Genetics of *Streptococcus thermophilus*: A Review

ANNICK MERCENIER and YVES LEMOINE
Transgene S. A.
11 Rue de Molsheim
F67082 Strasbourg, Cedex, France

ABSTRACT

The genetics and molecular biology of several species of lactic acid bacteria have progressed impressively during the last years with a major emphasis on the mesophilic cheese lactococci. More recent efforts were also addressed to streptococci and lactobacilli. Among them, *Streptococcus thermophilus* (*Streptococcus salivarius* ssp. *thermophilus*) has received particular attention because of its use in yogurt and cheese manufacture. Several gene transfer techniques – conjugation, conjugal mobilization, spheroplast transformation, transduction, and electrotumformation – have been developed for this species. This enabled examination of the use of existing or newly constructed cloning vectors and elaboration of a first generation of expression vectors. Studies on both homologous and heterologous gene expression are being pursued in *S. thermophilus*. Ongoing research covers the lactose utilization determinants and reporter genes such as the one encoding for α-amylase. Because of the potential problem of phage attack in dairy processes, collections of *S. thermophilus* bacteriophages have been analyzed and compared. In view of accumulating data collected for the thermophilic dairy streptococci, it may be concluded that their genetic modification is now feasible.

INTRODUCTION

During the past decade, impressive progress has been made in the genetics of lactic acid bacteria. Since the first report of the transformation of a *Lactococcus lactis* strain (39), efforts have been extended to the development of recombinant DNA technology for several lactic acid bacterial species (7, 15, 21, 40, 46). Genetic engineering has been used to investigate the molecular organization of primordial fermentation pathways, like lactose utilization (16, 17) or technologically important traits such as phage resistance (13, 33, 56), proteolytic activities (17, 37), or bacteriocin production (34). These technologies were also developed with the aim of constructing improved or modified dairy, meat, or silage starters. Most of the studies were first conducted with mesophilic lactococci for which we may consider that all basic genetic tools and techniques are now truly functional, at least for a limited number of strains.

Studies involving both classical genetics (conjugation, transduction, plasmid curing, and mutagenesis) and recombinant DNA technology (transformation, vector construction, and gene cloning) are now being applied to other lactic acid bacteria such as certain streptococci, lactobacilli, leuconostocs, or even pediococci (7, 21, 48). Among these, *Streptococcus thermophilus* represents an important dairy starter, which is used on a large scale, often in symbiosis with other species, for the manufacture of yogurt and certain cheese varieties (69). Despite its economical importance, this bacterium had been the subject of very few studies until recently. The interest in constructing improved starter or developing new dairy products stimulated investigations into the genetics of *S. thermophilus*.

This paper will update the current knowledge concerning *S. thermophilus* genetics and, particularly, recent advances in the genetic manipulation of this microorganism. It should be mentioned that most of the reported data were obtained for industrial yogurt starter strains and sometimes included the type strain ATCC-19258. However, very little information refers to strains used in cheese production.
Recent taxonomical studies (20, 42, 58, 60) have reclassified thermophilic dairy streptococcal strains into the *S. salivarius* species (*S. salivarius* ssp. *thermophilus*). However, for ease of presentation we will keep the original name throughout this review. Without going into the detailed description of the species, it is probably worthwhile to recall certain properties of *S. thermophilus*, which clearly differentiate it from the *L. lactis* genus. The optimal growth temperature of *S. thermophilus* is 42 to 44°C. Lactose is hydrolyzed via a β-galactosidase (65), which is similar to the *Escherichia coli* enzyme. Until recently, there was no evidence about the fate of this sugar during its uptake by the cell. Poolman et al. (53) have isolated a *S. thermophilus* gene (the lacS locus), which complemented an *E. coli* strain defective in the lactose permease and mediated nonphosphorylated sugar transport in this mutant. Amino acid sequence comparisons have shown that the product of lacS differs markedly from the *E. coli* lactose permease, since it is an hybrid protein that shares a region of similarity with the phosphoenolpyruvate (PEP)-dependent phosphotransferase systems (53). Nevertheless, lactose does not seem to be phosphorylated before hydrolysis in thermophilic streptococci. The DNA-DNA hybridization (20, 31) and numerical taxonomy (5) studies have shown that *S. thermophilus* possesses limited homology with the mesophilic lactococci (previously named group N streptococci). As a consequence, it lacks the group N surface antigen. A very striking difference was also observed in the plasmid contents: although essential metabolic traits are often carried on plasmids in the lactococci (40, 46), they are usually chromosomally encoded in *S. thermophilus*. Indeed, most of the strains belonging to the latter species appear to be plasmid-free (26, 49, 66, 67) except for a few isolates that contained a single plasmid of relatively small size (26, 67). In a limited number of cases, up to two replicons may coexist in a single cell (67). The *S. thermophilus* plasmids studied so far have not been linked to any particular phenotypic trait and thus remain cryptic for the time being [26, 67]; Mercenier et al., unpublished data].

