Natural genetic transformation: development of a new genetic tool for *Streptococcus thermophilus*

Naturlig genetisk transformasjon: utvikling av ett nytt genetisk verktøy for *Streptococcus thermophilus*

Philosophiae Doctor (PhD) Thesis

Trinelise Blomqvist

Department of Chemistry, Biotechnology and Food Science
Norwegian University of Life Sciences

Ås 2010

Thesis number 2010:16
ISSN 1503-1667
ISBN 978-82-575-0928-6

Til Erik og Haldor

ACKNOWLEDGEMENTS

The present work was carried out at the Department of Chemistry, Biotechnology and Food Science at the Norwegian University of Life Sciences. The study was supported by The Research Council of Norway.

First and foremost, I would like to express my gratitude to my supervisor Leiv Sigve Håvarstein for giving me the opportunity to work with natural genetic transformation in streptococci. Your undying enthusiasm and tireless work with this fascinating phenomenon has been a good inspiration to me during these years. Thank you for excellent scientific guidance.

I would also like to thank present and former members of the streptococci research group and especially those who have been working with *S. thermophilus*: Ingunn Alne Hoell, who started the work with *S. thermophilus* in our lab, thank you for much valuable scientific advice and for your wonderful sense of humour, I really appreciate your friendship. I want to thank Hilde Steinmoen for sharing many good laughs with me, and for interesting discussions regarding streptococci and life in general. I am very grateful for your support when life was tough. Johan Biørnstad, who was the perfect gentleman in the Netherlands and in Italy, thank you so much, I will never forget it! Maren Harby Backstrøm, who I had the pleasure of co-supervising during her Master's degree, you really brought a lot of sunshine into the *S. thermophilus* project and have become a very dear friend. Last, but not least, my roommate at the office Hua Wei: although you did not work with *S. thermophilus*, thank you so much for many wonderful stories from China, and for lots of delicious Chinese food.

Finally, to Erik and Haldor; Dere er det mest verdifulle i livet mitt.

Takk for at du alltid er der for meg, Erik.

Thank you!

TABLE OF CONTENTS

ABSTRACT

Lactic acid bacteria (LAB) have been used in fermentation of foods for hundreds or even thousands of years. During the last decade the genome sequences of a number of commercially important LAB species and strains have been published. These sequences provide new insights into the genetic and metabolic capacities of the species/strains concerned. To be able to fully exploit the wealth of information produced by the genomic revolution, efficient tools for genetic manipulation of bacterial genomes are required. Such tools have been lacking for *Streptococcus thermophilus*, an important bacterium for the dairy industry, which is used in the manufacture of yoghurt and Italian- and Swiss-type cheeses. As several species in the genus *Streptococcus* are competent for natural genetic transformation, we hypothesized that *S. thermophilus* possesses this property as well. Naturally transformable bacteria take up naked extracellular DNA and incorporate it into their genomes by homologous recombination. This mechanism provides an ideal tool for genetic engineering in bacteria. The major goal of the present work was therefore to determine whether *S. thermophilus* is naturally transformable. If so, a second goal was to investigate the usefulness of this mechanism as a tool for genetic engineering in *S. thermophilus*.

Early on in the study, we identified and sequenced the gene encoding the *S. thermophilus* homologue of ComX, an alternative sigma factor controlling competence development in *Streptococcus pneumonia*. As we did not succeed in finding growth conditions that provoked spontaneous competence development in *S. thermophilus*, overexpression of *comX* was chosen as an alternative strategy. For this purpose, we used a pheromone-inducible two-component signal transduction system that regulates bacteriocin production in *S. thermophilus*. We found that overexpression of *comX* induced the competent state in the LMG 18311 strain, demonstrating for the first time that at least one strain of the species *S. thermophilus* is naturally transformable. Further investigations showed that the transformation efficiency of our system was high enough to allow genetic manipulation of the *S. thermophilus* genome without the use of a selectable marker. Instead, transformants could be identified by colony-lift hybridization with a specific oligonucleotide probe. The advantage of this procedure is that the bacterial genome can be altered at preselected sites without introduction of foreign DNA. In sum, the genetic tools developed in this thesis has opened up

new research opportunities that will lead to a better understanding of the metabolism and physiology of *S. thermophilus,* and perhaps also to the development of novel starter strains with improved properties.

SAMMENDRAG

Melkesyrebakterier har blitt benyttet til fermentering av mat i hundrevis og sannsynligvis tusenvis av år. I løpet av de siste ti årene har genomsekvensen til et betydelig antall kommersielt viktige arter og stammer av melkesyrebakterier blitt publisert. Dette har gitt ny innsikt i disse bakterienes genetiske og metabolske egenskaper og potensialer. For at forskerne skal kunne utnytte den stadig økende mengden av sekvensinformasjon på en best mulig måte, trengs det effektivt genetisk verktøy. Slikt verktøy har manglet for *Streptococcus thermophilus*, en viktig bakterie for meieriindustrien der den benyttes i produksjonen av yoghurt samt noen harde oster av italiensk og sveitsisk type. Siden flere arter i slekten *Streptococcus* er kompetente for naturlig genetisk transformasjon antok vi at det er en mulighet for at også *S. thermophilus* har denne egenskapen. Naturlig transformerbare bakterier tar opp "nakent" ekstracellulært DNA og inkorporerer det i sitt eget genom ved hjelp av homolog rekombinasjon. Denne mekanismen representerer et ideelt verktøy for genetisk manipulasjon av bakterier. Hovedmålet med arbeidet som er presentert i denne avhandlingen var derfor å undersøke om *S. thermophilus* er naturlig transformerbar. Dersom dette skulle vise seg å være tilfelle, var neste mål å undersøke anvendeligheten av denne mekanismen som et verktøy for genetisk manipulasjon av *S. thermophilus*.

Tidlig i studiet identifiserte og sekvenserte vi et gen fra *S. thermophilus* som koder for en homolog til ComX, den alternative sigma faktoren som kontrollerer kompetanseutvikling hos *Streptococcus pneumoniae*. Siden vi ikke klarte å finne vekstforhold som utløste spontan kompetanseutvikling hos *S. thermophilus*, valgte vi som en alternativ strategi å overuttrykke *comX*. Et induserbart tokomponent signaloverføringssystem som regulerer bakteriosin-produksjon hos *S. thermophilus* ble benyttet til dette formålet. Overuttrykk av *comX* induserte kompetanse i LMG 18311 stammen, og for første gang ble det vist at minst en stamme av arten *S. thermophilus* er naturlig transformerbar. Videre undersøkelser viste at transformasjonseffektiviteten vi oppnådde med vårt system var høy nok til at det var mulig å innføre mutasjoner direkte på genomet til *S. thermophilus*, uten å være avhengig av antibiotikaresistensgener eller andre markører til seleksjon av ønskede mutanter. Transformanter kunne i stedet identifiseres ved hjelp av kolonihybridisering med en spesifikk oligonukleotidprobe. Den store fordelen med denne framgangsmåten er at det bakterielle

genomet kan endres på et hvilket som helst sted uten at det samtidig introduseres fremmed DNA.

