APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Apr. 2001, p. 1522–1528 Vol. 67, No. 4
0099-2240/01/$04.00+0 DOI: 10.1128/AEM.67.4.1522–1528.2001

Characterization of a Novel Type II Restriction-Modification System, *Sth*368I, Encoded by the Integrative Element ICE*St1* of *Streptococcus thermophilus* CNRZ368

VINCENT BURRUS, CYRIL BONTEMPS, BERNARD DECARIS,* AND GÉRARD GUÉDON

Laboratoire de Génétique et Microbiologie (INRA UA952), Faculté des Sciences, Université Henri Poincaré (Nancy 1), 54506 Vandoeuvre-lès-Nancy, France

Received 8 August 2000/Accepted 7 January 2001

A novel type II restriction and modification (R-M) system, *Sth*368I, which confers resistance to ϕST84, was found in *Streptococcus thermophilus* CNRZ368 but not in the very closely related strain A054. Partial sequencing of the integrative conjugative element ICE*St1*, carried by *S. thermophilus* CNRZ368 but not by A054, revealed a divergent cluster of two genes, *sth368IR* and *sth368IM*. The protein sequence encoded by *sth368IR* is related to the type II endonucleases R.*Lla*KR2I and R.*Sau*3AI, which recognize and cleave the sequence 5′-GATC-3′. The protein sequence encoded by *sth368IM* is very similar to numerous type II 5-methylcytosine methyltransferases, including M.*Lla*KR2I and M.*Sau*3AI. Cell extracts of CNRZ368 but not A054 were found to cleave at the GATC site. Furthermore, the C residue of the sequence 5′-GATC-3′ was found to be methylated in CNRZ368 but not in A054. Cloning and integration of a copy of *sth368IR* and *sth368IM* in the A054 chromosome confers on this strain phenotypes similar to those of CNRZ368, i.e., phage resistance, endonuclease activity of cell extracts, and methylation of the sequence 5′-GATC-3′. Disruption of *sth368IR* removes resistance and restriction activity. We conclude that ICE*St1* encodes an R-M system, *Sth*368I, which recognizes the sequence 5′-GATC-3′ and is related to the *Sau*3AI and *Lla*KR2I restriction systems.

Streptococcus thermophilus is extensively used as a starter in the manufacture of cheese and yogurt with other lactic acid bacteria, like *Lactococcus lactis* or *Lactobacillus delbrueckii* subsp. *bulgaricus*. The proliferation of bacteriophages is one of the main reasons for the failure of these fermentation processes. Since it is difficult to avoid contamination, the strains used as starters should be highly resistant to a large array of phages. In the best known lactic acid bacterium, *L. lactis*, four types of natural defense mechanisms against bacteriophages have been identified on the basis of their modes of action: blocking of phage adsorption, blocking of phage DNA penetration, abortive infection, and restriction-modification (R-M) systems (11). In this species, the resistance is generally encoded by plasmids, and several different mechanisms can be carried on one plasmid (18). The genes of 10 R-M systems have been cloned from *L. lactis* strains: eight of the systems are encoded by plasmids, and only two are encoded by the chromosome (11). Some of these plasmids, like pTR2030, are conjugative, allowing easy introduction by conjugative transfer into phage-sensitive strains of commercial importance. The resulting strains have been used successfully by the dairy industry (1, 33).

In contrast, very few phage defense mechanism have been described in *S. thermophilus*. This could be due to the scarcity of plasmids in this species and/or to the more recent progress in its genetics. Most of the strains of *S. thermophilus* appear to be plasmid free except for a few isolates that contain a single relatively small plasmid (25). None is conjugative. Four type II R-M systems have been well characterized in *S. thermophilus*: *Sth*134I (35) is an isoschizomer of *Hpa*II and *Msp*I, and *Sth*117I (36), *Sth*455I (15), and *Ssl*I (3) are isoschizomers of *Bst*NI and *Eco*RII. However, their genes have been neither cloned nor sequenced.

A site-specific integrative element, ICE*St1*,was found to be integrated in the 3′ end of *fda* of *S. thermophilus* CNRZ368, an open reading frame (ORF) encoding a putative fructose-1,6-bisphosphate aldolase (6). It excises by site-specific recombination. Partial sequencing of the right end of this element reveals ORFs encoding proteins related to those of some conjugative plasmids and conjugative transposons. Therefore, ICE*St1* could be an integrative conjugative element.

The results presented in this study show that ICE*St1* carries the genes encoding a type II R-M system, *Sth*368I, which recognizes the sequence 5′-GATC-3′. These genes were cloned and sequenced. They are related to those encoding *Lla*KR2I of *L. lactis* (38) and *Sau*3AI of *Staphylococcus aureus* (34), two type II R-M systems which also recognize GATC sequences.

MATERIALS AND METHODS

Bacterial strains and media. The *Escherichia coli, S. thermophilus*, and *L. lactis* strains used in this study are listed in Table 1. *E. coli* strains were grown at 37°C on Luria-Bertani medium supplemented with 170 μg of chloramphenicol/ml [strains containing pBC KS(+) (Stratagene, La Jolla, Calif.)]-derived plasmids], 50 μg of ampicillin/ml [strains containing pBluescript SK(−) (Stratagene)-derived plasmids], or 150 μg of erythromycin/ml (strains containing pG+Host9-derived plasmids). The *S. thermophilus* strains were grown at 42°C in M17 broth containing 5 g of lactose/liter (M17L) supplemented when appropriate with 2 μg of erythromycin/ml (strains containing integrated pG+Host9-derived plasmids) or at 30°C in M17L broth containing 5 μg of erythromycin/ml (strains containing free pG+Host9-derived plasmids). *L. lactis* MG1363 was grown at 30°C in M17 broth containing 0.2 M glucose (M17G) supplemented with 5 μg of erythromycin/ml (strains containing pG+Host9-derived plasmids).

* Corresponding author. Mailing address: Laboratoire de Génétique et Microbiologie (INRA UA952), Faculté des Sciences, Université Henri Poincaré (Nancy 1), BP239, 54506 Vandoeuvre-lès-Nancy, France. Phone: (33) 3 83 91 21 93. Fax: (33) 3 83 91 25 00. E-mail: decaris@nancy.inra.fr.