An additional feature of *S. thermophilus* deals with its limited range of usable carbon substrates. Most strains ferment as little as three sugars: lactose, glucose, and sucrose; raffinose can sometimes be used as an alternative energy source.

Gene Transfer Methods

The capacity to change the genetic content of a new bacterial host relies primarily on the ability to introduce genes into the microorganisms of choice. Subsequent challenges deal with the expression and stabilization of incoming DNA constructs and, finally, with the elaboration of a fully food-grade recombinant microorganism. This section will briefly summarize published data concerning the introduction of foreign DNA into *S. thermophilus*, with emphasis on significant points, and in the next section, the first steps toward controlled gene expression in this host will be described.

Several methods have been used by geneticists to introduce DNA into bacterial cells. Their application to lactic acid bacteria has been reviewed recently (21, 48). Although most of the work has been focused on the cheese lactococci, both naturally occurring mechanisms and artificially induced gene transfer are now well-established also for *S. thermophilus*.

Protoplast or Spheroplast Transformation. Interestingly, although bacterial transformation was first reported in 1928 for *Streptococcus pneumoniae*, a naturally competent lactic acid bacterial species (25), success in the field of dairy starters was only met more than 50 yr later when Kondo and McKay (39) introduced plasmid DNA into protoplasts of *L. lactis* LM3302. This was, however, achieved using a completely different methodology than the natural competence, which has not been demonstrated to occur in lactococci. The crucial observation that bacterial protoplasts, obtained by digestion of the cell wall with lytic enzymes, are able to incorporate DNA when incubated in the presence of polyethylene glycol (PEG) (4), has led to considerable advances in the genetics of several gram-positive bacterial species. It thus seemed promising to follow a similar approach (PEG/protoplast techniques) in order to develop a transformation system for *S. thermophilus*.

Success in this kind of procedure relies on the ability to prepare protoplasts or permeabilized cells, subsequently able to regenerate an
intact cell wall, together with the establishment of optimal DNA uptake conditions. Both aspects, which require the adjustment of several parameters and are often species- or even strain-dependent, have been reviewed (48). Appropriate cell wall lysis conditions (choice of the lytic enzyme, strength and duration of the lytic treatment, composition of the lysis buffer), definition of an adequate regeneration medium (choice and combination of the main osmoprotector and ingredients), determination of the DNA entrance conditions (polymerization and concentration of PEG, duration and temperature of the PEG shock, addition of liposomes), choice of a selective marker, and cell wall recovery conditions all markedly influence the final outcome of a protoplast PEG-mediated transformation procedure. For *S. thermophilus*, regeneration of digested cell wall represented the limiting step, as the strains studied grow very poorly, if at all, in presence of high concentrations of osmotic protectors (49). Consequently, currently described regeneration media seemed useless for yogurt starters. The problem could be solved with the aid of tranfection studies where phage DNA is used instead of plasmids (48, 49). Using the DNA from virulent phages of *S. thermophilus*, optimization of DNA uptake conditions (polymerization and concentration of PEG, duration and temperature of the PEG shock, addition of liposomes), definition of an adequate regeneration medium, determination of the DNA entrance conditions, choice of a selective marker, and cell wall recovery conditions all markedly influence the final outcome of a protoplast PEG-mediated transformation procedure. For *S. thermophilus*, regeneration of digested cell wall represented the limiting step, as the strains studied grow very poorly, if at all, in presence of high concentrations of osmotic protectors (49). Consequently, currently described regeneration media seemed useless for yogurt starters. The problem could be solved with the aid of tranfection studies where phage DNA is used instead of plasmids (48, 49). Using the DNA from virulent phages of *S. thermophilus*, the DNA uptake conditions have been optimized for three yogurt starter strains (47, 49). They were shown to vary significantly from strain to strain. The addition of liposomes (lecithine vesicles) was either beneficial or essential for the transfection whose efficiency reached $10^5$ to $10^6$ pfu/μg DNA under optimized conditions for bacterial cells treated with mutanolysin. However, the regeneration of protoplasts prepared in this way remained dramatically low [around 0.02% after the PEG shock, as calculated following Kondo and McKay (39)]. Further transfection trials led to a milder lytic treatment in which mutanolysin was replaced by lysozyme, generating less fragile cells still highly permeable to DNA (i.e., spheroplasts). Concomitantly, a large variety of potential regeneration media were screened for their ability to support the appearance of infective centers immediately after the PEG shock (48). Taken together, the results of these assays led to a functional PEG-protoplast transformation system for at least three *S. thermophilus* strains. The best efficiency reached was $7.5 \times 10^4$ transformants (tr)/μg of DNA; however, there were significant variations depending on the host strain and plasmid used (49). As will be mentioned, the chloramphenicol acetyltransferase gene carried by the vector pCK17 and its derivatives appears as the most efficient selective marker in the spheroplast transformation of our working strains (49).

