 11 amino acids (from residue 187 to 188) in the predicted GalE protein of strain A147. It remains to be established whether this frameshift is the result of sequencing errors since the A147 GalE was found to be functional [17]. Apart from these unique amino acids, all other substitutions in GalE appeared to be rather uniformly distributed. This contrasts with the substitutions in the LacS protein sequence, many of which were found to be located in the hydrophobic membrane spanning regions while none were located in the regulatory IIA domain. Not only the regulatory gene *galR* but also all regulatory sequences showed little sequence variation. The *galR-galK* promoter region was found to be identical in the strains LMG18311, CNZ302 and SMQ-301. Furthermore, the *lacS* promoter region was identical in all strains and contained an almost perfect *cre* site that has been implicated in catabolite repression [16, 25]. A single nucleotide substitution was present in the putative ribosome binding site (RBS) of A147 *LacS* relative to LMG18311 and SMQ-301, making it less homologous to postulated consensus RBS sequences in LAB [3]. These results indicate that not only the organisation of the *gal-lac* gene cluster is very well conserved but notably all regulatory regions.

Sugar utilisation of different *S. thermophilus* strains

To investigate the degree of conservation in the *gal-lac* gene cluster and correlate this with the capacity to utilise lactose, galactose and other sugars, we screened a set of 18 *S. thermophilus* strains from a variety of sources (Table 1). To verify the taxonomic identity of these strains the V1, V2 and V3 variable regions of 16S rRNA were sequenced and compared to known sequences in databases (data not shown). The obtained 16S rRNA sequence were completely identical, confirming that all strains belonged to the species *S. thermophilus*. In a growth assay all strains tested were able to utilise glucose, lactose and sucrose, and four were able to utilise fructose (Table 4). Besides the earlier described Gal+ mutant of CNRZ302, NZ302G [26], NIZO B112 and NIZO B113 were the only strains that showed significant growth on galactose, as was also suggested by the API-50 CH assay

Table 3. List of nucleotide mismatches and amino acid substitutions from four (partial) sequenced *S. thermophilus gal-lac* gene clusters using the cluster of LMG18311 as template. All shown nucleotide mismatches and amino acid substitutions were also found in the reported sequence of *S. thermophilus* SMQ-301 (24) except for these indicated with an asterix. Conservative amino acid changes (score ≥ 0 in Blusum62 matrix) are underlined. A minus indicates the genes of which no sequence data was available.

LMG18311 template		A147		CNRZ302	
gene	gene size (bp)	nucleotide mismatches	AA subst	nucleotide mismatches	AA subst
galR	995	–	–	2	0
galK	1166	–	–	3	A290T *
galT	1481	–	–	7	V42I*, N71K, A243V*, E314G*, N476S, D480A*
galE	1010	22	I100V, T238A, L251N, D252N, C269R, L284R, R285P, F302L, D312E, N313K, N324S	–	–
galM	1046	6	T55M*, E101D, E125D, I174V	–	–
lacS	1904	16	F105L, L107F, I110L, D183V, L185F, H190N, Y203F, D208A, N210I, H230R, M251L,H547R	–	–
lacZ	3080	4	H284Q, T964S*, Q973H*	–	–

(data not shown). Remarkably, these natural Gal$^+$ strains were also capable of utilising fructose (Table 4). For most of these strains tested the results of the API-50 CH assay could be confirmed by growth on a specific sugar. However, strain NIZO B112 showed reproducible acidification of the API-50 CH indicator medium containing maltose, mannose or cellobiose, while no significant growth was observed on these sugars. Acidification in absence of growth could be due to the presence of transport proteins and metabolic pathways with insufficient efficiency or specificity. These results indicate that all *S. thermophilus* strains are able to utilize glucose, lactose and sucrose, except for strain NCIMB 701968 that showed no growth on sucrose.