Det genetiske verktøyet som er utviklet i denne avhandlingen har gitt forskere som studerer *S. thermophilus* nye muligheter som på sikt vil resultere i bedre forståelse av denne bakteriens metabolisme og fysiologi, og kanskje også til utviklingen av nye starterstammer med forbedrede egenskaper.

LIST OF PAPERS

List of papers included in the thesis:

Paper I:
Blomqvist, T., Steinmoen, H. and Håvarstein, L.S. (2006). Pheromone-controlled overexpression of recombinant proteins in *Streptococcus thermophilus*. Archives of Microbiology 186 (6): 465-473.

Paper II:
Blomqvist, T., Steinmoen, H. and Håvarstein, L.S. (2006). Natural genetic transformation: a novel tool for efficient genetic engineering of the dairy bacterium *Streptococcus thermophilus*. Applied and Environmental Microbiology 72 (10): 6751-6756.

Paper III:
Blomqvist, T., Steinmoen, H. and Håvarstein, L.S. (2010). A food-grade site-directed mutagenesis system for *Streptococcus thermophilus* LMG 18311. Letters in Applied Microbiology 50 (3): 314-319.

List of other related papers by the author, not included in the thesis:

Johnsborg, O., Blomqvist, T., Kilian, M. and Håvarstein, L.S. (2006). Biologically active peptides in streptococci. In: The Molecular Biology of Streptococci Eds. Hakenbeck, R. and Chhatwal, S., Horizon Scientific Press.

Johnsborg, O., Kristiansen, P.E., Blomqvist, T. and Håvarstein, L.S. (2006). A hydrophobic patch in the pneumococcal competence pheromone CSP is essential for specificity and biological activity. Journal of Bacteriology, 188 (5): 1744-1749.

Kilian, M., Poulsen, K., Blomqvist, T., Håvarstein, L.S., Bek-Thomsen, M., Tettelin, H. and Sørensen, U.B.S. (2008). Evolution of *Streptococcus pneumoniae* and its close commensal relatives. PLoS ONE 3(7): e2683.

INTRODUCTION

The genus *Streptococcus*

Morphology, metabolism and habitats

Streptococci are Gram-positive spherical or ovoid cells arranged in chains or pairs. They are non-motile, catalase negative and unable to sporulate. Streptococci have a complex nutritional requirement that includes amino acids, peptides, purines, pyrimidies, vitamins and salts. All species are facultatively anaerobic and generate energy through fermentation of carbohydrates. Glucose is degraded through the Embden-Meyerhof glycolytic pathway and lactic acid is the main product. In the absence of glucose, most streptococci can ferment various other sugars and sugar alcohols by means of enzymes synthesized in the presence of the available substrate (Kilian, 1998).

The genus *Streptococcus* comprises a very diverse group of bacteria. Many species are members of the commensal microflora of mucosal membranes, such as the oral cavity, upper respiratory tract, intestine and genital tracts of humans or other animals. Some species are important pathogens. *Streptococcus pneumoniae* is one of the major bacterial agents of childhood pneumonia, the dominant cause of death in children worldwide (Garcia-Suarez *et al.*, 2006; Mulholland, 2007). *S. pyogenes* causes infections such as necrotizing fasciitis, scarlet fever and rheumatic fever (Kilian, 1998), and *S. agalactiae* (group B streptococci) is a leading cause of neonatal pneumonia, septicaemia and meningitis (Herbert *et al.*, 2004). *S. thermophilus* is a member of this genus with notably different characteristics. It is extensively used by the dairy industry for yogurt and cheese manufacture and is regarded as the second most important industrial dairy starter microorganism after *Lactococcus lactis* (Hols *et al.*, 2005).

Taxonomy of streptococci

The genus *Streptococcus* belongs to the Streptococcaceae family in the Lactobacillales order within the Bacilli class of the phylum Firmicutes, containing Gram-positive bacteria which have a low DNA G + C content in their genomes. Traditionally, classification of streptococci

was focused on the pathogenic species, and was in the beginning largely based on haemolysis. Streptococci were grown on blood agar to differentiate strains that were β-haemolytic from those which were not (Facklam, 2002). Haemolysis is still currently used as a guide in classification of the bacterium to the species level (Kilian, 1998). Serological grouping of β-hemolytic streptococci was introduced by Rebecca Lancefield in 1933. Based on carbohydrate antigens in the cell wall, streptococci were divided into serogroups A-W (with the exception of I and J). This became a valuable tool for identifying pathogenic streptococci and is still used in clinical microbiology. Despite its usefulness in classifying pathogenic streptococci, Lancefield's method is insufficient for classifying the majority of streptococci, in which group-specific antigens may be absent or shared by several distinct taxa (Kilian, 1998). By comparative analyses of 16S rRNA, more than 50 species in the genus *Streptococcus* have been divided into the following phylogenetically-related clusters: the Pyogenic, Anginosus, Mitis, Salivarius, Bovis and Mutans groups, plus one group containing species with uncertain or unknown genetic relationship (Köhler, 2007 and references therein).

Streptococcus thermophilus

Streptococcus thermophilus was described for the first time in 1919 by Sigurd Orla-Jensen in the book "Lactic acid bacteria" (Sherman, 1937). Due to its long history as a thermophilic lactic starter microorganism in the dairy industry, it has been classified as "Generally Recognized As Safe" (GRAS) for food production and has achieved "Qualified Presumption of Safety" (QPS) status in Europe, as the only species in the genus *Streptococcus*. Rapid conversion of lactose into lactate is the main role of *S. thermophilus* in dairy fermentations, but the production of other compounds, such as acetaldehyde and exopolysaccharides, to improve flavour and texture, is also an important contribution to dairy products (Behare *et al.*, 2009; Pastink *et al.*, 2008; Smit *et al.*, 2005). *S. thermophilus* is always used together with *Lactobacillus delbrueckii* subsp. *bulgaricus* (*Lb. bulgaricus*) in yogurt production. In the manufacture of cheese it is used together with lactobacilli for so-called "hard" cheeses (e.g. Emmental or Grana), and alone or together with lactobacilli for production of Mozarella and Cheddar (Hols *et al.*, 2005). *S. thermophilus* is classified as a member of the Salivarius group of streptococci which also includes *Streptococcus salivarius* and *Streptococcus vestibularis*.