A limited number of attempts have been made to apply an "average" spheroplast transformation protocol to other *S. thermophilus* strains, including type strain ATCC19258. Their failure is in agreement with numerous reports on the limited range and the irreproducibility of such protocols (48, 72). Therefore, it appears that each laboratory would have to adapt published protocols to its own host strain and working conditions. These drawbacks have prompted us, like others, to design more flexible gene transfer techniques and to turn to in vivo mechanisms. Natural transformation, transposition, conjugation, and transduction were all reported to occur in lactic acid bacteria, in different lactococcal or streptococcal species (9, 11, 21, 40, 48), but usually not in *S. thermophilus*. However, conjugation (55) and transduction (50) have been demonstrated quite recently for this microorganism. Conjugation and Conjugative Mobilization. The broad host-range plasmids pAMβ1 (12, 41), pIP501 (19, 30), and its derivative pVA797 (19), could be conjugated from various *L. lactis* or *Streptococcus faecalis* strains into a variety of *S. thermophilus* recipients (55). The transfer efficiencies appeared somewhat lower than those generally reported for lactococci, varying from $10^{-4}$ to $10^{-6}$ per recipient cell. They were usually not higher in intraspecies crosses than interspecific ones. Notably, both erythromycin and chloramphenicol resistance can be used equally well as selective traits in conjugal transfer. The conjugative mobilization system described by Smith and Clewell for *S. faecalis* (64) also worked with *S. thermophilus* strains, including type strain ATCC19258 (55). This procedure, based on the formation and subsequent resolution of a cointegrate intermediate between the self-transmissible plasmid (pVA797) and the mobilized vector (pVA838), allowed the introduction of pVA838 or its handler derivative pTG222 (55) into various dairy starters (55). Cloned genes ($β$-galactosidase, for example) could efficiently be intro
duced by this means in certain yogurt strains without prior mutation of the recipient cell; counterselection of the donor strain had been achieved on the basis of physiological properties (lactose utilization capacity and growth temperature of the recipient). The vector pVA797 (14) could also be mobilized by the plasmid pVA797, but the cointegrate formed did not seem to resolve properly. Although the pVA797::pVA838 system represents a workable and useful tool, further improvements would be required to establish a more elegant delivery system.

As in the case of lactococci, the precise mechanism of conjugal transfer and the possible existence of a pheromone-mediated system similar to \textit{S. faecalis} (10, 18) have not yet been elucidated for \textit{S. thermophilus}. Nevertheless, the initial success with conjugation was of importance since it represented the first DNA transfer to this microorganism (24, 55). It also allowed assessment of the expression of the selective markers and origins of replication carried on well-described plasmids in \textit{S. thermophilus}. This facilitated study of the stability of vectors in this species and provided a source of properly modified DNA for subsequent transformation trials.

\textit{Transduction}. Whereas natural competence has not been described so far for \textit{S. thermophilus}, transduction has been exploited as an alternative means to move genes from one host strain to another. Although the first phage-mediated DNA transfer was carried out for \textit{L. lactis} in 1962 (59), similar experiments were conducted only recently with \textit{S. thermophilus} (50). Two closely related virulent phages transduced various plasmids among their respective host strains. The transfer efficiencies (10^{-2} to 10^{-6}, as expressed by the number of transductants carrying plasmid per input phage particle) were independent of the plasmid size or the selected marker (chloramphenicol or erythromycin resistance) but mainly reflected the restriction capacity of the recipient cell. A remarkably high transfer was obtained with pTG219, a vector based on a small cryptic plasmid of \textit{S. thermophilus} (50), leading to efficiencies up to 10^{-1} (i.e., 10^6 transductants/ml of transduction suspension). The molecular reason for this high frequency transfer (HFT) is presently unknown; DNA-DNA hybridization studies between the plasmid DNA and the phage genome have not allowed detection of homologous regions between them. As was reported for other systems, this could lead to efficient packaging of vector DNA into the transducing phage head through homologous recombination with the viral genome (1). Strikingly, a similar HFT phenomenon also occurred with other shuttle vectors based on a different cryptic plasmid of \textit{S. thermophilus} (O'Regan and Mercenier, unpublished observations).