TABLE 1. Bacterial strains, phages, and plasmids

Strains, phages, and plasmids	Relevant characteristics	Reference or source
E. coli		
SURE	Host for plasmids derived from pBC KS(+) and pBluescript SK(−); *dam*$^+$	Stratagene
HB101	Propagation strain for pBC KS(+) plasmid; *dam*$^+$	Stratagene
KW251	Propagation strain for λ recombinant bacteriophages; *dam*$^+$	Stratagene
VEC6831	Host for plasmids derived from pG+Host9; *dam*$^+$	Provided by E. Maguin
S. thermophilus		
A054	Strain closely related to CNRZ368 which does not possess ICE*St1*	Industrial
CNRZ368	Strain closely related to A054 possessing ICE*St1*	CNRZ[a]
NST1010	Derivative of CNRZ368 with *sth368IR* gene disrupted by pNST154 integration	This work
NST1013A	Derivative of A054 containing a copy of pNST153IS integrated into IS*1195*L	This work
L. lactis MG1363	Plasmid free; transformation host	14
Phages[b]		
λNST101	λ recombinant phage containing the left end of ICE*St1*	6
λNST106	λ recombinant phage containing an internal fragment of ICE*St1*	6
λNST107	λ recombinant phage containing an internal fragment of ICE*St1*	6
λNST108	λ recombinant phage containing an internal fragment of ICE*St1*	6
λNST113	λ recombinant phage containing the right end of ICE*St1*	6
ϕST84	Pac site lytic group II bacteriophage of *S. thermophilus*	5
Plasmids[b]		
pBC KS(+)	Cloning vector; Cam^r; 3.4 kb	Stratagene
pBluescript SK(−)	Cloning vector; Amp^r; 2.96 kb	Stratagene
pG+Host9	Thermosensitive shuttle vector; Ery^r; 3.8 kb	23
pNST132.2	2.2-kb *Eco*RI/*Eco*RV fragment of λNST101 (1132.2) cloned into pBluescript SK(−)	This work
pNST141	3.3-kb *Xba*I fragment of λNST107 (I141) cloned into pBC KS(+)	This work
pNST142	2.2-kb *Xba*I fragment of λNST107 (I142) cloned into pBC KS(+)	This work
pNST144	5.1-kb *Eco*RI fragment of λNST107 (I144) cloned into pBC KS(+)	This work
pNST144.1	674-bp *Xmn*I/*Cla*I fragment of pNST144 (I144.1) cloned into *Eco*RV/*Cla*I-digested pBC KS(+)	This work
pNST153	3.8-kb *Hind*III/*Sal*I fragment of λNST106 (HS38) cloned into pG+Host9	This work
pNST153IS	3.8-kb *Hind*III/*Sal*I fragment of λNST106 (HS38) and 0,9-kb *Hind*III fragment of pNST132.2 (H09) cloned into pG+Host9	This work
pNST154	*Eco*RI fragment of pNST144.1 cloned into pG+Host9	This work
pNST155	75-bp *Sau*3AI fragment of pBC KS(+) cloned into *Bam*HI-digested pBC KS(+)	This work
pNST156	1104-bp *Sau*3AI fragment of pBC KS(+) cloned into *Bam*HI-digested pBC KS(+)	This work

[a] CNRZ, Centre National de la Recherche Zootechnique.
[b] Locations of the chromosomal fragments on the map of ICE*St1* are indicated in Fig. 1. Amp^r, ampicillin resistance; Cam^r, chloramphenicol resistance; Ery^r, erythromycin resistance.

DNA extractions and cloning. pBC KS(+) DNA and pBC KS(+) derived plasmid DNAs were extracted from *E. coli* HB101 and *E. coli* Sure cells, respectively, by the alkaline lysis method (31). Plasmid DNA was extracted from *L. lactis* and *S. thermophilus* cells according to the method described by J. Frère (13). A Quantum Prep plasmid miniprep kit (Bio-Rad, Marnes-la-Coquette, France) was used to isolate plasmid DNA for sequencing from *E. coli*. λ bacteriophage DNA was isolated from *E. coli* KW251 lysates according to the method described by Sambrook et al. (31). *S. thermophilus* genomic DNA extractions were performed as previously described (8). The construction of the genomic library of *S. thermophilus* CNRZ368 in bacteriophage λGEM11 (Promega, Lyon, France) has been described previously (28).

Three restriction fragments of the insert of recombinant λ were cloned into the corresponding restriction sites of the plasmid pBC KS(+) to give pNST141, pNST142, and pNST144 (Table 1). The 674-bp *Xmn*I/*Cla*I fragment of pNST144 (the region encoding residues 190 to 406 of the endonuclease R.*Sth*368I) was cloned into *Eco*RV/*Cla*I-digested pBC KS(+), resulting in pNST144.1. The 3.8-kb thermosensitive *E. coli-Lactococcus-S. thermophilus* shuttle vector pG+Host9 (23) was used to construct pNST153, which contains the 3.8-kb *Hind*III/*Sal*I fragment of λNST106 (HS38) including *sth368IM* and *sth368IR*. pG+Host9 was also used to construct pNST154, which contains the *Eco*RI fragment of pNST144.1 encompassing a 674-bp fragment of *sth368IR*. Cloning of the *Sth*368I R-M system (pNST153) was based on selection of methylated plasmid DNA. In this way, *Hind*III/*Sal*I-digested λNST106 DNA was ligated to *Hind*III/*Sal*I-digested pG+Host9. The ligation mixture was used to transform *L. lactis* MG1363 by electroporation. Plasmid DNA was extracted, treated with *Sau*3AI endonuclease to cleave unmethylated DNA, which does not carry the *sth368IM* gene, and then used to transform *L. lactis* MG1363. The 2.1-kb *Eco*RI/*Eco*RV fragment of λNST101 was cloned into *Eco*RI/*Eco*RV-digested pBluescript SK(−) to give pNST132.2 containing the right end of IS*1195*L. The 920-bp *Hin*dIII fragment of pNST132.2 (the region containing 872 bp of IS*1195*L) was cloned into the *Hin*dIII restriction site of pNST153 to give pNST153IS.

Digested vector DNAs were dephosphorylated with alkaline phosphatase (Roche Diagnostics, Meylan, France) prior to ligation. Ligations were performed with T4 DNA ligase (Roche Diagnostics) according to the manufacturer's instructions.

Bacterial transformation. *E. coli* was transformed by electroporation according to the method of Dower et al. (9). *L. lactis* MG1363 was transformed by electroporation according to the method described by Holo and Nes (20). *S. thermophilus* A054 and CNRZ368 were transformed by electroporation by a method adapted from Marciset et al. (24) with the following modifications. Cells were grown at 42°C in HJGL medium (3% tryptone, 1% yeast extract, 0.5% KH_2PO_4, 0.5% beef extract, 1% glucose, 1% lactose) to an optical density at 600 nm (OD_{600}) of 0.3. Threefold-concentrated cells were electroporated in EPM medium (5 mM KH_2PO_4 [pH 6.1], 0.3 M raffinose, 0.5 M $MgCl_2$) and then resuspended in 1 ml of sucrose M17, 1.2-fold concentrated, and incubated for 4 h at 30°C. Electroporations were performed using a Bio-Rad Gene Pulser apparatus set at 25 μF, 200 Ω, and 2.5 kV.