Genomics of S. thermophilus

The first whole genome sequences of *S. thermophilus* strains were completed and published in December 2004 (Bolotin *et al.*, 2004). At the time of writing, the genome sequences of three strains (LMG 18311, CNRZ 1066 and LMD-9) have been completed and are available on the webpage of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). All three strains are used as starter strains by the dairy industry. The size of these *S. thermophilus* genomes are ≈ 1.8 Mb, placing them among the smallest compared to those of other lactic acid bacteria (LAB), which range from 1.8-3.3 Mb (Pfeiler & Klaenhammer, 2007). The *S. thermophilus* genome contains approximately 1900 protein-coding sequences, of which 80 % are orthologues of other streptococcal genes. In recent years, comparative genome analyses have provided significant new insights into the molecular biology and physiology of this important dairy bacterium (Bolotin *et al.*, 2004; Hols *et al.*, 2005; Makarova *et al.*, 2006; Pfeiler & Klaenhammer, 2007). The *S. thermophilus* genome has a very large number of pseudogenes; approximately 10 % of its genes are non-functional (Bolotin *et al.*, 2004). Interestingly, the most degenerated genes are related to carbohydrate utilisation and uptake (Bolotin *et al.*, 2004). Most *S. thermophilus* strains are able to utilise only lactose, glucose and sucrose. A limited number of strains are also able to utilize galactose and fructose (van den Bogaard *et al.*, 2004). In LAB, the phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS) is the main sugar uptake system, catalyzing transport concomitantly with phosphorylation of the carbohydrates. Four out of seven PTSs are non-functional in *S. thermophilus* (Bolotin *et al.*, 2004), and sucrose and fructose are the only sugars that are taken up by PTS (Hols *et al.*, 2005). It is also remarkable that most of the virulence genes, which are used by other streptococci to escape the host immune system or achieve cell adhesion, have been lost in *S. thermophilus* (Hols *et al.*, 2005).

The Ami system is the only oligopeptide transport system characterized in *S. thermophilus* (Hols *et al.*, 2005). It consists of five elements: AmiA, -C, -D -E and -F. The substrate is recognized by the extracellular oligopeptide binding protein (AmiA), the permease is formed by two trans-membrane proteins (AmiC and AmiD), and two membrane-bound cytoplasmic ATPases (AmiE and AmiF) provide the energy for the peptide translocation (Garault *et al.*, 2002). In *S. thermophilus* the Ami system is essential for growth in milk as well as internalization of signal peptides involved in quorum sensing pathways (Fontaine *et al.*, 2010; Garault *et al.*, 2002).

Adaptation to a life in milk

For hundreds of years, *S. thermophilus* has been used in the fermentation of yogurt and cheeses, and its genome has slowly adapted to a life in milk. Lactose is the major carbon source in milk, while the source of nitrogen is mainly casein. The observed evolution of *S. thermophilus* is clearly a response to this environment. Unlike other streptococci, *S. thermophilus* prefers lactose as its carbon and energy source (van den Bogaard *et al.*, 2000). For this purpose *S. thermophilus* has acquired a dedicated lactose transporter, the lactose permease LacS. Although lactose is the preferred sugar, most *S. thermophilus* strains utilise only its glucose moiety. The galactose moiety of lactose is released to the medium by LacS, which functions as a lactose/galactose antiporter (Hols *et al.*, 2005), illustrating the evolution towards a life in milk and specialised lactose utilisation.

Most LABs are auxotrophic for several amino acids. The genes involved in biosynthesis of amino acids are highly conserved in *S. thermophilus*, and the amino acid auxotrophy is strain specific, but in general minimal compared to other LABs (Garault *et al.*, 2000; Hols *et al.*, 2005; Letort & Juillard, 2001; Pastink *et al.*, 2009). The amounts of free amino acids and peptides are limited in milk, and the optimal growth in milk is therefore dependent on hydrolysis of milk caseins into peptides and amino acids. The first step of casein degradation is performed by cell-wall proteases. The cell wall protease PrtS has been shown to be essential for *S. thermophilus* growth in milk (Delorme *et al.*, 2010; Fernandez-Espla *et al.*, 2000). Only a few *S. thermophilus* strains (e.g. LMD-9) contain the *prtS* gene. The PrtB protease from *L. bulgaricus* is necessary for growth of *S. thermophilus* strains lacking the *prtS* gene during co-cultivation of *S. thermophilus* and *Lb. bulgaricus* in milk (Courtin *et al.*, 2002). Recently, a study addressing the emergence of the *prtS* gene in *S. thermophilus* was performed on strains from the historical collection of the French Institut National de Recherches Agronomiques (INRA) (Delorme *et al.*, 2010). This collection contains strains collected from diverse products in different countries from 1956 to 2008. Only 21 out of 135 *S. thermophilus* strains from the INRA collection contained the *prtS* gene (Delorme *et al.*, 2010). Interestingly 80 % of the strains harbouring *prtS* were collected after 1999. This indicates that this key enzyme, which increases the fitness of *S. thermophilus* in milk, may have become more common as a result of selection for its occurrence in commercial starters (Delorme *et al.*, 2010).

Horizontal gene transfer in S. thermophilus

Even though *S. thermophilus* has evolved mainly through loss-of-function events (Bolotin *et al.*, 2004), horizontal gene transfer has also shaped its genome through acquisition of foreign genes. As mentioned above, the cell wall protease gene, *prtS*, confers an important metabolic trait for milk adaptation, and seems to have become more widespread among *S. thermophilus* strains during the last 10 years (Delorme *et al.*, 2010; Rasmussen *et al.*, 2008). Although *prtS* is part of a genomic island flanked by two tandem insertion sequence elements that may facilitate its integration into new genomes, no elements for transfer have been identified (Delorme *et al.*, 2010). The *S. thermophilus* PrtS displays 95 % identity with PrtS from *Streptococcus suis*, and appears to have been transferred laterally to *S. thermophilus* (Delorme *et al.*, 2010). Several small genomic islands including genes involved in bacteriocin production (*blp*) or synthesis of polysaccharides (*eps*) are also thought to be the result of horizontal gene transfer events (Hols *et al.*, 2005). A 17 kb region, found within a truncated *pepD* gene, consists of fragments with more that 90 % identity to DNA from *Lb. bulgaricus* and the *L. lactis* subspecies *lactis* and *cremoris*. The genes believed to originate from *Lb. bulgaricus* enable *S. thermophilus* to synthesize methionine, a rare amino acid in milk (Bolotin *et al.*, 2004; Hols *et al.*, 2005).