The ease of this technique added to the fact that both strains and phages originated from yogurt factories makes the described transduction system a very attractive tool for the construction of modified \textit{S. thermophilus} starters.

\textit{Electrotransformation}. Both in vivo gene transfer systems described herein represent invaluable genetic tools, but they still require the transformation of one intermediary strain, preferably belonging to either the \textit{L. lactis} or the \textit{S. thermophilus} species. As outlined previously (38), PEG-mediated protoplast transformation presents a number of drawbacks and difficulties which could be circumvented by the use of whole cells. An attractive PEG-induced transformation of intact \textit{L. lactis} strains has been described by Sanders and Nicholson (57). However such a protocol could not be adapted to \textit{S. thermophilus} (Slos and Mercenier, unpublished observations).

The availability of a completely new technology, called electrotransformation (8, 54, 62) solved many of the problems mentioned. Like numerous bacterial species, cells of \textit{S. thermophilus} are permeable to plasmid or phage DNA when they are submitted to high electric field pulses of controlled duration and intensity [see (48, 68) Mercenier et al., unpublished results]]. The maximal efficiencies reported for this species are equal to 10^3 to 10^6 tr/μg DNA. They could probably be increased with technically improved electroporation equipment. Indeed, despite numerous variations of the established protocols (including partial cell wall digestion), the number of transformants per microgram of DNA increases with the voltage applied to the electroporation chamber and does not seem to level off with the machines tested so far (Mercenier et al., unpublished data). This is in contrast with the situation existing for lactobacilli or other streptococcal species (8). It may actually reflect the existence of a "competent" subpopulation, the proportion of which in the
whole cell population could be smaller in the case of *S. thermophilus*. Yet electrotransformation presents the advantages of being quick, easy and adaptable to numerous strains belonging to a given species (8, 71). For instance, a standardized protocol was applied in parallel to 11 *S. thermophilus* strains; of these, 9 strains, including *S. thermophilus* ATCC19258, were transformed at the first attempt with efficiencies sometimes reaching $10^3$ to $10^4$ tr/µg DNA. It is suspected that restriction represented the major barrier to DNA entrance into the three nontransformable isolates (Kehrer and Mercenier, unpublished data). Similar data have been reported by Somkuti and Steinberg (68).

In conclusion, gene transfer no longer represents a limiting step in the genetic manipulation of *S. thermophilus* strains. As for lactococci, efficient and complementary techniques have been established for this microorganism.

**Cloning Vectors for *Streptococcus thermophilus***

The ability to introduce plasmid DNA into a number of *S. thermophilus* strains has allowed a careful survey of the available streptococcal vectors and their comparison to newly developed hybrid replicons. As a general rule, the vectors used for the mesophilic lactococci, *Streptococcus sanguis*, or *S. faecalis* seem to function very well in *S. thermophilus*. Several constructions have been tested so far (Table 1) (49, 68). For transformation, pCK17 (23) or pMG4536 (22) and pGK12 (36) are the most useful replicons for our strains (i.e., industrial yogurt starters and type strain ATCC19258). This, at least in part, because the chloramphenicol resistance marker was by far more appropriate for the selection of transformants than the erythromycin resistance gene. Both selective pressures were, however, equally efficient in conjugation or transduction. The reason for this difference remains unexplained at present. Similar situations have been described also in lactococcal species [see (49)] or other streptococci [(70); Solioz, personal communication].

The plasmid pCK17 has recently been modified by cloning the polylinker and α-lacZ fragment of the *E. coli* vector pUC18 (73) into a nonessential region (O’Regan and Mercenier, unpublished results). The new construction, pTG262, retains a broad host-range, possesses multiple unique cloning sites, and allows for the rapid screening of recombinant clones in *E. coli* (white versus blue colonies), a bacterial host in which such replicons are equally able to replicate (23).

