Analysis of the genetic polymorphism between three *Streptococcus thermophilus* strains by comparing their physical and genetic organization

Yvonne Roussel,† Florence Bourgoin, Gérard Guédon, Mireille Pébay and Bernard Decaris

Author for correspondence: Bernard Decaris. Tel: +33 383 91 21 93. Fax: +33 383 91 25 00. e-mail: decaris@nancy.inra.fr

Laboratoire de Génétique et Microbiologie, UA INRA, Université Henri Poincaré, BP 239 54506 Vandoeuvre-lès-Nancy Cedex, France

The physical maps of *Streptococcus thermophilus* CNRZ368 and NST2280 strains were constructed by analysing PFGE patterns obtained with the low-frequency-cutting enzymes SmaI, BssHII and SfiI. Their chromosomes are 1864 and 1840 kb circular molecules, respectively. Comparison of their physical maps with that of the reference A054 strain revealed a relatively conserved organization of the restriction sites. Three variable regions were detected with the map of CNRZ368 whereas 15 were found with the map of NST2280. To construct the genetic maps, probes corresponding to 10 single-copy genes, the *rrn* genes and the insertion sequences IS1191, IS981 and IS51 were hybridized to Southern blots of chromosomal DNA digested with the different mapping enzymes. Comparison of the genetic maps of the three strains showed a conserved location of the mapped single-copy genes. However, six *rrn* loci were present in the chromosome of A054 and CNRZ368 whereas five were present in the NST2280 chromosome. A polymorphism was also found in the copy number of the insertion sequences between the three strains.

**Keywords**: *Streptococcus thermophilus*, physical and genetic map, variable region, IS element

**INTRODUCTION**

*Streptococcus thermophilus* is a Gram-positive bacterium used as a starter in the manufacture of fermented dairy products such as yoghurt and some hard cheeses. Despite its importance in the dairy industry, little is known about the genetic diversity within this species. The genetic variability within *S. thermophilus* was first investigated by analysis of restriction fragment length polymorphism. Analysis of patterns produced after digestion with high-frequency-cutting endonucleases and separation by conventional agarose gel electrophoresis revealed the existence of a great deal of genetic polymorphism among *S. thermophilus* strains (Colmin et al., 1987, 1991; Salzano et al., 1993). The intraspecific polymorphism of precise regions of the genome was also demonstrated by hybridization with an rDNA fragment (Pebay et al., 1992a) and with a DNA fragment specific to the two closely related species *S. thermophilus* and *Streptococcus salivarius* (Colmin et al., 1991). However, the extent of the genetic polymorphism in the *S. thermophilus* species and its genetic background have not yet been clarified.

Pulsed-field gel electrophoresis (PFGE), which enables separation of large DNA fragments, provides an alternative and complementary tool for studying genetic polymorphism among bacterial strains. Strains of the same species can be identified and characterized by comparing PFGE fingerprints produced by digestion of genomic DNA with rare-cutting enzymes and resolution of resultant fragments by PFGE. *SmaI* and *NotI* fingerprints were used to characterize strains of *S. thermophilus* (Salzano et al., 1993). In order to correlate the differences observed in fingerprints with major changes in the chromosome structure, the macrorestriction fragments have to be aligned in a physical map. A chromosomal map of *S. thermophilus* A054 was constructed (Roussel et al., 1994). The genome of the A054...
strain is a unique circular chromosome of 1824 kb. Twenty-three genetic markers, including 11 single-copy genes, the six rRNA operons and six tRNA genes, were located on the restriction map by performing Southern hybridizations with specific probes.

In the present work, the PFGE approach was used to analyse genomic diversity within the *S. thermophilus* species. This report describes the physical and genetic organization of the chromosomes of two additional strains of *S. thermophilus*, CNRZ368 and NST2280. Alignment of the chromosomal maps of these two strains with that of strain A054 identified and located variable physical regions. Ten probes specific to single genes, rDNA fragments and internal sequences of the genes, were hybridized with specific probes.

Comparison of the genetic maps of the three strains revealed an important similarity in the genome organization of the three strains. However, six *rrn* loci are present in the chromosome of A054 and CNRZ368 whereas five are present in the NST2280 chromosome. Moreover, the IS1191, IS981 and ISS1 elements were found in different copy numbers in the chromosome of the three strains.