The genome of *S. thermophilus* also contains fragments derived from phages and plasmids (Figure 1). Clustered, regularly interspaced short palindromic repeats (CRISPR) loci consist of a leader sequence followed by short, highly conserved DNA repeats (23-47 bp) interspaced with variable spacers (21-72 bp). The spacers are not conserved and are usually found to be homologous to DNA from bacteriophages or plasmids. Adjacent to CRISPR is a heterogeneous family of *cas* (CRISPR associated) genes encoding proteins that often contain nuclease, helicase, polymerase or RNA- or DNA-binding domains (Horvath & Barrangou, 2010; Waters & Storz, 2009). CRISPR arrays, together with associated *cas* genes, provide acquired immunity against phages and plasmids (Horvath & Barrangou, 2010; Sorek *et al.*, 2008). Phage infection and subsequent phage proliferation will lead to cell death in the majority of sensitive bacteria. In a small fraction of the population, the foreign DNA (of phages or plasmids) invading the cell is recognized by a Cas complex that cleaves off a novel spacer unit and integrates it into the leader end of the CRISPR locus (Figure 1A). This results in acquired immunity against subsequent infections because the foreign DNA is targeted in a sequence-specific manner (Horvath & Barrangou, 2010). The CRISPR arrays are initially

transcribed as a long RNA (Figure 1B), which is subsequently processed to individual spacer units flanked by two partial repeats called crRNAs. These small RNA units are used by a Cas complex to recognize invading DNA to be inactivated by the complex (Horvath & Barrangou, 2010). Since incoming CRISPR spacers are integrated in the leader end of the locus, these fragments also provide a historical perspective of previous infections. CRISPR systems are assumed to be transferred horizontally by plasmids, megaplasmids or prophages (Horvath & Barrangou, 2010; Sorek *et al.*, 2008). The number of CRISPR loci is variable between strains. Four CRISPR/cas systems are present in S. *thermophilus* DGCC7710, and three of the systems are homologous to systems of *Staphylococcus epidermidis*, *Streptococcus mutans* and *Escherichia coli* (Horvath & Barrangou, 2010 and references therein).

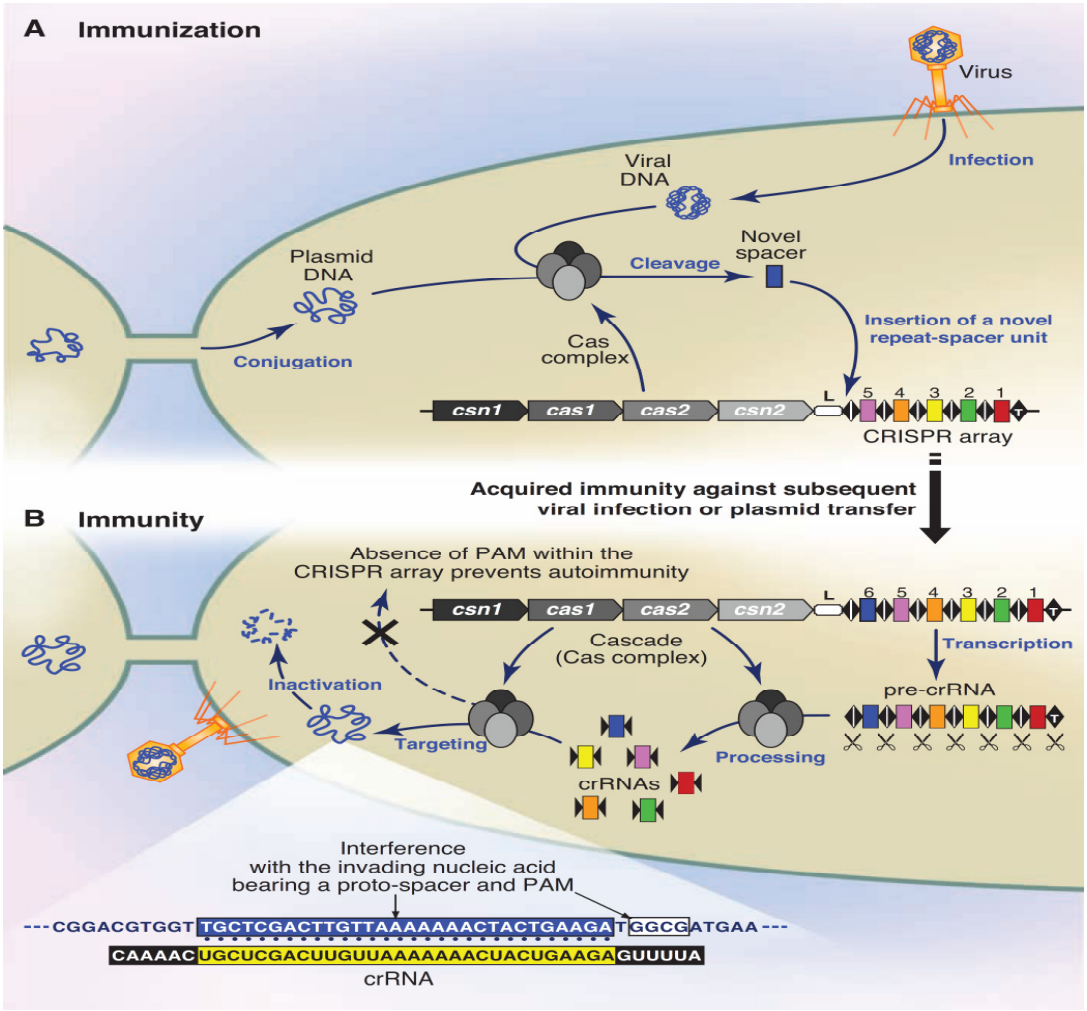

Figure 1. The CRISPR/Cas system. A) Immunization: Incoming exogenous DNA is recognized by a Cas complex that integrates a novel spacer unit at the leader (L) end of the CRISPR locus. B) Immunity: transcription of CRISPR arrays (pre-crRNA) are processed into mature crRNAs which are used by a Cas complex to recognize incoming DNA to be inactivated by the Cas complex. From (Horvath & Barrangou, 2010). Reprinted with permission from AAAS.