Organisation and location of the *S. thermophilus gal-lac* operon

S. thermophilus is highly adapted to lactose fermentation. However, only few strains are able to ferment the galactose moiety of lactose. We used a diagnostic PCR-based approach to assess the conservation of the *gal-lac* gene cluster organisation in *S. thermophilus* (Fig. 1). Primers were designed based on strategic regions in the cluster and flanking sequences that allowed the determination of the position of each gene and the length of the intergenic regions from the generated PCR products. The

Table 4. Growth of the strains used in this study on various sugars (1% (w/v), in M17 medium). Full growth (OD600nm <1): +, No growth: –, In some cases significant growth was observed but the final OD600nm was between 0.3 and 0.6; this growth was scored at +/–.

Strain	Glucose	Lactose	Sucrose	Galactose	Fructose
LMG18311	+	+	+	–	–
CNRZ302	+	+	+	–	+/–
NZ302G	+	+	+	+	+/–
ATCC 19258	+	+	+	–	+/–
A054	+/–	+	+	–	–
ST8	+	+	+	–	–
ST11	+	+	+	–	–
NIZO B110	+	+	+	–	–
NIZO B112	+	+	+	+	+
NIZO B113	+	+	+	+	+
NIZO B126	+	+	+	–	–
NIZO B132	+	+	+	–	–
NIZO B1122	+	+	+	–	–
NCIMB 701968	+/–	+	–	–	–
NCIMB 702533	+	+	+	–	–
NCIMB 702560	+	+	+	–	–
NCIMB 702561	+	+	+	–	–
NCIMB 702562	+	+	+	–	–
NCIMB 702637	+	+	+	–	–