The segregational stability of vectors like pVA838 (44) or pTG222 (55) was strikingly high in *S. thermophilus*: no plasmid loss was measured after more than 100 generations under nonselective growth at 37 or 42°C. The plasmids pCK17 and pMG4536 were about as stable as pVA838. In contrast, when the antibiotic selection was dropped, hybrid constructions like pTG219, which are based on cryptic replicons of *S. thermophilus*, segregated quickly (Romero and Mercenier, unpublished data). In addition, they also exhibited a more narrow host range than those derived from small cryptic plasmids of lactococci (15, 23, 36). It appears that the host range of these vectors is restricted to certain streptococci in that they are unable to replicate in *Bacillus subtilis*, for example. When the introduction of bacilli or staphylococci cloning vehicles was tried, the only success obtained was with the staphylococcal plasmid pC194 (29) and its high copy number variants isolated by Espinosa (personal communication) in *S. pneumoniae*. Introduction of them by protoplast fusion, as was done in the case of *L. lactis* (2), was not attempted. Consequently, it cannot be ruled out that negative results were due to a particularly low copy number or poorly selectable markers (like kanamycin or tetracyclin resistance genes, for example) rather than a nonfunctional origin of replication. Incoming plasmids may also have been restricted efficiently by the studied host cells. In addition, cloning vectors should possess regulatory signals necessary to express a foreign gene. Hence, promoter sequences have been isolated from *S. thermophilus* chromosomal DNA by a shotgun approach. The first generation of expression vectors constructed in this way has been used to express model genes in this species.

Expression and cloning vectors should ultimately be composed exclusively of food-grade sequences. This would require the replacement of antibiotic resistance markers by acceptable selection traits (for example, carbohydrate catabolic genes) and the removal of any *E. coli* vector part. Therefore, the sequence of a 7-kilobase (kb) endogenous plasmid from *S. thermophilus* (Mercenier and O’Regan, unpublished results).
TABLE 1. Useful cloning vectors for *Streptococcus thermophilus*.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Antibiotic resistance marker</th>
<th>Size (kb)</th>
<th>Replication in <em>E. coli</em></th>
<th>Replication in <em>B. subtilis</em></th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pVA838</td>
<td>Em¹</td>
<td>9.0</td>
<td>+</td>
<td>M</td>
<td>(44)</td>
</tr>
<tr>
<td>pTG222</td>
<td>Em</td>
<td>7.0</td>
<td>+</td>
<td>-</td>
<td>(55)</td>
</tr>
<tr>
<td>pTG244³</td>
<td>Em, promoterless cat-86</td>
<td>9.3</td>
<td>+</td>
<td>-</td>
<td>unpublished</td>
</tr>
<tr>
<td>pTG219</td>
<td>Em</td>
<td>7.4</td>
<td>+</td>
<td>-</td>
<td>(49)</td>
</tr>
<tr>
<td>pSA3</td>
<td>Em</td>
<td>10.2</td>
<td>+</td>
<td>+</td>
<td>(14)</td>
</tr>
<tr>
<td>pVA736</td>
<td>Em</td>
<td>7.6</td>
<td>-</td>
<td>ND²</td>
<td>(43)</td>
</tr>
<tr>
<td>pAMB1</td>
<td>Em</td>
<td>26.5</td>
<td>-</td>
<td>+</td>
<td>(12, 41)</td>
</tr>
<tr>
<td>pIP501</td>
<td>Em, Cm²</td>
<td>30</td>
<td>-</td>
<td>ND</td>
<td>(19, 30)</td>
</tr>
<tr>
<td>pC9</td>
<td>Cm</td>
<td>≥4.0</td>
<td>-</td>
<td>ND</td>
<td>Gasson [see (49)]</td>
</tr>
<tr>
<td>pCK17</td>
<td>Cm, (Km)²</td>
<td>5.9</td>
<td>+</td>
<td>+</td>
<td>(23)</td>
</tr>
<tr>
<td>pMO4536</td>
<td>Cm, (Km)</td>
<td>5.4</td>
<td>+</td>
<td>+</td>
<td>(22)</td>
</tr>
<tr>
<td>pTG262⁴</td>
<td>Cm, (Km)</td>
<td>5.6</td>
<td>+</td>
<td>+</td>
<td>unpublished</td>
</tr>
<tr>
<td>pNZ12</td>
<td>Cm, (Km)</td>
<td>4.3</td>
<td>+</td>
<td>+</td>
<td>(15)</td>
</tr>
<tr>
<td>pGK12</td>
<td>Em, Cm</td>
<td>5.3</td>
<td>+</td>
<td>+</td>
<td>(36)</td>
</tr>
<tr>
<td>pC194</td>
<td>Cm</td>
<td>2.9</td>
<td>+</td>
<td>+</td>
<td>(29)</td>
</tr>
<tr>
<td>pC194cop</td>
<td>Cm</td>
<td>2.9</td>
<td>+</td>
<td>+</td>
<td>Espinosa, unpublished</td>
</tr>
</tbody>
</table>