Quorum sensing and two-component signal transduction

Cell to cell communication in bacteria is accomplished by extracellular signalling molecules. If the signal is produced by the bacteria themselves, a response can be attuned to the population density. This process, termed quorum sensing, allows bacterial populations to coordinate gene expression. Two-component signal transduction systems in bacteria function by transferring a phosphoryl group from a receptor to a response regulator. The receptor in a classical two-component system (TCS) is usually a membrane bound histidine protein kinase (HK) designed to sense a particular signal and in response activate its cognate response regulator (RR), which in turn is responsible for the cellular response. The quorum sensing mechanism depends on a low-molecular-weight autoinducer that can be synthesized and secreted by the bacteria themselves. The autoinducer is sensed by a dedicated HK which responds by autophosphorylating a conserved histidine (His) residue. Subsequent transfer of the phosphate group to an aspartic acid (Asp) residue in the cognate RR will induce a conformation change that increases the affinity of the RR for its target promoters.

Sensing changes in the environment may be vital for the ability of bacteria to adapt to new conditions. A classical TCS is commonly used for this purpose to regulate features such as bioluminescence, virulence, sporulation and competence for natural transformation (Waters & Bassler, 2005). Both the HK and the RR have a modular architecture, and many different combinations of the modules have been reported, reflecting the widespread use of TCS in bacteria (Gao & Stock, 2009). In the following I will concentrate on TCS responding to unmodified peptide pheromones in LAB. These systems are known to regulate production of class II bacteriocins and competence for natural genetic transformation in streptococci (Diep *et al.*, 1995; Håvarstein *et al.*, 1995a). Genome analyses of the three completed genome sequences from *S. thermophilus* strains have identified eight complete two-component systems that are potentially involved in quorum sensing (Hols *et al.*, 2005). Some of these systems have an unknown function, while others display significant homology to systems from other species and might be responsible for production of bacteriocins or to mediate stress responses.

Histidine kinases

Histidine kinases (HKs) consist of two functional domains: a highly variable sensor domain and a conserved cytoplasmic kinase core domain. The amino acid sequence diversity observed in the N-terminal sensor domains reflects the diversity of environmental signals sensed by different TCS. Based on the membrane topologies of their sensor domains HKs are divided into three major groups (Mascher *et al.*, 2006). The largest group (classical HKs), sensing solutes and nutrients, typically consists of a relatively large extracellular sensing domain residing between two transmembrane helices. In the second group, the membrane-embedded sensor domain consists of 2-20 membrane-spanning helices. This is the smallest group but contains most of the TCSs involved in quorum sensing in Gram-positive bacteria. The third group has a cytoplasmic sensor domain and detects internal stimuli. The C-terminal conserved kinase core domain contains several amino acid motifs (H, N, D, F and G boxes), named after the most conserved amino acid residue in each box. The kinase core can be divided into two conserved subdomains; 1) the dimerization and histidine phosphotransfer domain, containing the H box with the phosphorylation site, and 2) the catalytic and ATP-binding domain containing the rest of the boxes (Mascher *et al.*, 2006). Together these subdomains define a highly conserved ATP binding cavity (Bilwes *et al.*, 2001; Marina *et al.*, 2001). There is some variability in the conserved H, N, D, F and G boxes, and based on these differences HPKs have been divided into 11 distinct subfamilies (Grebe & Stock, 1999).

As mentioned above, unmodified peptides are known to regulate production of class II bacteriocins and competence for natural genetic transformation (Diep *et al.*, 1995; Håvarstein *et al.*, 1995a). Most of the HKs responding to unmodified peptide pheromones belong to the histidine protein kinase subfamily 10 (HPK_{10}) (Grebe & Stock, 1999). HPK_{10} is membrane localized and has 5-8 transmembrane segments (Håvarstein, 2003 and references therein). The HPK_{10} family is characterized by having a kinase domain lacking a D box, an N box containing only one aspargine (D), no Proline (P) in the H box, and tyrosine (Y) as the second residue downstream of the conserved histidine in the H box (Grebe & Stock, 1999). ComD, the HK of the TCS regulating competence in *S. pneumoniae*, is classified as a HPK_{10} (Grebe & Stock, 1999). To the best of my knowledge, the histidine kinase BlpH (a paralogue of ComD), which regulates bacteriocin production in *S. pneumoniae*, has not been assigned any HPK subfamily. Comparison of the H, N, F and G boxes from the kinase domains of BlpH

and its homolog StbH (stu1687) from *S. thermophilus* LMG 18311 with the same boxes from ComD reveals that both BlpH and StbH most likely belong to the HPK_{10} subfamily (Figure 2). The conserved boxes in the kinase domains of BlpH and StbH have high homology to the kinase domain of ComD, and, except for one extra aspargine in the N box, all the above-mentioned criteria characterizing the HPK_{10} family are present.

```
Box:              H                                        N
StbH   LYNSVRSFRHDYSNLLVTL - 76 aa - EILDFIIVTSIFLDNAIEGA -
BlpH   LYKEIRSFRHDYTNLLTSL - 76 aa - SLLDFLTIVSILCDNAIEAS -
ComD   LYNEIRGFRHDYAGMLVSM - 76 aa - ELLDLVRIMSVLLNNAVEGS -

Box:                F            G
StbH   - 36 aa -IFERGVSNKGRERGIGLANVTEILDNYINVNLETQSNNFSFTQQLTI
BlpH   - 36 aa -IFSFGASSKGEERGVGLYTVMKIVESHPNTNLNTTCQNQVFRQVLTV
ComD   - 36 aa -LFALGFSTKGRNRGVGLNNVKELLDKYNNIILETEMEGSTFRQIIRF
```

Figure 2. Comparison of H, N, F and G boxes from the kinase domain of StbH from *S. thermophilus* LMG 18311 with BlpH and ComD from *S. pneumoniae* Rx. ComD belongs to the HPK_{10} family. Red letters = conserved residues in HPK_{10} family, blue letters = conserved hydrophobic residues in HPK_{10} family. Aspartic acid (D) is underlined in the N box. aa = amino acids.

HPKs function as dimers. For HPK_{10} type of receptors it is presumed that binding of their ligands, the unmodified peptide pheromones, brings about a conformational change in the HPK transmembrane domain that leads to autophosphorylation of the conserved His residue in the histidine kinase domain (Johnsborg & Håvarstein, 2009). Upon phosphorylation the configuration of the kinase will allow the response regulator access to the phosphorylation site (Gao & Stock, 2009). Transfer of the phosphoryl group to the response regulator will induce a conformational change in its DNA binding domain that increases its efficiency as a transcriptional activator of target genes (Gao *et al.*, 2007; Robinson *et al.*, 2000).