**METHODS**

**Bacterial strains and gene probes.** *S. thermophilus* A054, which originated from the dairy industry, was provided by A. Mercenier (Institut Pasteur, Lille, France). The CNRZ368 strain was obtained from the culture collection of the Centre National de la Recherche Zootechnique (France) and the NST2280 strain descended from the industrial NST7 strain (Roussel et al., 1994). *Escherichia coli* KW251 (Promega) and Sure (Stratagene) were used as recipient of recombinant phages and plasmids, respectively. DNA fragments used as hybridization probes for genetic mapping studies are listed in Table 1.

**DNA preparations, endonuclease restriction and gel electrophoresis.** For high-molecular-mass DNA preparation, *S. thermophilus* cells were grown in modified M17 medium (Terzaghi & Sandine, 1975) supplemented with 20 mM DL-threonine and with chloramphenicol at a final concentration of 170 µg ml⁻¹ during the last hour of incubation. The cells were mixed with low-melting-temperature agarose and further treated to purify DNA as described previously (Roussel et al., 1994). For use in conventional electrophoresis, DNA was prepared according to a modified version of the Marmur procedure (Colmin et al., 1991). Digestions using restriction endonucleases *SmaI*, *BstHII* and *SfiI* purchased from Boehringer Mannheim were performed as previously described (Roussel et al., 1994). High-molecular-mass fragments were separated by contour-clamped homogeneous electric field (CHEF) gel electrophoresis (Chu et al., 1986) in a CHEF-DRII apparatus (Bio-Rad). Lambda DNA concatemers purchased from Bio-Rad and XhoI- or HindIII-digested lambda DNA were used as size markers. Fragments shorter than 30 kb were identified by conventional gel electrophoresis. Bacteriophage DNA was extracted according to Sambrook et al. (1989). Plasmid DNA was extracted by the alkaline lysis method described by Hopwood et al. (1985) and purified on a CsCl gradient.

**Table 1. DNA fragments used for genetic mapping and assignment to the physical map of the CNRZ368 and NST2280 chromosomes**

<table>
<thead>
<tr>
<th>Probes</th>
<th>Gene (function)</th>
<th>Reference/source</th>
<th>Hybridizing fragments</th>
</tr>
</thead>
<tbody>
<tr>
<td>pNST1</td>
<td>parA (adeno-succinate synthetase)</td>
<td>Pébay et al. (1993)</td>
<td>SmaI</td>
</tr>
<tr>
<td>pERF1</td>
<td>locS (lactose permease)</td>
<td>Poolman et al. (1990)</td>
<td>1</td>
</tr>
<tr>
<td>pGST1</td>
<td>galM (aldose 1-epimerase)</td>
<td>Galimin et al. (1991)</td>
<td>6</td>
</tr>
<tr>
<td>pSLH1</td>
<td>galE (UDPglucose 4-epimerase)</td>
<td>Delcour</td>
<td>9</td>
</tr>
<tr>
<td>pTIL4</td>
<td>pepC (aminopeptidase C)</td>
<td>Chaupet-Chartier et al. (1994)</td>
<td>3</td>
</tr>
<tr>
<td>pTGF21</td>
<td>rec-like (DNA recombination)</td>
<td>A. Mercenier et al.</td>
<td>4</td>
</tr>
<tr>
<td>pVE104</td>
<td>recA (DNA recombination)</td>
<td>Druwat et al. (1992)</td>
<td>13</td>
</tr>
<tr>
<td>pR224</td>
<td>hexA (DNA repairation)</td>
<td>Poolman et al. (1991)</td>
<td>13</td>
</tr>
<tr>
<td>pST1</td>
<td>23S gene encoding rRNA</td>
<td>Pébay et al. (1992a)</td>
<td>3, 6, 11, 12, 17, 18</td>
</tr>
<tr>
<td>pST2</td>
<td>16S gene encoding rRNA</td>
<td>Pébay et al. (1992b)</td>
<td>3, 4, 11, 13, 14, 17</td>
</tr>
<tr>
<td>pNST19</td>
<td>gor (glutathione reductase)</td>
<td>Pébay et al. (1995)</td>
<td>6</td>
</tr>
<tr>
<td>pSLS2</td>
<td>IS191 (insertion sequence)</td>
<td>Guedon et al. (1995)</td>
<td>4</td>
</tr>
<tr>
<td>pSLH1</td>
<td>IS981 (insertion sequence)</td>
<td>Guedon et al. (1995)</td>
<td>2</td>
</tr>
<tr>
<td>pSLH1</td>
<td>IS591, IS51 and unknown</td>
<td>Bourgeois et al. (1996)</td>
<td>ND</td>
</tr>
<tr>
<td>pSLH1</td>
<td>IS591, IS51 and unknown</td>
<td>Bourgeois et al. (1996)</td>
<td>ND</td>
</tr>
</tbody>
</table>

ND, Not determined.