Integration of pG+Host9-derived plasmid by homologous recombination. pNST153IS and pNST154 were used to transform *S. thermophilus* A054 and CNRZ368, respectively (Table 1). Plasmid DNA from several transformants was extracted and verified by agarose gel electrophoresis. Integration of pNST153IS and pNST154 into the chromosome by single crossover was performed according to the method described by Biswas et al. (4) with the following modifications: the cultures were shifted to 42°C for 3 h, and samples were diluted and plated at 42°C on M17 agar with 2 μg of erythromycin/ml. Total DNA of several integrants was extracted and submitted to Southern blot analyses to verify the location and copy

TABLE 2. Activity of restriction endonucleases recognizing sequences containing GATC on A054 and CNRZ368 DNAs

Restriction endonuclease	Recognition sequence[a]	Activity[b]	
		A054	CNRZ368
*Bam*HI	G ↓ G A T C C (A: o, C: +)	C	NC
*Bcl*I	T ↓ G A T C A (A: +, C: o)	C	C
*Bgl*II	A ↓ G A T C T (A: o, C: +)	C	NC
*Dpn*I	G A ↓ T C (A: m, C: +)	NC	NC
*Nde*II	↓ G A T C (A: +, C: o)	C	C
*Sau*3AI	↓ G A T C (A: o, C: +)	C	NC

[a] ↓, cleavage site of a restriction endonuclease; +, inhibition of the endonuclease by the methylated residue within the recognition sequence; o, digestion of the DNA with the endonuclease is not at all influenced by the methylated residue within the recognition sequence; m, the methylated residue is a prerequisite for the enzymatic activity of the endonuclease.

[b] C, DNA is cleaved by restriction endonuclease; NC, DNA is not cleaved by restriction endonuclease.

number of the integrated plasmids. NST1010 was obtained by integration of a unique copy of pNST154 into *sth368IR* of CNRZ368 (Table 1). NST1013A was obtained by integration of a unique copy of pNST153IS into IS*1195*L of A054 (Table 1).

Phage propagation and assays. The R-M phenotype of streptococcal hosts was monitored by plaque assays using the bacteriophage ϕST84 (lytic group II) (5). To determine the titer of the phage, 100 μl of the relevant phage dilution was added to 0.4 ml of Elliker medium containing 200 μl of an exponentially grown culture (OD_{650}, 0.4) of the appropriate host and 12.5 μM $CaCl_2$. The suspension was mixed and incubated for 10 min at 42°C to allow phage adsorption; 1.5 ml of prewarmed Elliker medium (0.5% agar) supplemented with 1.5% milk was then added, and the mixture was poured onto Elliker medium (1.6% agar) and incubated anaerobically for 20 h at 42°C. Phage DNA modification was established by purification of phage from single-plaque isolates and propagation on the same host culture. Phage lysates were obtained by infecting at a multiplicity of infection of 0.3 1 ml of prewarmed (42°C) Elliker medium containing 400 μl of an exponentially grown culture (OD_{650}, 0.4) of the appropriate host and 12.5 μM $CaCl_2$. After 15 min at 37°C, 9 ml of Elliker medium was added and the mixture was incubated at 42°C until complete lysis occurred. Cell debris was removed by centrifugation at 4,000 × *g* for 10 min at 4°C. The supernatant was treated with lysozyme (50 μg/ml), DNase I (5 μg/ml), and RNase A (10 μg/ml) for 30 min at 37°C, filtered through a 0.45-μm-pore-size cellulose nitrate filter, and stored at 4°C.

Isolation, partial purification, and test of endonuclease extracts. Partial purification of endonuclease extracts was performed according to the method described by Su et al. (37). The reaction mixture (20 μl) contained 0.2 μg of pBC KS(+) DNA or 0.5 μg of recombinant λ or genomic DNA, 10 μl of cell extract, and reaction buffer B (Roche Diagnostics). After incubation at 37°C for 3 h, the reactions were stopped by adding gel loading dye, and each mixture was applied to a 1.2% agarose gel for electrophoresis.

DNA sequencing and sequence analysis. Automatic DNA sequencing was performed on double-stranded template from a recombinant plasmid with an ABI Prism BigDye Terminator cycle-sequencing ready-reaction kit (PE Applied Biosystems, Paris, France) using a GeneAmp PCR system 2400 thermal cycler (Perkin-Elmer Cetus). Sequencing products were run on an ABI Prism 310 genetic analyzer. Related sequences were detected in the GenBank-EMBL database by using the BLASTX, BLASTP, and PSI-BLAST local alignment search tools (2). Searches of ORFs were performed with GeneMark (http://genemark.biology.gatech.edu/GeneMark/) using known codon preferences of *Lactococcus* spp., *Streptococcus pyogenes*, and *Streptococcus pneumoniae*. DNA Strider 1.2 was used to find direct or inverted repeats.

Nucleotide sequence accession number. The GenBank accession number for the nucleotide sequence reported in this paper is AJ271594.

RESULTS

Methylation of GATC sequence of CNRZ368. Several attempts at digestion of CNRZ368 DNA with some restriction enzymes recognizing the sequence 5′-GATC-3′, i.e., *Bam*HI and *Sau*3AI, were unsuccessful, indicating the presence of a methyltransferase. To determine if adenine or cytosine had been methylated, digestion assays were performed on DNAs of CNRZ368 and on the closely related A054 with various endonucleases recognizing sequences containing GATC (Table 2). *Bam*HI, *Bcl*I, *Bgl*II, *Nde*II, and *Sau*3AI cleave A054 DNA, showing that neither the A nor the C residues of the GATC sites of this strain are methylated. Furthermore, *Dpn*I, which cleaved only the $G^{6m}ATC$ site, does not cut A054 DNA. In the same way, CNRZ368 DNA is cleaved by *Bcl*I and *Nde*II endonuclease but not by *Dpn*I, indicating that 5′-GATC-3′ sequences do not contain N^6-methyladenine. However, *Bam*HI, *Bgl*II, and *Sau*3AI, which are inhibited by 5-methylcytosine in 5′-GATC-3′ sequences, do not cleave CNRZ368 DNA. These results showed that the C residue of 5′-GATC-3′ sequences is methylated in CNRZ368 but not in A054, a closely related strain. Therefore, this indicated that CNRZ368 carries a functional methyltransferase absent from A054.