Response regulators

Response regulators (RRs) usually consist of two domains; the N-terminal conserved receiver (REC) domain accepting the phosphoryl group from the HPK, and the C-terminal variable effector domain. Phosphorylation of the RR occurs at a conserved Asp residue in the REC domain resulting in a conformational change that elicits a change in the structure of the DNA binding site of the effector domain(Gao *et al.*, 2007). RRs are classified according to their type of effector domains. About two thirds of all RRs contain a DNA binding output domain and function as transcription factors (Galperin, 2006). The remaining third consist of only the REC domain or have RNA binding, protein binding, enzymatic or uncharacterized output domains (Galperin, 2006; Gao *et al.*, 2007; Gao & Stock, 2009). Based on their function or structure the RRs are further divided into subfamilies (Figure 3).

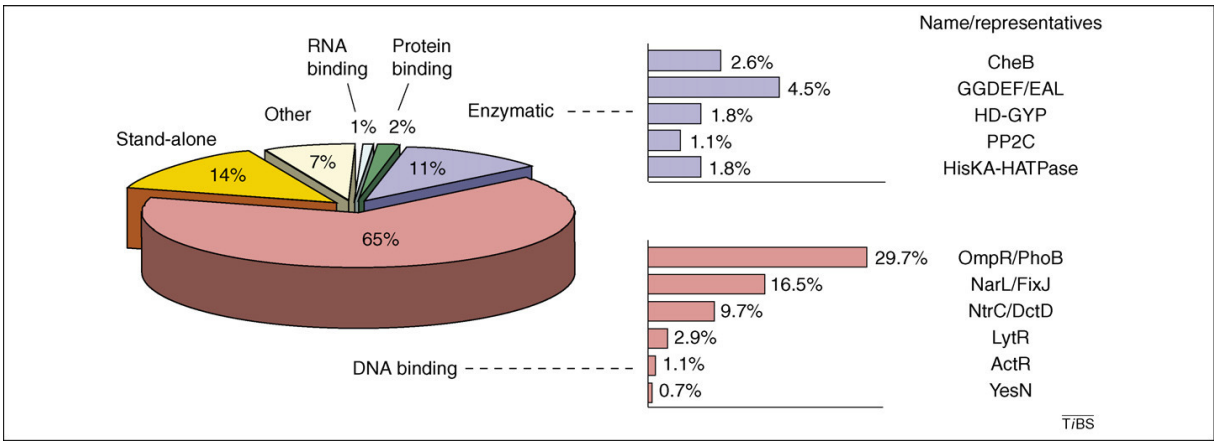

Figure 3. Bacterial RRs classified according to function. Distribution of percentage values are based on analyzes of ≈ 9000 bacterial RRs. Reprinted from (Gao *et al.*, 2007), with permission from Elsevier.

RRs activated by HPK_{10} type histidine kinases have DNA binding domains belonging to the LytTR (or LytR) family (Grebe & Stock, 1999; Johnsborg & Håvarstein, 2009; Nikolskaya & Galperin, 2002). In this family the DNA binding domain is composed of β-strands, unlike the typical helix-turn-helix structure found in most response regulators containing a DNA binding output domain (Nikolskaya & Galperin, 2002; Sidote *et al.*, 2008). RRs of the LytTR family are known to bind as dimers to degenerated 9-10 bp direct repeats in their target promoters (Risøen *et al.*, 1998; Ween *et al.*, 1999). *S. pneumoniae* harbours two RRs of the LytTR family type: ComE, regulating natural competence, and BlpR, regulating bacteriocin production (Galperin, 2006). The BlpR homologue StbR from *S. thermophilus* belongs to the LytTR family as well (Galperin, 2006).

Natural genetic transformation in the genus *Streptococcus*

Natural genetic transformation in bacteria is a natural mechanism for active uptake and heritable integration of naked DNA from the environment. Frederick Griffith discovered this phenomenon in *Streptococcus pneumoniae* as early as 1928, based on his observation that mice did not survive injections with a non-pathogenic mutant strain if they were mixed together with heat-killed pathogenic pneumococci (Griffith, 1928). We now know that the explanation for this surprising result is that the avirulent strain took up DNA from the dead wild type strain, and thereby reverted to a virulent strain. Even though many years have passed and lots of research has been done since Griffith reported his observations, scientists worldwide are still studying this fascinating feature of streptococci. Extensive studies have been done on natural genetic transformation in the model species *S. pneumoniae*. Figure 4 displays a simplified sketch of the regulation of natural genetic transformation in *S. pneumoniae*.

S. pneumoniae

Natural genetic transformation in *Streptococcus pneumonia* is regulated by quorum sensing in response to environmental signals. When grown under laboratory conditions, *S. pneumoniae* is naturally transformable during a brief period of time at the beginning of the exponential growth phase and is then referred to as being competent. The development of competence in *S. pneumoniae* depends on two sets of essential genes (Figure 4 and 5). Some of the proteins encoded by the so-called early competence genes make up the quorum sensing system controlling the competent state. A subfraction of the late competence genes encodes proteins involved in binding, uptake and integration of extracellular DNA into the competent cell's genome. ComX is the alternative sigma factor specifically required for transcription of the late competence genes, linking expression of the early and late competence genes.