* All probes used were from *S. thermophilus*, with the exception of the pR224 probe, which was from *Streptococcus pneumoniae*.
† Fragment numbers refer to those given in Table 2.
‡ Personal communication.
§ Hybridization data are given in Table 3 for the E, 149.2 and AL probes.
|| Probe that revealed several fragments of different intensities. For each of these probes, the most strongly revealed band is indicated.
Fragment nomenclature. Each restriction fragment produced by
digestion with a single endonuclease was denoted by the
initial letter(s) of the enzyme used to produce it (Sm for SmaI;
Sf for SfiI, and B for BssHII). Fragments produced by digestion
with two endonucleases were denoted by the initial letters of
the two enzymes (BSm for BssHII/SmaI; SfSm for SfiI/SmaI;
and BSf for BssHII/SfiI). The fragments from each digest were
numbered in order of decreasing size.

DNA Southern blot analysis. High-molecular-mass fragments
separated by CHEF electrophoresis were transferred onto
positively charged nylon membranes (Hybond N+; Amersham)
by alkaline Southern transfer (Reed & Mann, 1985), whereas fragments separated by conventional electrophoresis
were transferred onto nylon membranes (Hybond N; Amersham)
following the manufacturer's instructions. The probes were labelled by random priming systems with
[a-32P]dCTP (Megaprime DNA labelling kit; Amersham) or with
digoxigenin (Dig DNA Labelling and Detection kit; Boehringer Mannheim). Hybridizations were carried out as
previously described by Larbi et al. (1990) for 32P-labelled
probes or by following the manufacturer’s instructions for
Dig-labelled probes.

RESULTS

PFGE of A054, CNRZ368 and NST2280 DNA digested
with low-frequency-cutting restriction enzymes and
estimation of their genomic sizes.

The three restriction endonucleases SfiI, BssHII and
SmaI previously used in the physical mapping of the S.
thermophilus A054 chromosome (Roussel et al., 1994)
were also used to digest the genomic DNA of strains
CNRZ368 and NST2280. Fig. 1 displays the PFGE
patterns of A054, CNRZ368 and NST2280 genomic
DNA digested with SmaI. The similarity of the A054 and
CNRZ368 SmaI PFGE patterns (lanes 1 and 2; 12
comigrating fragments among 15) suggests that only a
slight genetic polymorphism exists between the two
strains and probably involves a few regions in the
chromosome. The comparison of A054 and NST2280
SmaI PFGE patterns (lanes 1 and 3; six comigrating
fragments among 15) revealed a more important re-
striction polymorphism and suggested that the polymor-
phism between the two strains involves numerous
regions of the chromosome. SmaI, SfiI and BssHII
cleaved the DNA of CNRZ368 into 24, five and eight
fragments and produced 21, four and eight fragments
from NST2280 DNA, respectively (Table 2). Mean
chromosome sizes have been estimated as 1864 kb for
CNRZ368 and 1840 kb for NST2280 by summing-up
fragment sizes obtained in single and double digestions.

Comparison of the physical maps of the CNRZ368
and NST2280 chromosomes with that of A054

Physical maps of the CNRZ368 and NST2280 chromo-
somes were established by using the strategy previously
described for the map construction of the A054 strain
(Roussel et al., 1994). The SfiI, SmaI and BssHII
restriction sites were located by analysing partial di-
gestion products with SfiI, by analysing double-digest
patterns and by identifying cross-hybridizing fragments
with DNA probes. Double digestions were performed
according to two different procedures. The first method
was simultaneous digestions and the second method was
digestions of PFGE-purified fragments with a second
restriction enzyme. Thirty-seven restriction sites were
located on the physical map of CNRZ368: 24 SmaI sites,
eight BssHII sites and five SfiI sites (Fig. 2). Alignment of
the CNRZ368 and A054 physical maps showed that the
37 restriction sites are conserved in the two strains. The
40 kb difference in the CNRZ368 and A054 genome
sizes was found distributed into three variable regions,
named var1A, var1B and var1C (Fig. 2). The poly-
orphic region var1A was detected within the
BssHII–SfiI fragment BS56 of CNRZ368 (common part
of B1 and Sf4) where 6 kb is missing in comparison to
the BS56 fragment of A054. The variable region var1B
was located in the BS5 fragment (common part of Sm2
and B2) of CNRZ368 which has an additional 11 kb.
The third variable region (var1C) was located in the
Sm4 fragment of CNRZ368 which contains an addi-
tional 35 kb sequence compared to the equivalent Sm5
fragment of A054.