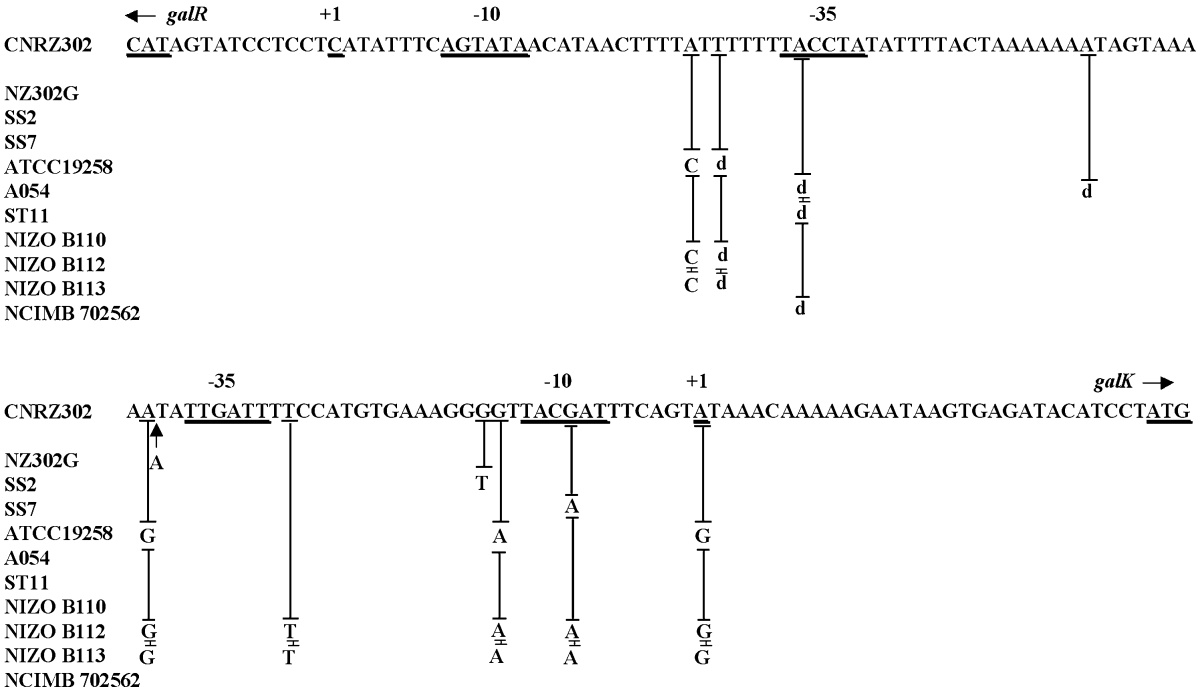

Fig. 2. Sequence of the *galR-galK* intergenic region. The intergenic region containes the promoter sequences for the *galK* and *galR* genes from the strains used in this study are shown as deviations from the CNRZ302 sequence. The –10 and –35 regions of the promoter are underlined, and the transcriptional start sites are indicated as +1. The nucleotide mismatches (relative to the CNRZ302 sequence) are depicted underneath. Nucleotide insertion is indicated with a vertival arrow and nucleotide deletions are indicated with d.

anticipated PCR products (fragments D to J in Fig. 1) were readily generated from all strains, indicating that in all the *gal-lac* gene order (*galKTEM-lacSZ*) is conserved. Furthermore, no differences in length of these PCR products could be detected on agarose gels, indicating that individual gene sizes were highly similar.

To analyse the location of the *gal-lac* cluster in the genome of the various *S. thermophilus* strains, we amplified its flanking regions. In all *S. thermophilus* strains, the *gal-lac* cluster was found to be flanked by the same genes (*zmpB*-transposase and *sbcD* homologues). Only for strain NIZO B126, the PCR fragments A and B (Fig. 1) were similarly increased in size by approximately 0.5 kb, indicating a small insertion between the *zmpB* and the transposase homologues in this strain. These results indicate that the organisation of the *S. thermophilus gal-lac* gene cluster as well as the flanking regions of this locus in the genome of *S. thermophilus* are very well conserved among strains isolated from different sources. This also holds for the galactose-fermenting strains NIZO B112 and NIZO B113.

Sequence analysis of the *galR-galK* promoter region

To study the relation between the natural ability to ferment galactose of strains NIZO B112 and NIZO B113 and the previously identified promoter-up mutations within the *galK* promoter [26] the *galR-galK* intergenic regions of these and the other *S. thermophilus* strains were amplified by PCR, cloned in pGEM-T, and characterised by sequence analysis (Fig. 2). The *galR-galK* promoter sequences of 10 of the analysed strains were completely identical to that of LMG18311 and the rest showed in between one to five nucleotide mismatches or single nucleotide deletions. The promoter sequences of NIZO B112, NIZO B113 and ATTC19258 showed several substitutions relative to LMG18311 (Fig. 2). Common substitutions were a single base pair deletion in the proposed –35 sequence of the *galR* promoter, and an A to G substitution at the predicted transcript initiation site *galK*. In contrast, only the galactose-fermenting strains NIZO B112 and NIZO B113 contained a G to A substitution in the –10 sequence. In analogy, an identical substitution in the –10 sequence of the *galKTE* promoter was identified in one class of galactose-fermenting mutants of CNRZ302 that conformed better to consensus –10 sequences known [13, 26]. These results suggest that this optimised –10 promoter region is responsible for the observed ability to ferment galactose. Various attempts to validate this suggestion by measuring the activity of the NIZO B112, and NIZO B113 derived *galK* promoters using established promoter probe vectors for lactic acid bacteria failed since the resulting constructs appeared to be highly unstable in *S. thermophilus*. Similar results have been described for the promoter-up *galK* promoter variants isolated from galactose fermenting CNRZ302 cells [26]. Taken together, these results strongly suggest that an optimised *galK* promoter sequence most likely conferred the ability to ferment galactose to NIZO B112 and NIZO B113.