Restriction at 5′-GATC-3′ sequence by crude cell extract of CNRZ368. The methyltransferase encoded by *S. thermophilus* CNRZ368 could be the methylation protein of a type II R-M system. A hypothetical restriction activity was searched for in this strain. In this way, crude cell extracts of A054 and CNRZ368 were used to perform digestion assays of A054 and CNRZ368 DNAs (data not shown). The crude cell extract of CNRZ368 cuts A054 DNA but not CNRZ368 DNA. This result indicates that CNRZ368 produces an endonuclease which is active on A054 DNA but not on CNRZ368 DNA. On the other hand, the crude cell extract of A054 does not cut either A054 DNA or CNRZ368 DNA, so no endonuclease activity was detected in this strain closely related to CNRZ368.

Furthermore, the pattern of undigested pBC KS(+) DNA was compared with those of pBC KS(+) DNA digested by crude cell extracts of A054 and CNRZ368 (Fig. 1). The results confirmed that the cell extract of A054 does not cut DNA. However, this cell extract has a retarding effect on the migration of DNA (Fig. 1, lanes 1 and 3). CNRZ368 cell extract partially cuts pBC KS(+) DNA (Fig. 1). The unpurified protein extract and the unoptimized conditions of the experiment are responsible for the partial digestion of the DNA. It could also be the result of competition of restriction with methylation of DNA, since corresponding methyltransferase is probably present in the crude extract.

The comparison of pBC KS(+) DNA digested by *Sau*3AI, which recognizes and cleaves the GATC sequence, and pBC KS(+) DNA digested by CNRZ368 cell extract revealed numerous small fragments common to both DNAs (Fig. 1). The patterns of pBC KS(+) DNA digested by CNRZ368 cell extract and those of the same DNA partially digested by *Sau*3AI are similar (data not shown).

Furthermore, the pBC KS(+) DNA used in these experiments was produced from the dam^+ bacterium *E. coli* HB101, which methylates the A residue of the GATC sequence. Since crude cell extract of CNRZ368 cuts $G^{6m}ATC$, the restriction endonuclease of this strain is not inhibited by methylation at this position.

Fragments generated from digesting pBC KS(+) with partially purified endonuclease extracts were cloned into the *Bam*HI site of pBC KS(+) by direct ligation, assuming that like R.*Sau*3AI, R.*Sth*368I generated sticky-end termini compat-

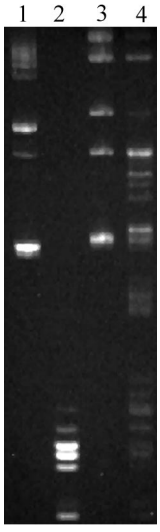

FIG. 1. Electrophoresis of DNAs digested by crude cell extracts. Comparison of patterns from digestion assays of pBC KS(+) DNA by A054 and CNRZ368 crude cell extracts. Lane 1, pBC KS(+) native DNA; lane 2, pBC KS(+) DNA digested by *Sau*3AI; lane 3, pBC KS(+) with A054 cell extract; lane 4, pBC KS(+) with CNRZ368 cell extract.

ible with *Bam*HI cohesive ends. Sequence analysis of the the ligation junctions of two plasmids (pNST155 and pNST156) isolated from transformants picked randomly showed that R.*Sth*368I recognizes and cleaves the same sequence as R.*Sau*3AI. Indeed, pNST155 contains a 75-bp insert corresponding to the sequence localized on the pBC KS(+) map at coordinates 1719 to 1794 and flanked by two *Sau*3AI sites. The cloned sequence allows the regeneration of two *Bam*HI sites on both sides. The 1,104-bp insert of pNST156 is localized at coordinates 1927 to 3031 on the pBC KS(+) map and corresponds to three adjacent *Sau*3AI fragments. In pNST156, this insert is bordered by two *Sau*3AI sites but does not regenerate *Bam*HI sites.

These results indicate that endonuclease produced by CNRZ368 has the same recognition and cleavage specificity as *Sau*3AI.

Identification of an R-M system in *S. thermophilus* CNRZ368. *S. thermophilus* CNRZ368 and A054 were tested for phage resistance against ϕST84 to detect activity of a putative R-M system. ϕST84 propagated on CNRZ368 was not restricted by A054 and CNRZ368 (Table 3). ϕST84 propagated on A054 was found to be restricted by CNRZ368 with an efficiency of plating (EOP) of 1.8×10^{-4}. Therefore, this temporary host-specific immunity of ϕST84 indicates that the strain CNRZ368 encodes a classical R-M system absent from A054. It was named *Sth*368I according to the standard R-M nomenclature (H. O. Smith and D. Nathans, Letter, *J. Mol. Biol.* **81:**419–423, 1973). Furthermore, an EOP of 0.62 was obtained when ϕST84 propagated on CNRZ368 (ϕST84.CNRZ368) was plated on CNRZ368. Moreover, bacteriophage plaques obtained on this strain are very small and appear hazy (data not shown). This could be due to the presence of a low-efficiency abortive infection mechanism or to physiological differences between the two strains A054 and CNRZ368.

Localization of *sth368IM*. Comparison of the physical maps revealed only two regions present in *S. thermophilus* CNRZ368 and absent from the closely related strain A054 (29). One of them is the integrative and potentially conjugative element ICE*St1* (6). Therefore, this 34.7-kb element could encode the *Sth*368I R-M system.

The inserts of five λ recombinant bacteriophages, isolated from a λGEM11 genomic library of *S. thermophilus* CNRZ368, entirely overlap the ICE*St1* element and the flanking regions (Fig. 2). The DNA of three of these recombinant bacteriophages, λNST101, λNST108, and λNST113, were found to be restricted by *Sau*3AI, whereas λNST106 and λNST107 DNAs were not (Fig. 3). These results show that λNST106 and λNST107 inserts carry the *sth368IM* gene encoding a methyltransferase which is expressed in *E. coli* KW251 and protects DNA against cleavage at the GATC site by *Sau*3AI. Moreover, pNST144, which contains a 5.1-kb *Eco*RI fragment common to the λNST106 and λNST107 inserts (Fig. 2), is digested by *Sau*3AI (Fig. 3). This suggested that the genes encoding the *Sth*368I R-M system are localized in the right region of the λNST106 insert.