Thirty-three restriction sites were located on the chromo-
somal map of NST2280: 21 SmaI sites, eight BssHII sites
and four SfiI sites (Fig. 2). Alignment of NST2280 and

![Fig. 1. PFGE gel (CHEF configuration) of SmaI digests of
genomic DNA of S. thermophilus strains. Lanes: 1, strain A054;
2, strain CNRZ368; 3, strain NST2280; 4, lambda DNA
concatemers. The sizes of lambda DNA concatemers are given in
kb on the right. The ramped pulse times were 5–15 s at 200 V for
27 h. The gel was 1.0% SeaKem LE agarose in 0.5x TBE (1x TBE
is 0.1 M Trisborate, 0.2 mM EDTA, pH 8) held constant at 14 °C.
With the conditions used, fragments with sizes smaller than
30 kb were not detectable in the gel but were observed by
conventional electrophoresis.]
Table 2. Size of the restriction fragments of S. thermophilus CNRZ368 and NST2280 DNA

The mean sizes for the CNRZ368 and NST2280 genomes were 1864 and 1840 kb, respectively. Fragments in bold type have identical sizes to those of A054 DNA (Roussel et al., 1994). The sizes indicated are the mean of at least three independent experiments.

<table>
<thead>
<tr>
<th>Band</th>
<th>CNRZ368</th>
<th>NST2280</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Smal</td>
<td>SfiI</td>
</tr>
<tr>
<td>1</td>
<td>324</td>
<td>828</td>
</tr>
<tr>
<td>2</td>
<td>223</td>
<td>622</td>
</tr>
<tr>
<td>3</td>
<td>221</td>
<td>261</td>
</tr>
<tr>
<td>4</td>
<td>193</td>
<td>158</td>
</tr>
<tr>
<td>5</td>
<td>182</td>
<td>181</td>
</tr>
<tr>
<td>6</td>
<td>146</td>
<td>68</td>
</tr>
<tr>
<td>7</td>
<td>107</td>
<td>64</td>
</tr>
<tr>
<td>8</td>
<td>85</td>
<td>48</td>
</tr>
<tr>
<td>9</td>
<td>72</td>
<td>68</td>
</tr>
<tr>
<td>10</td>
<td>61</td>
<td>64</td>
</tr>
<tr>
<td>11</td>
<td>58</td>
<td>58</td>
</tr>
<tr>
<td>12</td>
<td>46</td>
<td>58</td>
</tr>
<tr>
<td>13</td>
<td>46</td>
<td>58</td>
</tr>
<tr>
<td>14</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>15</td>
<td>36</td>
<td>46</td>
</tr>
<tr>
<td>16</td>
<td>9.2</td>
<td>36</td>
</tr>
<tr>
<td>17</td>
<td>4.85</td>
<td>17</td>
</tr>
<tr>
<td>18</td>
<td>3.4</td>
<td>10.1</td>
</tr>
<tr>
<td>19</td>
<td>0.91†</td>
<td>9.2</td>
</tr>
<tr>
<td>20</td>
<td>8.5</td>
<td>24</td>
</tr>
<tr>
<td>21</td>
<td>8.5</td>
<td>24</td>
</tr>
<tr>
<td>24</td>
<td>0.91†</td>
<td>5.80</td>
</tr>
<tr>
<td>25</td>
<td>3.34</td>
<td>5.80</td>
</tr>
<tr>
<td>26</td>
<td>0.91†</td>
<td>9.2</td>
</tr>
<tr>
<td>27</td>
<td>1861</td>
<td>1867</td>
</tr>
</tbody>
</table>

* Presence of five 0.91 kb fragments.
† Presence of six 0.91 kb fragments.