Discussion

Characteristic for *S. thermophilus* is the adaptation to lactose as primary carbon and energy source and the inability to ferment a large number of sugars that can be utilised by other lactic acid bacteria [14]. Moreover, although *S. thermophilus* readily ferments lactose, most strains can not utilise galactose, neither the free sugar nor the moiety derived from lactose hydrolysis by β-galactosidase activity [10]. In this study we used a combination of diagnostic PCR, Southern hybridisation and nucleotide sequence analysis to evaluate the diversity of sugar utilisation and the conservation of *gal-lac* gene cluster in a collection of 18 *S. thermophilus* strains isolated from different sources.

The carbon sources that most *S. thermophilus* strains are able to utilise are limited to glucose, lactose, and sucrose, while some strains can also grow on fructose. If utilised, glucose and fructose are relative poor carbon sources on which the growth rates are several-fold lower than that on lactose or sucrose, which are readily metabolised [16, 25]. Although fructose is taken up by a PTS, utilisation of this sugar as a carbon source requires fructose-bisphosphatase (FBPase) activity to form essential biomass precursors such as nucleotide sugars. In the genome sequence of *S. thermophilus* LMG18311 no homologue for a FBPase encoding gene could be found [8]. It could very well be that FBPase activity levels in most *S. thermophilus* strains are limiting since a lack of this activity in *L. lactis* uncoupled fructose fermentation from growth [12]. Strains NIZO B112 and B113 were able to utilise galactose and fructose as sole carbon source while only NIZO B112 was able to also convert maltose and cellobiose into acids, although these carbon sources did not support growth.

In many organisms the *galK*, *galT* and *galE* genes that encode the enzymes of the Leloir pathway of galactose metabolism are clustered or organised in a single operon, but their order may vary. The organisation of the *gal-lac* genes appears to be identical in all *S. thermophilus* strains tested in this study. Moreover, the flanking regions of the *S. thermophilus gal-lac* cluster appeared to be conserved among the tested strains reflecting conservation for that chromosomal locus. The remnant of one or more transposase-like genes upstream of the *gal-lac* cluster indicates the insertion and loss of a transposable element early in *S. thermophilus* evolution, which did not disrupt the operon itself. At least 13 of these transposase remnants with a homology of 86% or higher (at the amino acid level) can be found in the genome sequence of strain LMG18311 [8]. The gene order for the *S. thermophilus gal-lac* operon and the divergently encoded *galR* gene is identical to that of *S. salivarius*, which reflects the evolutionary relationship between these bacteria [24]. However, the *S. salivarius gal-lac* operon has different flanking regions compared to those of *S. thermophilus* indicating that the *gal-lac* gene cluster has evolved as one single unit in these phylogenetically closely related streptococci. Alignment of all reported *S. thermophilus gal-lac* operon sequences showed that the overall homology was highest between SMQ-301, A147

and CNRZ302 relative to LMG18311. The observed substitutions were randomly distributed over the *gal-lac* cluster and although the substitution frequency was higher in the *galE* and the *lacS* genes, no significant hotspots could be identified. The regulatory gene *galR* and all regulatory sequences showed no sequence variations.

Two naturally isolated strains NIZO B112 and NIZO B113 were able to ferment and use galactose for growth. Apparently, organisational changes are not responsible for conferring the ability for galactose utilisation to *S. thermophilus*. Only limited sequence variation was found between the complete *galRK* intergenic regions of the Gal- strains used in this study. However, in the galactose-fermenting strains NIZO B112 and NIZO B113 a single nucleotide substitution in the –10 promoter sequence was found that most likely results in an improved *galK* promoter activity. The finding that an identical substitution was identified in class III of the previously described *galK* promoter-up mutations derived from galactose fermenting-mutants of CNRZ302 [26] corroborates this explanation for the observed galactose fermenting phenotype.