Nucleotide sequence of the *Sth*368I R-M system. pNST141 and pNST142 were obtained by cloning the 3.3- (I141) and 2.2 (I142)-kb *Xba*I fragments of λNST107 into pBC KS(+) (Fig. 2). However, cloning of the 1.3-kb *Xba*I fragment localized between I141 and I142 failed. The inserts of pNST141 and pNST142 were entirely sequenced. The unclonable 1.3-kb *Xba*I fragment was sequenced by primer walking on λNST107 DNA. The nucleotide sequence revealed three ORFs (Fig. 2). BLAST searches on databases failed to find protein sequences related to the putative protein encoded by *orfS*. *orfS* is preceded by an AAAGGAAA ribosome binding site (RBS) and by a putative promoter sequence similar to those of *S. pneumoniae*, including a −10 sequence (TATAAT) and a −35 sequence (TCAATA) separated by a 17-bp consensus spacer (26). *orfS* is convergent with the next ORF, *sth368IR*. *sth368IR* encodes a putative 494-amino-acid protein with 23% identity with the endonuclease R.*Lla*KR2I of pKR223 of *L. lactis* KR2 (38) and 22% identity with the endonuclease R.*Sau*3AI of *S. aureus* (34). A putative RBS (ATGAGAGG) was found 7 bp upstream from the AUG start codon of this ORF (Fig. 4). *sth368IM* encodes a putative 421-amino acid protein with similarities to a large array of 5-methylcytosine methyltransferases including M.*Lla*KR2I (85% identity) and M.*Sau*3AI (52% identity). A suitable RBS (ACAGGAGA) was found 5 bp upstream of the AUG start codon of *sth368IM* (Fig. 4).

A putative promoter sequence is located upstream from the AUG start codon of *sth368IM*. This promoter includes a −10 sequence (TATAAT) and a −35 sequence (TTGATC) sepa-

TABLE 3. Effect of R-M system on plaque-forming ability of phage ϕST84

Phage	Relative EOP[a]			
	A054	CNRZ368	NST1010	NST1013A
ϕST84.CNRZ368[b]	1	0.62	0.45	1
ϕST84.CNRZ368.A054[c]	1	1.8×10^{-4}	0.35	7.1×10^{-4}
ϕST84.CNRZ368.A054.NST1013A[d]	1	0.51	0.42	1

[a] EOP of ϕST84 on the host of interest relative to plaquing ability on the nonrestricting host, *S. thermophilus* A054.
[b] ϕ ST84 propagated on *S. thermophilus CNRZ368.*
[c] ϕ ST84.CNRZ368 propagated on *S. thermophilus* A054.
[d] ϕ ST84.CNRZ368.A054 propagated on *S. thermophilus* NST1013A.

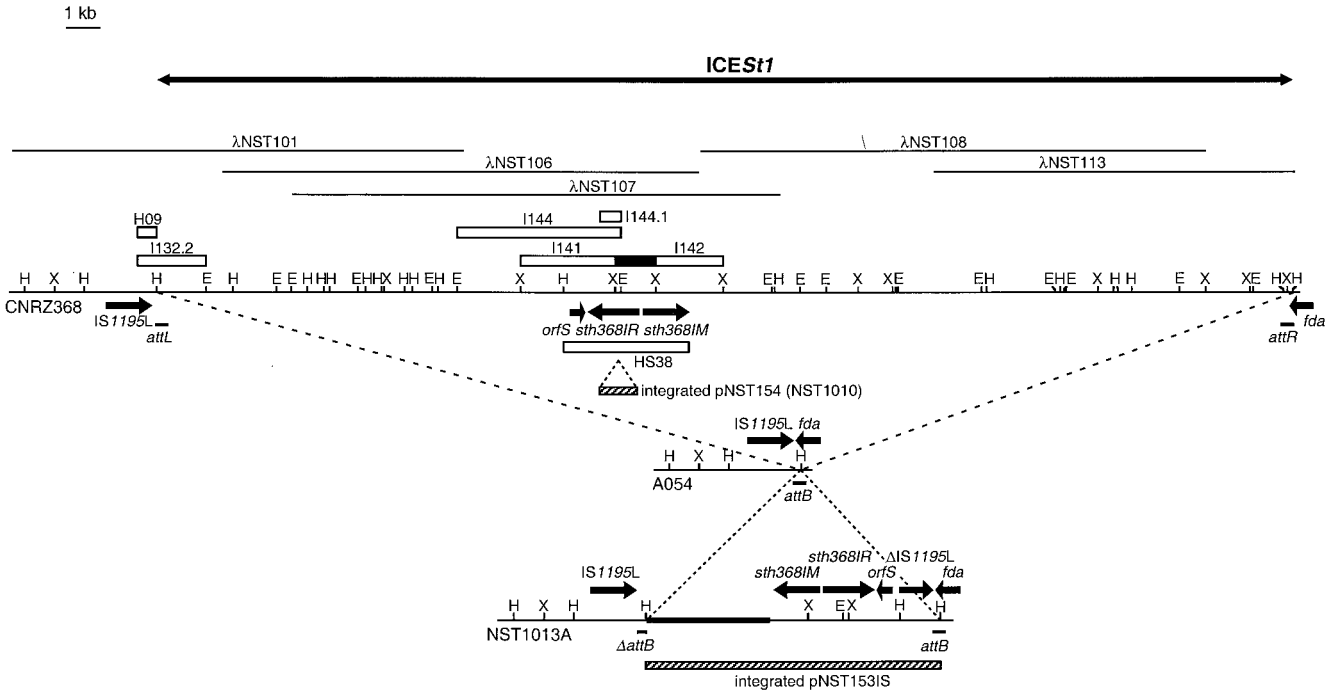

FIG. 2. Localization and maps of genes encoding the *Sth*368I R-M system and of cloned fragments on an ICE*St1* physical map. λ recombinant bacteriophage inserts are indicated by thin lines. The open boxes correspond to plasmid inserts. The unclonable 1.3-kb *Xba*I fragment is indicated by a solid box. The hatched boxes represent integrated plasmids. *attL* and *attR* show the left and right attachment sites, respectively, corresponding to the ends of ICE*St1*. *attB* corresponds to the chromosomal integration site of ICE*St1* found in A054. ORFs are marked by arrows indicating the direction of transcription. The pG+Host9 sequence is indicated by a thick line on the NST1013A restriction map. E, *Eco*RI; H, *Hind*III; X, *Xba*I.

rated by a 17-bp consensus spacer (Fig. 4). The −10 hexamer fits perfectly the consensus sequence of *S. pneumoniae*. An interesting point is the presence of a 5′-GATC-3′ sequence in the −35 hexamer. *sth368IR* is preceded by a canonical extended −10 promoter sequence (TNTGNTATAAT) (16) localized 33 bp upstream from the AUG start codon (Fig. 4). However, unlike *sth368IM*, no putative promoter sequence was found at the −35 region upstream from the −10 sequence. Such arrangements have also been found in numerous *S. pneumoniae* and *Bacillus subtilis* promoters (16, 30). This might constitute a transcriptional regulatory effect resulting in a more efficient expression of *sth368IM* than *sth368IR*. No suitable transcriptional-terminator structure was found between *orfS* and *sth368IR*. On the contrary, *sth368IM* is immediately followed by a perfect 10-bp inverted repeat and by a stretch of Ts which could be used as a rho-independent transcriptional terminator ($\Delta G_{37} = -10.3$ kcal · mol^{-1}) (12).