A054 physical maps allowed 15 variable loci to be detected. They were named var2A–var2O. Eleven of them correspond to the modification in the size of a fragment delimited by conserved restriction sites. These variable regions are var2A, var2B, var2C, var2D, var2G, var2I, var2J, var2L, var2M, var2N and var2O (Fig. 2). The small difference of 0.4 kb detected in var2O between the Sm14 fragments was determined by analysing conventional electrophoresis patterns of double digestions with SfiI/Smal. The 8.5 kb fragment of A054 (SfiSm20) is replaced by a fragment of 8.9 kb in the genome of strain NST2280. Other size differences estimated by comparing the fragment sizes determined by PFGE extend from 3 kb to 13 kb. Physical map comparison also allowed us to detect the appearance and disappearance of Smal, BssHII or SfiI restriction sites in the variable regions var2E, var2F, var2H, var2K and var2M of the NST2280 chromosome (Fig. 2). var2M is a variable region which shows a modification of its size and the disappearance of two Smal sites.

Comparison of the genetic maps of the CNRZ368 and NST2280 chromosomes with that of A054

The lacS, galE, galM, ldb, gor, pepC, recA, hexA, rec-like, purA and rrr genes were located on the physical maps of the CNRZ368 and NST2280 chromosomes by hybridizing gene-specific probes to Smal, BssHII and SfiI patterns. The probes used and the fragments to which they hybridized are listed in Table 1. Both 16S-specific probe 143.1 and 23S-specific probe 141 were used to locate the rrr loci. Each of these genes was found to have the same location in the chromosome of the three strains within the limits of precision of the genetic mapping experiments (Fig. 2). The only difference in the genetic organization was the presence of five rrr loci in the chromosome of NST2280 instead of the six in the A054 and CNRZ368 chromosomes. We demonstrated that the NST2280 strain contained five rrr loci whereas the NST7 strain from which it is derived was previously shown to contain six rrr loci (Pébay et al., 1992b). The rrrD–rrrE tandem in the NST7 strain had been replaced by the single rrr locus in the NST2280 strain. The deletion event removes the sequence containing the 3’ extremity of the rrrD locus, the interoperonic region and the 5’ extremity of the rrrE locus corresponding to the 6 kb variable locus var2M.

Location of IS981 copies

The number and location of IS981 copies in the chromosome of the three strains A054, CNRZ368 and NST2280 were determined by hybridizing the IS981-specific probe 149.2 to conventional and PFGE restric-
Fig. 2. Comparison of the chromosome maps of the S. thermophilus strains A054, CNRZ368 and NST2280. The circular genomes of the three strains are represented as linear maps to facilitate the comparison. Restriction sites for Smal, BssHII and Sfl are indicated at the upper, middle and lower part of each map, respectively. The zero coordinate corresponds to the conserved Sfl site located between the Sfl and the Sf3 fragments. The left to right direction of the chromosome maps corresponds to the clockwise direction of the A054 circular map (Roussel et al., 1994). Restriction fragments of CNRZ368 and NST2280 DNA are numbered according to their sizes, from the largest to the smallest (Table 2). The A054 restriction fragments were previously described in the same manner (Roussel et al., 1994). For the A054 and CNRZ368 physical maps, the fragments Sm10 and Sm15 could not be oriented, so the limits between the two fragments are indicated by dotted lines. Horizontal bars above each physical map indicate where the genes and IS copies are located. AlS981, A151191 and AlSS1 indicate truncated copies of IS981, IS1191 and IS51, respectively. Each horizontal arrowhead at the top of each map at the rrn loci corresponds to the two internal Sfl sites of an rrn locus. The orientation of the arrowheads indicates the transcription direction of the rrn loci. The loci of ΔNST302 (Bourgoin et al., 1996) and ΔNST101 (Guedon et al., 1995) are also indicated on the CNRZ368 map. The variable regions were deduced by comparing the restriction map of A054 to that of CNRZ368 (var1A–var1C) and of NST2280 (var2A–var2O). The variable regions are indicated by horizontal bars under the physical map when a restriction fragment has a different size (the size difference is indicated in kb in parentheses) or by a vertical arrow when restriction sites are additional or missing. The scale given at the top of the figure indicates the size in kb.
**Table 3.** Hybridization results of the E probe (specific to IS1191), the I49.2 probe (specific to IS981) and the AL probe (specific to \(\alpha IS1\)) with restriction fragments of S. thermophilus A054, CNRZ368 and NST2280 DNA

<table>
<thead>
<tr>
<th></th>
<th>A054</th>
<th>CNRZ368</th>
<th>NST2280</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Smal</td>
<td>BssHII</td>
<td>Sfl</td>
</tr>
<tr>
<td>IS1191</td>
<td>Sm1</td>
<td>B1</td>
<td>Sf1</td>
</tr>
<tr>
<td></td>
<td>Sm2</td>
<td>B2</td>
<td>Sf2</td>
</tr>
<tr>
<td></td>
<td>Sm3</td>
<td>B3</td>
<td>Sf3</td>
</tr>
<tr>
<td></td>
<td>Sm4</td>
<td>B6</td>
<td>SfSm5</td>
</tr>
<tr>
<td></td>
<td>Sm5</td>
<td>B7</td>
<td>B7</td>
</tr>
<tr>
<td>(\alpha IS1)</td>
<td>Sm3</td>
<td>NT</td>
<td>NT</td>
</tr>
</tbody>
</table>

ND, Not detected; NT, not tested.