In conclusion, lactose is the preferential sugar in *S. thermophilus* and due to continued growth on lactose-containing media such as milk, this organism has probably lost the ability to utilize a number of PTS and non-PTS sugars that can be fermented by other lactic acid bacteria. The *S. thermophilus gal-lac* operon is highly conserved in sequence, organisation and genomic location among strains isolated from a variety of sources. Galactose-fermenting mutants have lower growth-rates when growing on lactose compared to the Gal- parental strains [11]. Apparently, the fermentation of the galactose moiety via the relatively slow Leloir pathway is energetically less favourable and will lead to the outgrowth of the Gal- variants. The loss of the ability to ferment galactose can be attributed to the adaptation to milk in which the lactose levels are in excess. Apparently, the galactose-fermenting strains NIZO B112 and NIZO B113 that were fortuitously found in milk have not adapted as strongly to growth on lactose and are able to utilise additional sugars. At the molecular level this the adaptation to lactose is reflected not only by the highly efficient lactose/galactose exchange activity of the galactoside transporter LacS but also by the low activity of the *galKTE* promoter, due to accumulation of specific nucleotide substitutions that have been identified here.

Acknowledgements

We thank Arjen Wagendorp and Jan Hoolwerf for the *S. thermophilus* 16S rRNA sequencing. This work was partially supported by the BIOTECH Program of the European Community (contract no ERBBI04-CT96-0439 and ERBBI04-CT96-0498).

References

1. Auclair, J., and J. P. Accolas: Use of thermophilic lactic starters in the dairy industry. Antonie Van Leeuwenhoek 49, 313 (1983).
2. Benataya, A., P. Bracquart, and G. Linden: Galactose-fermenting mutants of *Streptococcus thermophilus*. Can. J. Microbiol. 35, 136–140 (1991).
3. De Vos, W. M., and G. Simons: Molecular cloning of lactose genes in dairy lactic streptococci: the phospho-beta-galactosidase and beta-galactosidase genes and their expression products. Biochimie 70, 461 (1988).
4. de Vos, W. M., and E. E. Vaughan: Genetics of lactose utilization in lactic acid bacteria. FEMS Microbiol. Rev. 15, 217 (1994).
5. Foucaud, C., and B. Poolman: Lactose transport system of *Streptococcus thermophilus*. Functional reconstitution of the protein and characterization of the kinetic mechanism of transport. J. Biol. Chem. 267, 22087 (1992).
6. Giraffa, G., A. Paris, M. Gatti, and E. Neviani: Genotypic and phenotypic heterogeneity of *Streptococcus thermophilus* strains isolated from dairy products. J. Appl. Microbiol. 91, 937–943 (2001).
7. Herman, R. E., and L. L. McKay: Cloning and expression of the beta-D-galactosidase gene from *Streptococcus thermophilus* in *Escherichia coli*. Appl. Environ. Microbiol. 52, 45 (1986).
8. Hols, P., D. Prozzi, S. Burteau, K. Nagui, F. Hancy, T. Ferain, and J. Delcour: The complete *Streptococcus thermophilus* LMG18311 genome sequence. www.biol.ucl.ac.be/gene/genome (2002).
9. Hutkins, R., H. A. Morris, and L. L. McKay: Galactose transport in *Streptococcus thermophilus*. Appl. Environ. Microbiol. 50, 772 (1985).
10. Hutkins, R., and H. A. Morris: Carbohydrate metabolism by *Streptococcus thermophilus*: a review. J. Food. Prot. 50, 876 (1987).
11. Levander, F., M. Svensson, and P. Radstrom: Enhanced exopolysaccharide production by metabolic engineering of *Streptococcus thermophilus*. Appl. Environ. Microbiol. 68, 784 (2002).
12. Looijesteijn, P. J., I. C. Boels, M. Kleerebezem, and J. Hugenholtz.: Regulation of exopolysaccharide production by *Lactococcus lactis* subsp. *cremoris* By the sugar source. Appl. Environ. Microbiol. 65, 5003 (1999).
13. McClure, W. R.: Mechanism and control of transcription initiation in prokaryotes. Annu. Rev. Biochem. 54, 171 (1985).
14. Mercenier, A.: Molecular genetics of *Streptococcus thermophilus*. FEMS Microbiol. Rev. 7, 61 (1990).
15. Mollet, B., J. Knol, B. Poolman, O. Marciset, and M. Delley: Directed genomic integration, gene replacement, and integrative gene expression in *Streptococcus thermophilus*. J. Bacteriol. 175, 4315 (1993).
16. Poolman, B.: Energy transduction in lactic acid bacteria. FEMS Microbiol. Rev. 12, 125 (1993).
17. Poolman, B., T. J. Royer, S. E. Mainzer, and B. F. Schmidt: Carbohydrate utilization in *Streptococcus thermophilus*: characterization of the genes for aldose 1-epimerase (mutarotase) and UDPglucose 4-epimerase. J. Bacteriol. 172, 4037 (1990).
18. Poolman, B., T. J. Royer, S. E. Mainzer, and B. F. Schmidt: Lactose transport system of *Streptococcus thermophilus*: a hybrid protein with homology to the melibiose carrier and enzyme III of phosphoenolpyruvate-dependent phosphotransferase systems. J. Bacteriol. 171, 244 (1989).
19. Postma, P. W., J. W. Lengeler, and G. R. Jacobson: Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. Microbiol. Rev. 57, 543 (1993).
20. Randazzo, C. L., S. Torriani, A. D. Akkermans, W. M. de Vos, and E. E. Vaughan: Diversity, dynamics, and activity of