Disruption of *sth368IR* leads to sensible phenotype. The involvement of *sth368IR* in the phage resistance phenotype was verified by insertion mutagenesis. The thermosensitive plasmid pNST154 containing a 674-bp fragment of *sth368IR* was constructed and integrated by homologous recombination into *sth368IR* (Fig. 2). The resulting strain, NST1010, which contains two truncated copies of *sth368IR*, was used to perform phage infection assays with ϕST84 (Table 3). The EOPs of the methylated phage ϕST84.CNRZ368 and ϕST84 propagated on A054 (unmethylated phage ϕST84.CNRZ368.A054) are not significantly different. Insertional mutagenis of *sth368IR* leads to an R^- phenotype, indicating that this gene encodes the

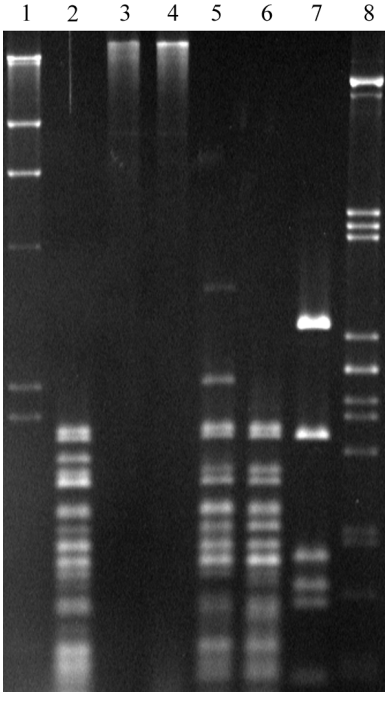

FIG. 3. Electrophoresis of *Sau*3AI digestion assays of DNA fragments overlapping ICE*St1*. Lane 1, λ DNA digested by *Pst*I; lane 2, λNST101; lane 3, λNST106; lane 4, λNST107; lane 5, λNST108; lane 6, λNST113; lane 7, pNST144; lane 8, λ DNA digested by *Hind*III.

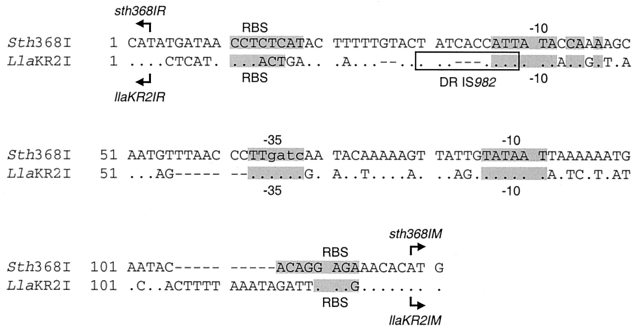

FIG. 4. Alignment of nucleotide sequences of *Sth*3681 and *Lla*KR2I R-M systems. *Sth*368I corresponds to the actual intergenic sequence of the *Sth*368I R-M system of *S. thermophilus* CNRZ368. *Lla*KR2I corresponds to the intergenic sequence of the *Lla*KR2I R-M system, where the entire IS*982* and a copy of the sequence duplicated after the IS*982* insertion (5′-TATCATT-3′) were deleted to allow comparison. The remaining copy of the duplication created by IS*982* transposition is indicated (DR IS*982*). The dots in the *Lla*KR2I sequence indicate the positions which are identical in the *Sth*368I and *Lla*KR2I intergenic sequences. The dashed lines indicates gaps. The ATG start codons of endonuclease- and methyltranferase-encoding genes are labeled by arrows indicating the direction of transcription. The shaded areas indicate putative RBSs or putative −10 and −35 consensus sequences. The sequence 5′-GATC-3′ found in the −35 motif preceding *sth368IM* and *llaKR2IM* is written in lowercase.

phage resistance phenotype of the strain CNRZ368. However, as for CNRZ368, the EOPs are not 1. Moreover, as for methylated ϕST84.CNRZ368 plated on CNRZ368, plaques of methylated ϕST84.CNRZ368 and unmethylated ϕST84.CNRZ368. A054 plated on NST1010 are small and appear hazy (data not shown). Furthermore, as previously shown, crude cell extracts of CNRZ368 cut pBC KS(+) DNA, whereas those of A054 and NST1010 failed to cleave this unmethylated DNA (data not shown).

Cloning of *Sth*368I R-M system in A054 strain. To confirm that *sth368IM* and *sth368IR* of ICE*St1* encode the R-M system involved in the phage resistance of CNRZ368, HS38 (Fig. 2) encompassing *sth368IM* and *sth368IR* was ligated to *Hin*dIII/*Sal*I-digested pBC KS(+). The ligation mixture was used to transform *E. coli* SURE and VEC6831 by electroporation. However, a recombinant plasmid was never obtained. Since attempts to subclone other R-M systems which modify the C residues of the sequence 5′-GATC-3′ were unsuccessful in various *E. coli* strains (34, 38), cloning in this species was given up. However, a similar experiment using the thermosensitive plasmid pG+Host9 as the vector and *L. lactis* MG1363 as the host and selection based on *Sau*3AI restriction was successful, resulting in plasmid pNST153 (21). The pNST153 DNA extracted from *L. lactis* MG1363 is not digested by *Sau*3AI, indicating that *sth368IM* is expressed at 30°C in this bacterium. A fragment containing the right end of IS*1195*L (sequence adjacent to ICE*St1*) was cloned into pNST153 to obtain the thermosensitive plasmid pNST153IS. Integration of this plasmid by homologous recombination in the A054 chromosome was selected. The integration site and the number of integrated copies were checked by hybridization of probe pG+Host9 to *Cla*I, *Eco*RI, *Pst*I, and *Sal*I patterns of several integrants (data not shown). The resulting strain, NST1013A, contains a single copy of the *Sth*368I R-M system integrated between IS*1195*L and the *fda* gene (Fig. 2). Phage assays were performed on this strain with ϕST84, and they gave results similar to those obtained with strain CNRZ368 (Table 3). However, the EOPs are about four times higher with NST1013A than with CNRZ368, and as on A054, phage plaques are larger, suggesting that CNRZ368 could encode another system of defense against bacteriophage infection, as has been previously suggested, or that the basis of regulation could be different.