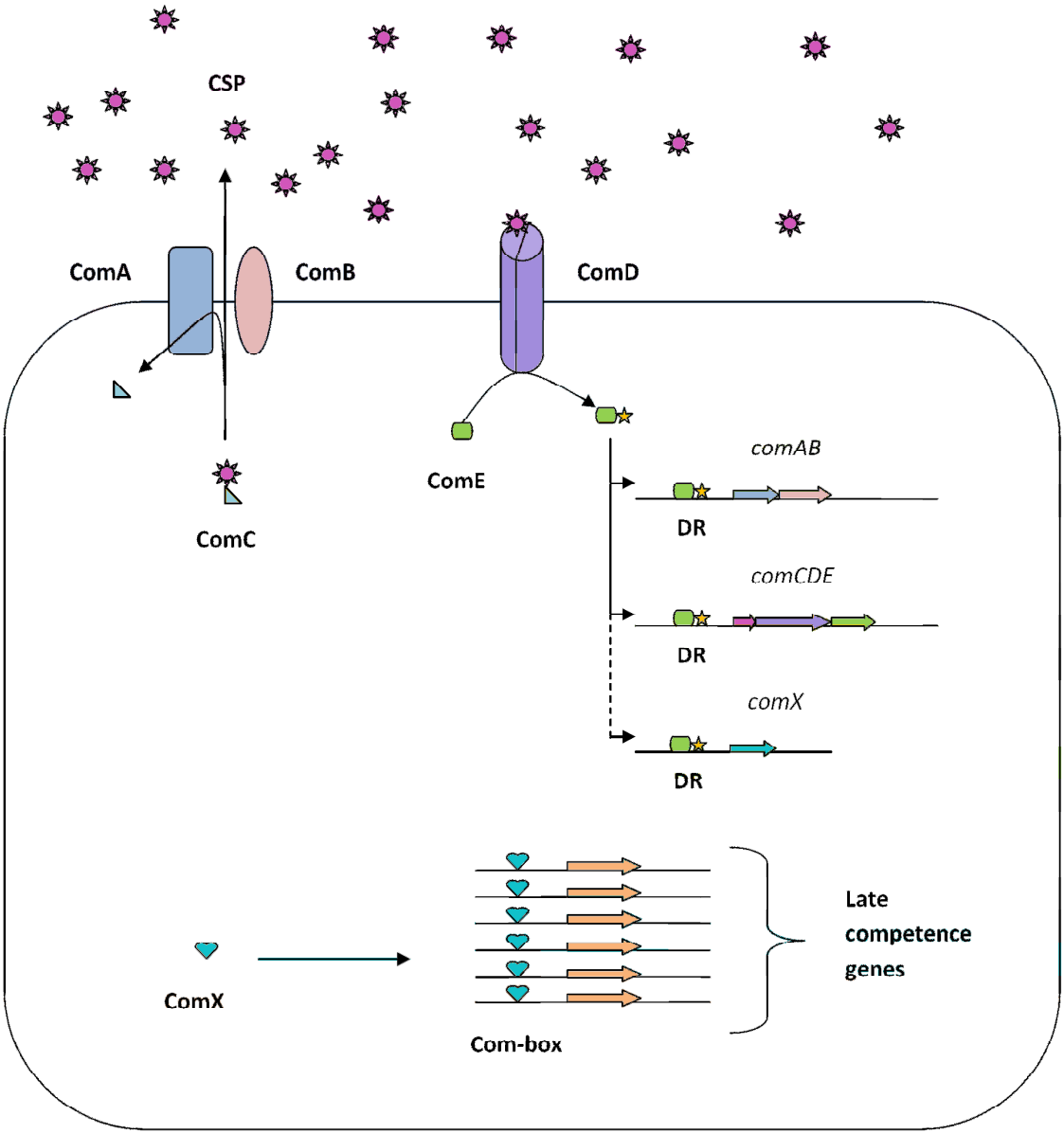

Figure 4. Regulation of natural genetic transformation in *S. pneumoniae*. The early competence genes (*comABCDE*) encode the quorum sensing system controlling the competent state. ComX is the alternative sigma factor responsible for the induction of the late competent genes encoding proteins involved in binding, uptake and integration of extracellular DNA. ◺ leader peptide, ★ phosphoryl group. For further details see text.

Five gene products (ComAB and ComCDE) make up the quorum sensing circuit that regulates competence development in *S. pneumoniae* by producing, secreting and monitoring the extracellular concentration of the competence stimulating peptide (CSP). The precursor of the competence stimulating peptide is encoded by *comC*, and contains a Gly-Gly type of leader peptide (Håvarstein *et al.*, 1995a). It is transported across the cytoplasmic membrane by the ABC transporter ComAB (Hui & Morrison, 1991; Hui *et al.*, 1995). The leader peptide is removed by ComA concomitant with export (Håvarstein *et al.*, 1995b; Ishii *et al.*, 2006; Kotake *et al.*, 2008). Binding of the secreted CSP to its histidine kinase receptor (ComD) is believed to stimulate transfer of a phosphoryl group from the histidine kinase to the cognate response regulator (ComE) (Håvarstein *et al.*, 1996; Pestova *et al.*, 1996). Phosphorylated ComE binds to a 9 bp imperfect direct repeat motif found in the promoter regions of early competence genes (designated DR in Figure 4 and 5) (Ween *et al.*, 1999). When ComE activates transcription of the early genes *comAB* and *comCDE* it generates an autoinduction circuit. As a result the amount of extracellular CSP will increase, which in turn will increase the amount of phosphorylated ComE inside the cell. At a critical CSP concentration ComE will elicit a chain reaction in which the alternative sigma factor ComX will induce transcription of the late competence genes (Lee & Morrison, 1999). ComX is regarded as the product of an early competence gene, but it has not been experimentally proven that the promoter region of the *comX* gene contains a ComE-binding site. However, a DR-like motif which deviates from the consensus sequence at two positions has been identified upstream of *comX* (Håvarstein, 2003). It has therefore been assumed that the promoter of *comX* will bind phosphorylated ComE (ComE-P) with lower affinity than the promoters of the other early genes. A high concentration of ComE-P in the cell may therefore be needed to turn on the expression of *comX* and the late genes (Håvarstein, 2003). In addition to ComE, ClpP and ComW are also known as regulators of ComX. ComX is stabilized in Δ*clpP* mutants indicating that the level of ComX is negatively regulated by ClpP-mediated proteolyses (Chastanet *et al.*, 2001; Robertson *et al.*, 2002; Sung & Morrison, 2005). The early competence gene *comW*, which is needed for induction of the competent state in *S. pneumoniae*, represents an additional level of regulation (Peterson *et al.*, 2004). ComW protects ComX from ClpP-mediated proteolyses, but it is not needed for *comX* expression (Sung & Morrison, 2005).