bacterial communities during production of an artisanal Sicilian cheese as evaluated by 16S rRNA analysis. Appl. Environ. Microbiol. 68, 1882 (2002).
21. Sambrook, J., E. F. Fritsch, and T. Maniatis: Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).
22. Schroeder, C. J., C. Robert, G. Lenzen, L. L. McKay, and A. Mercenier: Analysis of the *lacZ* sequences from two *Streptococcus thermophilus* strains: comparison with the *Escherichia coli* and *Lactobacillus bulgaricus* beta-galactosidase sequences. J. Gen. Microbiol. 137, 369 (1991).
23. te Giffel, M. C., R. R. Beumer, N. Klijn, A. Wagendorp, and F. M. Rombouts: Discrimination between *Bacillus cereus* and *Bacillus thuringiensis* using specific DNA probes based on variable regions of 16S rRNA. FEMS Microbiol. Lett. 146, 47 (1997).
24. Vaillancourt, K., S. Moineau, M. Frenette, C. Lessard, and C. Vadeboncoeur: Galactose and Lactose Genes from the Galactose-Positive Bacterium *Streptococcus salivarius* and the Phylogenetically Related Galactose- Negative Bacterium *Streptococcus thermophilus*: Organization, Sequence, Transcription, and Activity of the *gal* Gene Products. J. Bacteriol. 184, 785 (2002).
25. van den Bogaard, P. T., M. Kleerebezem, O. P. Kuipers, and W. M. de Vos: Control of lactose transport, beta-galactosidase activity, and glycolysis by CcpA in *Streptococcus thermophilus*: evidence for carbon catabolite repression by a non-phosphoenolpyruvate-dependent phosphotransferase system sugar. J. Bacteriol. 182, 5982 (2000).
26. Vaughan, E. E., P. T. van den Bogaard, P. Catzeddu, O. P. Kuipers, and W. M. de Vos: Activation of silent *gal* genes in the *lac-gal* regulon of *Streptococcus thermophilus*. J. Bacteriol. 183, 1184 (2001).
27. Vos, P., M. van Asseldonk, F. van Jeveren, R. Siezen, G. Simons, and W. M. de Vos: A maturation protein is essential for production of active forms of *Lactococcus lactis* SK11 serine proteinase located in or secreted from the cell envelope. J. Bacteriol. 171, 2795.

Corresponding author:
Michiel Kleerebezem, NIZO food research, Department of Flavour and Natural Ingredients, P.O. Box 20, 6710 BA Ede, The Netherlands
Tel.: ++31-318 659511; Fax: ++31-318 650400;
e-mail: michiel.kleerebezem@nizo.nl