DISCUSSION

The integrative and potentially conjugative element ICE*St1* of *S. thermophilus* CNRZ368 encodes the type II R-M system *Sth*368I, the first cloned in this species. *Sth*368I is related to *Lla*KR2I of *L. lactis* (38) and more distantly related to *Sau*3AI of *S. aureus* (34). Furthermore, *Sth*368I and *Sau*3AI recognize the sequence 5′-GATC-3′ and have identical specificities of methylation, i.e., 5-methylcytosine of GATC, whereas the specificity of *Lla*KR2I remains undetermined. The two genes encoding *Sau*3AI have identical orientations, whereas the genes encoding *Lla*KR2I and *Sth*368I are divergently transcribed. However, the *Lla*KR2I R-M system harbors an insertion sequence (IS) element (IS*982*) which is inserted between the −10 hexamer of a putative promoter sequence and the potential RBS sequence of *llaKR2IR*. The presence or absence of IS*982* does not seem to significantly alter *llaKR2IR* expression (38). The sequence similarities of *Sth*368I and *Lla*KR2I R-M systems and comparison of their structure strongly suggest that these two R-M systems could have evolved from a common ancestral R-M system. Since 7-bp target duplication flanks IS*982* (38), the structure of the *Lla*KR2I encoding sequence has probably evolved by transposition of IS*982* in the promoter region of *llaKR2IR*. Comparison of the intergenic sequences of the two R-M systems *Sth*368I and *Lla*KR2I, after the entire IS*982* element and one copy of the flanking direct repeats were removed, clearly showed that the putative promoters of the genes encoding the methyltransferases are related. They share the presence of a GATC site in the −35 putative promoter sequence of the methyltransferase gene. The transcription level of *sth368IM* could be modulated by the methylation state of this sequence 5′-GATC-3′, included in the −35 hexamer, as suggested by Twomey et al. for *llaKR2IM* (38). Thus, the expression of methyltransferase genes in both the *Sth*368I and *Lla*KR2I R-M systems would be identically regulated. Madsen and Josephsen also showed that the *Lla*DII R-M system has the recognition sites 5′-GCGC-3′ and 5′-GCCGC-3′, forming a putative stem-loop structure spanning part of the presumed −35 sequence and part of the intervening region between the −35 and −10 sequences preceding the methyltransferase-encoding gene (22).

In the same way, the promoter of *sth368IR* and of the ancestral *llaKR2IR* are related: both possess a perfect extended −10 promoter sequence (16) but no −35 sequence. Insertion of IS*982* disrupts this promoter in *Lla*KR2I, suggesting that the regulation of *llaKR2IR* and *sth368IR* are different.

Only a few *S. thermophilus* strains contain plasmids, and they are generally cryptic and small (25). To our knowledge, only one R-M system has been genetically characterized and sequenced in this species. Thus, the sequences of two ORFs of pSt0, a plasmid from *S. thermophilus* St0 which would encode

a type II R-M system with about 82% identity (nucleotide and protein sequences) with *Lla*DII of *Lactococcus lactis* subsp. *cremoris* (22) are available in databases (accession number AJ242480). However, the possible involvement of pSt0 in resistance against phage infection was not stated. Furthermore, pCI65st, a plasmid from *S. thermophilus* NDI-6 (27), was found to carry an ORF encoding a putative specificity subunit protein (hsdS) of a type I R-M system.

Numerous lactococcal conjugative plasmids are known to carry R-M systems (7, 10, 17, 19). *Sth*368I is the first R-M system to be carried by an integrative conjugative element and/or a similar element, like a conjugative transposon. The conjugative transposon Tn*5252* of *S. pneumoniae* encodes a type II methyltransferase (32) but no associated endonuclease. In vivo mutations in the gene encoding this methyltransferase were reported not to affect the transferability of the element (32). Sampath and Vijayakumar suggest that in this way Tn*5252* could be protected against a large array of recipient-encoded restriction endonucleases. On the other hand, we have shown here that *Sth*368I confers resistance against the ϕST84 bacteriophage. The presence of the *Sth*368I R-M system on ICE*St1* could favor the spread and maintenance of the element in the dairy industry, since phage attacks are frequent in this environment. The integrative system encoded by ICE*St1* would provide a stable site-specific integration of the element and, therefore, of the R-M system in the sensitive recipient strain.

ACKNOWLEDGMENTS

We are grateful to Harald Brüssow from Nestle Research Center (Vers-chez-les-Blanc, Lausanne, Switzerland) for providing the bacteriophage ϕST84. We thank E. Maguin for providing the thermosensitive plasmid pG+Host9 and *E. coli* strain VEC6831.

This work was supported by grants from the Institut National de la Recherche Agronomique, University of Nancy 1, and Ministère de l'Education Nationale, de la Recherche et de la Technologie, Paris, France.