* Fragments revealed with a higher intensity than the other fragments of the pattern.

hybridization patterns. CNRZ368 DNA digestions with five different restriction enzymes (Asp700, CfoI, Eco0109, EcoRI and EcoRV) revealed four hybridizing bands instead of three in the case of A054. Since these high-frequency-cutting restriction endonucleases do not cut within the probe, the chromosomes of A054 and CNRZ368 would contain three and four copies of IS981, respectively. For the five enzymes tested, the A054 and CNRZ368 restriction patterns showed three comigrating bands, strongly indicating that three IS981 copies have an identical location in the map of the two strains. For each enzyme, the additional band revealed in CNRZ368 DNA was shown to carry the truncated copy of IS981 previously cloned into \(\alpha\)NST101 (Guedon et al., 1995). Hybridization of the I49.2 probe to Smal, BssHII, Sfl, BssHII/Smal and Sfl/Smal product digestions revealed the presence of two and three fragments bearing IS981 copies in A054 and CNRZ368 DNA, respectively (Table 3). However, a higher intensity was obtained for the Sm3 and Bsm5 (common part of Sm3 and B2) fragments of A054 and for the Sm2 and Bsm5 (common part of Sm2 and B2) fragments of CNRZ368 in the Smal and BssHII/Smal digests, respectively. These data suggested the presence of at least two copies of IS981 in this region of both the A054 and CNRZ368 chromosomes. Two IS981 copies from the Bsm5 fragment of CNRZ368 were found within a 12 kb region which is included in the insert of \(\alpha\)NST302 (Bourgoin et al., 1996). The location of the IS981 copies on the chromosomal map of strains A054 and CNRZ368 is indicated in Fig. 2. The I49.2 probe did not hybridize with the DNA of strain NST2280, indicating the absence of IS981 in this strain.

**Location of IS1191 copies**

The E probe which is included in the IS1191 element cloned into \(\alpha\)NST101 (Guedon et al., 1995) was used to count and locate IS1191 copies. Hybridization of the E probe to CfoI-, DraI-, Eco0109- or HaeIII-digested DNA from A054 and CNRZ368 revealed seven and eight bands, respectively. Six bands were revealed when the E probe was hybridized to NST2280 DNA digested with CfoI, ClaI, Eco0109 or DraI. As these high-frequency-cutting enzymes do not cut within the probe, A054, CNRZ368 and NST2280 DNA would contain seven, eight and six copies of IS1191, respectively. For the four enzymes tested, both A054 and CNRZ368 hybridization patterns showed seven comigrating bands, strongly suggesting that the seven IS1191 copies have an identical location in the map of the two strains. The additional band only revealed in CNRZ368 DNA was shown to bear the IS1191 copy cloned into \(\alpha\)NST101. 173 and 174 are two adjacent EcoRI fragments which were subcloned from \(\alpha\)NST101 and sequenced (Guedon et al., 1995). 173 carries a complete IS1191 copy inserted in a truncated IS981 element and 174 carries a truncated copy of IS1191 which is not revealed by the E probe. Hybridization of 173 and 174 to genomic DNA digested with EcoRI and EcoRV showed that this 44 kb sequence of the CNRZ368 chromosome was completely absent in the A054 and NST2280 DNA.

In order to locate the different copies of IS1191 on the chromosomal map of the three strains, the E probe was hybridized to Smal, BssHII, Sfl, BssHII/Smal and Sfl/Smal restriction patterns. Hybridization on the BssHII/Smal pattern revealed six fragments containing IS1191 copies for each of the three strains (Table 3), indicating the location of at least six copies of IS1191. Moreover, the Sm4 fragment of CNRZ368 was revealed with greater intensity by the E probe than the five other fragments, suggesting that this Sm4 fragment could carry at least two IS1191 copies. The location of \(\alpha\)NST101 on the Sm4 fragment of CNRZ368 confirmed that this fragment contains the supplementary copy of \(\alpha\)NST101. The location of the IS1191 copies on the chromosomal maps of the three strains is indicated in Fig. 2. Although high-frequency-cutting enzyme data indicated the presence of seven and eight IS1191 copies, only six and seven copies were located on the A054 and CNRZ368 maps, respectively. This suggests that another fragment could bear two IS1191 copies. The six copies of IS1191 in the NST2280 chromosome had
similar locations in the A054 and CNRZ368 chromosomes, suggesting that they could be the same copies.