¹The plasmids listed in Table 1 were reported to be introduced by transformation in *S. thermophilus* [see notably (49, 68)]. Antibiotic markers in brackets could not be used to select *S. thermophilus* transformants.

²Em = Erythromycin; Cm = chloramphenicol; Km = Kanamycin; ND = not determined.

³pTG244 is a promoter-probe vector; Slos and Mercenier, unpublished data.

⁴pTG262 carries the α-lacZ screening marker and polylinker of pUC18 (69); O’Regan and Mercenier, unpublished data [see text].

The plasmids listed in Table 1 were reported to be introduced by transformation in *S. thermophilus* [see notably (49, 68)]. Antibiotic markers in brackets could not be used to select *S. thermophilus* transformants.

Gene Cloning and Expression

Genetic modification of a chosen host requires foreign DNA to be introduced and expressed, or existing genes to be amplified, in the cell. Because no transcription and translation signals had been characterized for *S. thermophilus*, random chromosomal sequences have been selected on the basis of their function in adequate in vivo tests. Analysis of their nucleotide sequence has shown 60 to 80% similarity with that of the classical bacterial promoters (Slos and Mercenier, unpublished results). The isolated fragments promoted expression when cloned in front of a promoter less cat-86 gene, either in *S. thermophilus*, or in *L. lactis*, or in both species. The molecular reason for this difference in "host range" is still being examined. The measurement of the amount of specific mRNA synthesized in each case has permitted classification of these promoters according to their transcriptional strength. They were then incorporated in appropriate vectors in order to direct the expression of model genes.

As a first choice, the *xylE* gene encoding for catechol 2,3-oxygenase (75) was cloned downstream of the potential promoter sequences. Although the gene was expressed properly in *B. subtilis*, no enzymatic activity could be detected in transformed *S. thermophilus* or *L. lactis* cells (Slos and Mercenier, unpublished data). As confirmed by other laboratories (Gasson, unpublished results; Chassy, personal communication), this reporter gene does not seem to function in lactic acid bacteria.

An attempt was made to express the α-amylase gene of *B. licheniformis* in *S. thermophilus*. Two chromosomal promoter sequences were linked to the α-amylase structural gene preceded by the original *Bacillus* ribosome-binding site and signal sequence [ (72); Speck and Zukowski, unpublished data]. Recombinant clones of *S. thermophilus* carrying these con-
structions synthesized detectable levels of the enzyme, as assessed by the formation of a clear halo on Cibacron-starch plates (35). It remains unclear at this stage whether the recombinant α-amylase is truly secreted in this host rather than released by lysed cells (Slos et al., unpublished results).

An alternative way to collect information on expression signals consisted of the analysis of the molecular organization of well-expressed genes of *S. thermophilus*. Research has focused on the lactose utilization pathway of this microorganism, which plays an essential role in dairy fermentations (27, 28, 53). The first *S. thermophilus* gene to be cloned in *E. coli* was the one encoding for the β-galactosidase of the strain ATCC19258 (27). In similar experiments, the corresponding locus of the yogurt starter A054 was isolated (Robert and Mercenier, unpublished results).

Comparative restriction enzyme analysis provided evidence that the chromosomal inserts cloned in each laboratory were very similar, if not identical. Both structural genes were sequenced separately. The alignment of the determined nucleotide sequences confirmed a level of DNA similarity equal to 99.1% (Schroeder et al., unpublished data). Out of the 1026 amino acids constituting the β-galactosidase open reading frame, only 9 differed between both strains. Among the 21 corresponding base changes, 12 were in third position of the codon having no effect on the deduced protein sequence.