REFERENCES

1. **Alatossava, T., and T. R. Klaenhammer.** 1991. Molecular characterization of three small isometric-headed bacteriophages which vary in their sensitivity to the lactococcal phage resistance plasmid pTR2030. Appl. Environ. Microbiol. **57:**1346–1353.
2. **Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
3. **Benbadis, L., J. R. Garel, and D. L. Hartley.** 1991. Purification, properties, and sequence specificity of *Ssl*I, a new type II restriction endonuclease from *Streptococcus salivarius* subsp. *thermophilus*. Appl. Environ. Microbiol. **57:**3677–3678.
4. **Biswas, I., P. Duwat, S. D. Ehrlich, A. Gruss, T. Hege, P. Langella, Y. Le Loir, and E. Maguin.** 1993. Efficient system for genetic modification of lactic bacteria: construction of food grade strains. Lait **73:**145–151.
5. **Brüssow, H., M. Frémont, A. Bruttin, J. Sidoti, A. Constable, and V. Fryder.** 1994. Detection and classification of *Streptococcus thermophilus* bacteriophages isolated from industrial milk fermentation. Appl. Environ. Microbiol. **60:**4537–4543.
6. **Burrus, V., Y. Roussel, B. Decaris, and G. Guédon.** 2000. Characterization of a novel integrative element, ICE*St1*, in the lactic acid bacterium *Streptococcus thermophilus*. Appl. Environ. Microbiol. **66:**1749–1753.
7. **Chopin, A., M. C. Chopin, A. Moillo-Batt, and P. Langella.** 1984. Two plasmid-***determined restriction and modification systems in*** *Streptococcus lactis*. Plasmid **11:**260–263.
8. **Colmin, C., M. Pebay, J. M. Simonet, and B. Decaris.** 1991. A species-specific DNA probe obtained from *Streptococcus salivarius* subsp. *thermophilus* detects strain restriction polymorphism. FEMS Microbiol. Lett. **81:**123–128.
9. **Dower, W. J., J. F. Miller, and C. W. Ragsdale.** 1988. High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. **16:**6127–6145.
10. **Forde, A., C. Daly, and G. F. Fitzgerald.** 1999. Identification of four phage resistance plasmids from *Lactococcus lactis* subsp. *cremoris* HO2. Appl. Environ. Microbiol. **65:**1540–1547.
11. **Forde, A., and G. F. Fitzgerald.** 1999. Bacteriophage defence systems in lactic acid bacteria. Antonie Leeuwenhoek **76:**89–113.
12. **Freier, S. M., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner.** 1986. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. **83:**9373–9377.
13. **Frère, J.** 1994. Simple method for extracting plasmid DNA from lactic acid bacteria. Lett. Appl. Microbiol. **18:**227–229.
14. **Gasson, M. J.** 1983. Plasmid complements of *Streptococcus lactis* NCDO712 and other lactic streptococci after protoplast curing. J. Bacteriol. **154:**1–9.
15. **Guimont, C., P. Henry, and G. Linden.** 1993. Restriction/modification in *Streptococcus thermophilus*: isolation and characterization of a type II restriction endonuclease *Sth*455I. Appl. Microbiol. Biotechnol. **39:**216–220.
16. **Helmann, J. D.** 1995. Compilation and analysis of *Bacillus subtilis* σ^A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. Nucleic Acids Res. **23:**2351–2360.
17. **Higgins, D. L., R. B. Sanozky-Dawes, and T. R. Klaenhammer.** 1988. Restriction and modification activities from *Streptococcus lactis* ME2 are encoded by a self-transmissible plasmid, pTN20, that forms cointegrates during mobilization of lactose-fermenting ability. J. Bacteriol. **170:**3435–3442.
18. **Hill, C.** 1993. Bacteriophage and bacteriophage resistance in lactic acid bacteria. FEMS Microbiol. Rev. **12:**87–108.
19. **Hill, C., K. Pierce, and T. R. Klaenhammer.** 1989. The conjugative plasmid pTR2030 encodes two bacteriophage defense mechanisms in lactococci, restriction modification (R+/M+) and abortive infection (Hsp+). Appl. Environ. Microbiol. **55:**2416–2419.
20. **Holo, H., and I. F. Nes.** 1995. Transformation of *Lactococcus* by electroporation. Methods Mol. Biol. **47:**195–199.
21. **Lunnen, K. D., J. M. Barsomian, R. R. Camp, C. O. Card, S. Z. Chen, R. Croft, M. C. Looney, M. M. Meda, L. S. Moran, D. O. Nwankwo, B. E. Slatko, E. M. Van Cott, and G. G. Wilson.** 1988. Cloning type-II restriction and modification genes. Gene **74:**25–32.
22. **Madsen, A., and J. Josephsen.** 1998. Cloning and characterization of the Lactococcal plasmid-encoded type II restriction/modification system, *Lla*DII. Appl. Environ. Microbiol. **64:**2424–2431.
23. **Maguin, E., H. Prévost, D. Ehrlich, and A. Gruss.** 1996. Efficient insertional mutagenesis in lactococci and other gram-positive bacteria. J. Bacteriol. **178:**931–935.
24. **Marciset, O., and B. Mollet.** 1994. Multifactorial experimental designs for optimizing transformation: electroporation of *Streptococcus thermophilus*. Biotechnol. Bioeng. **43:**490–496.
25. **Mercenier, A.** 1990. Molecular genetics of *Streptococcus thermophilus*. FEMS Microbiol. Rev. **87:**61–77.
26. **Morrison, D. A., and B. Jaurin.** 1990. *Streptococcus pneumoniae* possesses caononical *Escherichia coli* (sigma 70) promoters. Mol. Microbiol. **4:**1143–1152.
27. **O'Sullivan, T., D. van Sinderen, and G. Fitzgerald.** 1999. Structural and functional analysis of pCI65st, a 6.5 kb plasmid from *Streptococcus thermophilus* NDI-6. Microbiologie **145:**127–134.
28. **Pébay, M., Y. Roussel, J.-M. Simonet, and B. Decaris.** 1992. High-frequency deletion involving closely spaced rRNA gene sets in *Streptococcus thermophilus*. FEMS Microbiol. Lett. **98:**51–56.
29. **Roussel, Y., F. Bourgoin, G. Guédon, M. Pébay, and B. Decaris.** 1997. Analysis of the genetic polymorphism between three *Streptococcus thermophilus* strains by comparing their physical and genetic organization. Microbiology **143:**1335–1343.
30. **Sabelnikov, G. A., B. Greenberg, and S. A. Lacks.** 1995. An extended −10 promoter alone directs transcription of the *Dpn*II operon of *Streptococcus pneumoniae*. J. Mol. Biol. **250:**144–155.
31. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Sampath, J., and N. Vijayakumar.** 1998. Identification of a DNA cytosine methyltranferase gene in conjugative transposon Tn*5252*. Plasmid **39:**63–76.
33. **Sanders, M. E., P. J. Leonhard, W. D. Sing, and T. R. Klaenhammer.** 1986. Conjugal strategy for construction of fast acid-producing, bacteriophage-resistant lactic streptococci for use in dairy fermentations. Appl. Environ. Microbiol. **52:**1001–1007.
34. **Seeber, S., C. Kessler, and F. Götz.** 1990. Cloning, expression and characterization of the *Sau*3AI restriction and modification genes in *Staphylococcus carnosus* TM300. Gene **94:**37–43.
35. **Solaiman, D. K. Y., and G. A. Somkuti.** 1990. Isolation and characterization of a type II restriction endonuclease from *Streptococcus thermophilus*. FEMS Microbiol. Lett. **55:**261–265.
36. **Solaiman, D. K. Y., and G. A. Somkuti.** 1991. A type II restriction endonuclease of *Streptococcus thermophilus* ST117. FEMS Microbiol. Lett. **80:**75–80.
37. **Su, P., H. Im, H. Hsieh, S. Kang'a, and N. W. Dunn.** 1999. *Lla*FI, a type III restriction and modification system in *Lactococcus lactis*. Appl. Environ. Microbiol. **65:**686–693.
38. **Twomey, D. P., L. L. McKay, and D. J. O'Sullivan.** 1998. Molecular characterization of the *Lactococcus lactis Lla*KR2I restriction-modification system and effect of an IS*982* element positioned between the restriction and modification genes. J. Bacteriol. **180:**5844–5854.