**Location of ISS1 copies**

Two different groups of iso-ISS1 elements (α and β) were found in both Lactococcus lactis and S. thermophilus (Bourgoin et al., 1996). The AL probe which is an AluI fragment of pNST49 was used to count and locate the different α ISS1 copies in the chromosome of the three strains. Hybridization of the AL probe to DraI- and EcoRI-digested DNA of A054, CNRZ368 and NST2280 revealed three, three and one bands, respectively. As these enzymes do not cut within the probe, A054, CNRZ368 and NST2280 would contain three, three and one copies of α ISS1, respectively. The sizes of the three bands revealed by the AL probe on the A054 and CNRZ368 patterns are the same, suggesting that the three α ISS1 copies have an identical location on the chromosome of the two strains. The recombinant clone jNST302 isolated from a gene library of CNRZ368 was previously shown to contain three copies of the ISS1 element; two of them belong to the α subgroup and the third is an αβ mosaic element called ISS1SC (Bourgoin et al., 1996). The three ISS1 copies are located in the 12 kb region which also contains two IS981 copies. No β ISS1 element was found in NST2280 and ISS1SC was the only β ISS1 element found in A054 and CNRZ368 DNA (Bourgoin et al., 1996). In order to locate the three copies of the α ISS1 element on the chromosomal map, the AL probe was hybridized to Smal patterns of the three strains. Only one fragment was revealed for each strain (Sm3, Sm2 and Sm3 for A054, CNRZ368 and NST2280, respectively). A more precise location of the three ISS1 copies on the physical maps of A054 and CNRZ368 was determined by using 149, a 6 kb EcoRI fragment of jNST302 (Guedon et al., 1995), as a probe. This probe revealed the BSm5 fragments, which are the common parts of Sm3 and B2 for A054 and of Sm2 and B2 for CNRZ368.

**DISCUSSION**

In this work, we have analysed the genetic polymorphism existing between three S. thermophilus strains by PFGE. The size of the chromosomes of strains CNRZ368 and NST2280 has been estimated at 1864 and 1840 kb, respectively. These sizes are very close to those previously found for the A054 strain (1824 kb; Roussel et al., 1994). However, comparison of chromosome sizes is not a good indicator of genetic polymorphism. In the case of the three studied strains, the comparison of fingerprints obtained with both high-frequency-cutting enzymes or low-frequency-cutting enzymes revealed that CNRZ368 and A054 are very closely related to each other, and both are less closely related to NST2280.

The two strains A054 and CNRZ368 have chromosomes with similar restriction maps. All the 37 mapped sites (Smal, SfiI and BssHII) were found to be conserved. The only differences detected were the sizes of three restriction fragments located in the variable regions var1A, var1B and var1C. These modifications can be explained by the insertion or deletion of sequences in the chromosome of CNRZ368 or A054. The alignment of the NST2280 physical map with that of A054 revealed 15 variable loci which were found to be distributed all along the chromosome. Among them, 11 variable regions correspond to restriction fragments with modified size and five variable regions have additional or missing mapped restriction sites compared with the map of A054. The appearance or disappearance of restriction sites in the NST2280 chromosome can be explained either by point mutation within the restriction site or by insertion or deletion of a sequence containing a mapped site. For the var2M region, the disappearance of two Smal sites is associated with the deletion of a 6 kb sequence containing these two Smal sites (see later).