The *S. thermophilus* β-galactosidase exhibits 35% similarity with the corresponding *E. coli* and *Klebsiella pneumoniae* proteins, respectively. It also appears highly similar (48% at the protein level) to the enzyme of *Lactobacillus bulgaricus* (61), its partner in the yogurt symbioses. In the case of strain A054, the 3' and 5' regions flanking the β-galactosidase structural gene were sequenced in order to localize potential additional lactose genes. The nucleotide analysis of a 5.6-kb chromosomal fragment (Mercenier et al., unpublished data) showed that the β-gal gene [open reading frame (ORF) 1026] was immediately preceded by another reading frame (ORF 280), which was incomplete on the sequenced fragment. The ORF 280 clearly presented 35% amino acid similarity with the *Salmonella thyphimurium* enzyme III component of the glucose phosphotransferase system (51). The ORF 280 and ORF 1026 are separated by three nucleotides. These data are in complete agreement with the results published by Poolman et al. (53), who have cloned and sequenced about 5 kb of an overlapping chromosomal DNA fragment from another yogurt starter (strain A147). Molecular analysis and biochemical studies demonstrated that the sequence in front of the β-gal gene most probably encodes for the lactose transport system (*lacS*) of *S. thermophilus*. The unexpected structure of this lactose carrier and the fact that lactose is not phosphorylated during transport in this bacterium has led the authors to hypothesize that the enzyme III-like region plays a regulatory role in this protein (53). Poolman et al. (53) also mention that the *lacS* of *S. thermophilus* presents more than 60% similarity with the equivalent gene of *L. bulgaricus*. From the determined sequences, the lactose transport and β-galactosidase genes seem virtually identical for the two *S. thermophilus* yogurt starters studied. Since these starters coagulate milk much faster than the type strain ATCC19258 (4 to 5 h at 44°C versus more than 10 h, respectively), it would be interesting to study lactose transport in the latter as well.

From the data presented, it may be postulated that the lactose genes could be organized as an operon in *S. thermophilus*. Preliminary measurements of the length of the lactose messenger RNA agree with this hypothesis (Poolman and Mainzer, unpublished observations; Slos et al., unpublished observations).

**Bacteriophages of *Streptococcus thermophilus***

As outlined earlier in this paper, phages represent an excellent source of DNA and an effective tool for the first steps in the establishment of a new transformation system. Some also efficiently transduce plasmid DNA among their respective host strains. They could ultimately be used to construct very potent cloning vehicles, but to our knowledge, this line of research has not succeeded yet for lactic acid bacteria.

The strong interest, which is driving the analysis of phage collections, results from the economic burden phage represent when they contaminate starter cells causing fermentation failure. As a probable consequence of the working conditions specific to each manufacturing
procedure (duration and temperature), phage attack is undoubtedly a larger problem in the cheese process than it is in yogurt making. Nevertheless, premature arrest of acid production occurs regularly in the latter process as well. In many instances, these accidents were linked to the appearance of phages lysing the \textit{S. thermophilus} components of the yogurt starter mixture.

This situation has raised increasing interest in collecting and classifying representative phage samples and, in the case of the mesophilic lactococci, it has also stimulated elegant research on the molecular basis of phage defense mechanisms (13, 33, 56). The construction of more "phage-resistant" yogurt strains would represent undeniable progress. However, only the first steps toward this aim have been performed for \textit{S. thermophilus}. Ten phage isolates able to lyse 14 host strains were studied to establish their lytic spectra and the extent of their relatedness (Faelen and Mercenier, unpublished data). This study has further been scaled up by L. Benbadis et al. [unpublished data; (3)] to the screening of about 200 yogurt strains and 18 phage lysates collected from fermentation failures occurring at various times and in different yogurt factories. The main conclusions of these studies (L. Benbadis et al., unpublished data) could be summarized as follows: three major groups of \textit{S. thermophilus} strains may be established on the basis of their restriction-modification properties. Fifty-eight percent of the strains examined appeared insensitive to all phages tested; 30% were lysed only by a limited number of phages and were defined as "phage specific"; 12% were sensitive to a variety of phages. In this last group, crosspropagation was very frequent and the phage particles released from certain hosts exhibited an enlarged lytic spectrum.

Several phages were compared to each other by restriction analysis of their genomic DNA and DNA-DNA hybridization experiments; the morphology and protein profile of type phages has also been examined (L. Benbadis et al., unpublished data). It appears that all the samples included in our study behaved as virulent phages and possessed a double-stranded DNA genome of a size varying between 35 and 45 kb. They showed remarkably more relatedness than expected from their random sampling. They fell into two subgroups, one corresponding to the phage-specific host strains, the other attacking exclusively the phage permissive strains. From the DNA-DNA hybridization studies, the extent of homology seemed to be very high within a given subgroup (i.e., all DNA restriction fragments crosshybridized for all the phages belonging to that subgroup). The exact percentage of similarity could, however, not be determined by electronic microscopy, as no reproducible heteroduplex molecules were obtained for pairs of \textit{S. thermophilus} phages (Oudet et al., unpublished data). This latter observation has also been made by Neve et al. (personal communication) during the analysis of their own collection of \textit{S. thermophilus} phages. However, limited homology was detected between phages belonging to separate subgroups. It should be mentioned that none of the isolates examined seemed related to \textit{Ø870.ST27} obtained from A. Jarvis (Slos and Mercenier, unpublished observations). Classification of \textit{S. thermophilus} strains and phages has allowed us to explain the evolution of certain fermentation failures and will probably help to achieve the appropriate selection of strains and optimize strain rotation in yogurt fermentation.

Additional collections of \textit{S. thermophilus} phages have been analyzed by other research groups (32, 45, 52). The conclusions of these studies agree well with the results mentioned. Indeed, the \textit{S. thermophilus} phages were usually assembled into two to three subgroups. The genome of several isolates possess cohesive termini [(52); Benbadis et al., unpublished data]. All phages studied in detail were reported to be virulent. The lysogenic state of \textit{S. thermophilus} strains is presently very poorly documented (63) and still remains an open question for numerous isolates.

Very little work has been devoted so far to the systematic analysis of the natural phage resistance mechanisms that may occur in \textit{S. thermophilus}. Solely, the existence of several restriction-modification systems has been documented [(3, 47); Benbadis et al., unpublished data]. It is well-known that certain strains of \textit{S. thermophilus} produce polysaccharides (6), but this property has not yet been linked to their endogenous phage resistance or sensitivity.

\textbf{CONCLUSION}

The genetics of \textit{S. thermophilus} has advanced in recent years to a point where tools
and techniques now exist as for the better studied mesophilic cheese starters. Gene transfer methods, basic cloning vehicles, a first generation of expression vectors, and sequences of a few cloned genes are all available. This progress represents an essential, but primary, approach to a better knowledge of the thermophilic dairy strains. In the area of genetics itself, much remains to be done; more elegant tools should be elaborated, such as the modification of chromosomal determinants by transposon mutagenesis or site-directed integration. This would allow isolation of specific mutants [none have been described so far for S. thermophilus apart from spontaneous antibiotic-resistant isolates (55)], cloning of the targeted genes, and construction of stabilized or modified strains by integration in the appropriate chromosomal region or by allelic exchange. Mutants should also be sought by classical mutagenesis to provide adequate hosts for the return of cloned genes. Indeed, the two genetic determinants isolated so far from S. thermophilus have been cloned by complementation of appropriate E. coli mutants; their contribution to the original host's physiology is still poorly documented at this stage.

Recombinant DNA technology developed so far represents, however, an invaluable tool to investigate key metabolic functions or technologically important traits of S. thermophilus. Quite logically, investigation has started with the lactose utilization pathway, but it should be extended to sucrose breakdown, naturally occurring phage resistance mechanisms, texturing properties, protease or peptidase activities, and other essential traits. Improved or modified thermophilic dairy starters would include hosts with increased phage resistance, novel texturing properties, controlled proteolysis, or the ability to produce a completely new component. The first steps toward these aims have now been reached for a number of S. thermophilus yogurt starter strains.

ACKNOWLEDGMENTS

We are indebted to J.-P. Lecocq for his constant and enthusiastic support during Transgene's research program on S. thermophilus. We wish to thank P. Chambon, L. L. McKay, and B. M. Chassy for many helpful discussions in the course of our work. We especially want to acknowledge our colleagues D. A. Romero, P. Slos, M. O'Regan, C. Robert, and J.-F. Viret for their contribution to the reported data and M. E. Sanders for critical reading of the manuscript. The technical help of I. Castellino, N. Kehrer, G. Lenzen, and D. Villeval has been highly appreciated. We thank M. Kennel for expert typing of this manuscript. Some projects described in this review have been supported by BSN Gervais Danone (Centre de Recherche Carasso).

REFERENCES

20 Farrow, J. A. E., and M. D. Collins. 1984. DNA base composition, DNA-DNA homology and long-chain fatty acid studies on Streptococcus thermophilus and Streptococcus salivarius. J. Gen. Microbiol. 130:357.