Comparison of gene locations in strains A054, CNRZ368 and NST2280 showed that the genetic organization is similar within these strains. Because A054 and CNRZ368 have very similar physical maps, a conserved genetic organization was expected. The 10 single-copy genes and the six rRNA loci share identical positions on the chromosomes. The only differences detected between these two strains were two additional copies of IS1191 and one additional copy of IS981 present in the variable region var1C of CNRZ368. The location of the five other copies of IS1191, the three other copies of IS981 and the three copies of ISS1 were found to be the same. When the location of markers on the NST2280 chromosome was compared with that of the A054 chromosome, more differences were detected. The NST2280 chromosome was found to contain five rRNA loci instead of six, the rRNA tandem being replaced by a single rRNA locus. A homologous recombination between the two rRNA loci of the original tandem probably resulted in the deletion of one rRNA locus and the interoperonic region. The loss of one rRNA locus from the tandem rRNA locus has already been observed in the progeny of the CNRZ368 (Pebay et al., 1992b) and A054 (Roussel et al., 1994) strains. In the chromosome of the NST2280 strain, no copy of the element IS981 was detected and only a single copy of ISS1 was found instead of three in the two other strains.

In other bacteria, variability of genetic organization has also been analysed by comparing their physical and genetic maps. In some bacterial species, genetic maps are not conserved; for example, comparison of the genetic maps of the larger replicons of two Leptospira interrogans serovars showed evidence of large rearrangements (Zuerner et al., 1993). However, comparative studies of chromosomal maps in Mycoplasma mycoplasmoides (Pyle et al., 1990) and Mycoplasma hominis species (Ladefoged & Christiansen, 1992) revealed a conserved location of the mapped genes as well as between two strains each of Pseudomonas aeruginosa (Schmidt et al., 1996) and Thermus thermophilus (Tabata & Hoshino, 1996). Conservation of the gene order was also found when the genetic maps from 22 Lyme disease agents (Borrelia burgdorferi) were compared (Casjens et al., 1995). Comparison of the physical
and genetic maps of eight Clostridium perfringens strains also showed a similar genetic organization in seven of them. Three hypervariable regions were detected and found to be associated with a virulence gene or with an enterotoxin gene (Canard et al., 1992). A comparative analysis of five physical maps of different E. coli strains with that of E. coli MG1655 revealed 41 variable regions (Perkins et al., 1993). Whereas some of these modifications could be explained by rearrangements involving repeated sequences (rrn loci and IS elements), or by the movement of bacteriophages, the majority of variations had unknown causes.

Because of their ability to cause rearrangements such as insertions, inversions or deletions, IS elements were suspected to contribute to the genetic variability observed between the three S. thermophilus strains. The two polymorphic regions var1B and var1C are associated with IS elements in the CNRZ368 chromosome. Indeed, the BS3 fragment (common region of Sm2 and B2) carries the variable region var1B as well as two copies of IS981 elements, one copy of IS1191 and the three copies of ISS1 (Fig. 2). The 12 kb region of 2NST302 containing the two copies of IS981 and the three copies of ISS1 was shown to be present in the A054 chromosome, indicating that IS981 and ISS1 elements were not responsible for the variability observed in var1B. The quasi-identity of IS981 and ISS1 element sequences between S. thermophilus and Lactococcus lactis and their distribution in the two species suggest recent horizontal transfer of these elements from L. lactis to S. thermophilus (Bourgin et al., 1996; Guédon et al., 1995). The 12 kb chromosomal region which seems to be absent in almost all of the S. thermophilus strains tested could result from the integration of a lactococcal DNA fragment already containing the IS elements into the chromosome of the common ancestor of A054 and CNRZ368. In the Sm4 fragment of the CNRZ368 chromosome, two additional copies of IS1191 and one additional copy of IS981 are detected in the variable region var1C. We showed that these IS copies were present in var1C included in a 44 kb fragment of CNRZ368, which is completely absent in the A054 and NST2280 chromosomes. Whilst these three IS copies were shown to participate in the variability observed in var1C, they are not directly responsible for the observed modification. It is tempting to speculate that the 44 kb sequence is a part of the 35 kb which could be a temperate bacteriophage or a transposon inserted in the var1C region.

Our study is a first step in the analysis of the intraspecific variability within S. thermophilus species. The nature and the origin of the sequences present in the variable regions have to be further analysed to understand the speciation process in the species.

ACKNOWLEDGEMENTS

We are grateful to the scientists who kindly supplied cloned genes and probes: A. Mercenier for the rec-like probe, J. Delcour for the idb probe, A. Gruss for the recA probe, B. Poolman for the gal probe, J. P. Claverys for the hexA probe and M. P. Chartier for the pepC probe. Y.R. was supported by a fellowship from the Ministère de la Recherche et de la Technologie. We thank Lois Silk and Huw Parry for reading the manuscript.

REFERENCES


Polymorphism in *Streptococcus thermophilus*


Received 6 September 1996; revised 4 November 1996; accepted 8 November 1996.