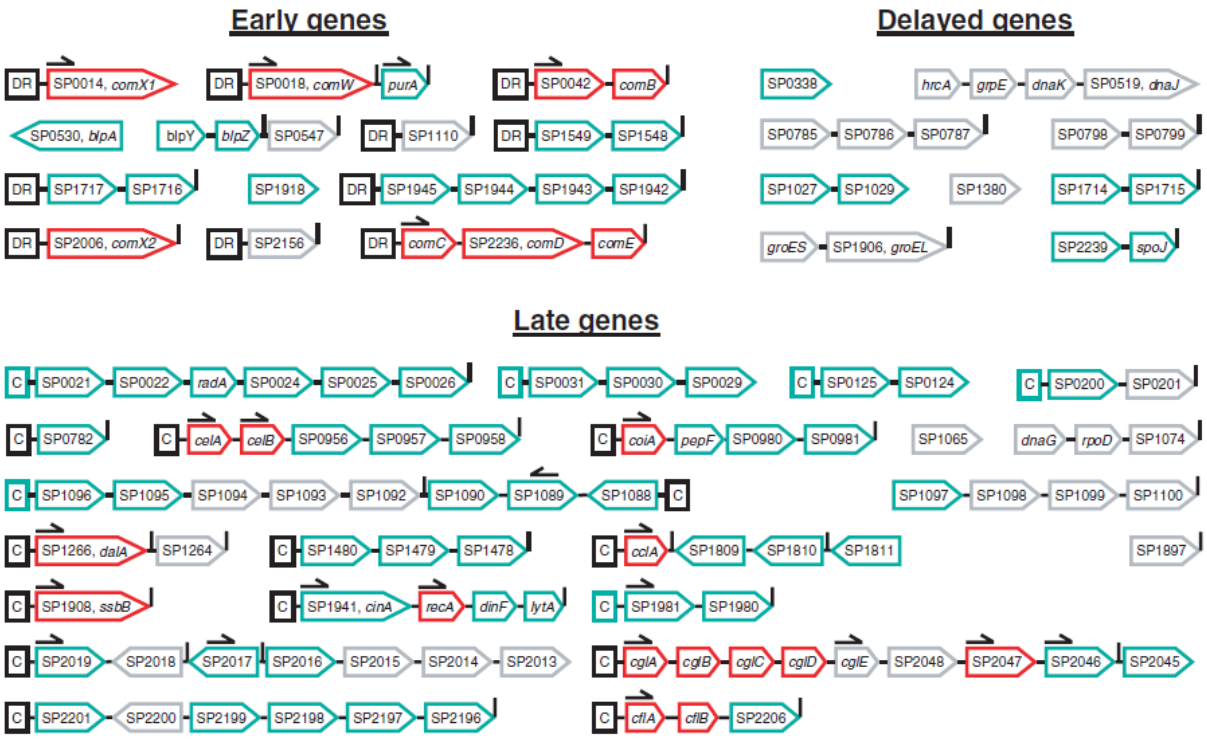

Figure 5. Organization and functional characterization of CSP-induced genes in *S. pneumoniae* identified by PCR amplicon-based microarrays. Genes required for competence are in red, those that are dispensable for competence are in green, and essential or otherwise unanalyzed genes are in grey. Pentagons (not in scale) indicate the open reading frame orientation within each cluster. In cases where sense direction of transcription for individual genes is determined by reporter fusions this is indicated by arrows. Vertical lines indicate location of putative terminators. DR: ComE binding site (direct repeat), c: ComX binding site (combox promoter). Reprinted from (Peterson *et al.*, 2004), with permission from John Wiley and Sons.

The promoters of the late competence genes contain a conserved -10 sequence (TACGAATA) called the combox or cinbox (marked as C in Figure 5), constituting the binding motif of ComX (Campbell *et al.*, 1998). When late competence genes are transcribed the alternative sigma factor ComX replaces the primary sigma factor (σ^A) in the RNA polymerase holoenzyme (Lee & Morrison, 1999).

In addition to *S. pneumoniae*, nine of twelve other naturally transformable *Streptococcus* species contain early competence genes (Gardan *et al.*, 2009; Johnsborg *et al.*, 2007). For these nine species (*S. oralis*, *S. peroris*, *S. mitis*, *S. gordonii*, *S. cristatus*, *S. sanguinis*, *S. anginosus*, *S. intermedius* and *S. infantis*) competence is presumably regulated in the same manner as for *S. pneumoniae*.

S. mutans

S. mutans is a major cause of human dental caries, and a leading cause of infective endocarditis (Banas, 2004). As in *S. pneumoniae*, the development of competence for natural transformation in *S. mutans* is regulated by quorum sensing in response to environmental signals (Ahn *et al.*, 2006; Perry *et al.*, 2009). Although the alternative sigma factor ComX and the late gene products are homologues of the corresponding competence proteins found in *S. pneumonia*, there are some differences between *S. mutans* and *S. pneumoniae* when it comes to regulation of competence induction. The early competence genes in *S. mutans*, whose products are involved in CSP secretion and sensing, are more closely related to the *blp* than the *com* system in *S. pneumoniae* (Martin *et al.*, 2006). In *S. mutans* there is no binding site for ComE in the promoters of the early genes *comAB*, *comC* or *comDE* (Martin *et al*., 2006; van der Ploeg, 2005). The autoinduction circuit driven by ComE in *S. pneumoniae* is therefore missing in *S. mutans*. This is assumed to be the reason for the delayed competence development in *S. mutans* compared to *S. pneumoniae*. While the competence stimulating peptide (CSP) induces competence only minutes after addition to cultures of *S. pneumoniae* (Morrison, 1997), it takes two hours to achieve the maximum competent state in *S. mutans* (Kreth *et al.*, 2005). ComX is needed for induction of competence both in *S. pneumoniae* and in *S. mutans* (Lee & Morrison, 1999; Li *et al.*, 2002), but the early competence gene *comW*, whose product is needed for stabilization of ComX in *S. pneumoniae*, is not present in *S. mutans* (Martin *et al.*, 2006). The lack of a ComE binding site in the promoter of *comX* (van der Ploeg, 2005), and the fact that a Δ*comE* mutant of *S. mutans* is still transformable at a fairly high residual level (Li *et al.*, 2001b), indicate that *comX* expression is presumably not directly regulated by ComE in *S. mutans*.

The ecological niche of *S. mutans* is a multispecies biofilm known as dental plaque, a major cause of human dental caries. The early competence genes encoding the quorum sensing system essential for genetic competence in *S. mutans* have been demonstrated to be involved in biofilm formation (Li *et al.*, 2002; Zhang *et al.*, 2009). While the *comCDE* genes are situated on the same transcription unit in *S. pneumoniae*, the *comC* gene in *S. mutans* is encoded on the complementary strand relative to *comDE* and is transcribed in the opposite direction (Li *et al.*, 2001b). Interestingly, with regard to biofilm formation, a *S. mutans* Δ*comC* mutant behaves very differently from a Δ*comD* (or Δ*comE* or Δ*comX*) mutant. While

biofilm formed by the Δ*comC* mutant contains very long chains of cells and lacks the wild type architecture, the Δ*comD*, Δ*comE* or Δ*comX* mutants form biofilm with reduced biomass. It is therefore believed that transcription of *comC* may not only be correlated with *comDE* for development of competence, but may also be regulated by a yet unknown regulator (Li *et al.*, 2002). In addition to regulating competence (Ahn *et al.*, 2006) and biofilm formation (Li *et al.*, 2002; Zhang *et al.*, 2009), the ComC-ComDE signal transduction pathway in *S. mutans* has also been shown to regulate features like bacteriocin production (Kreth *et al.*, 2006; van der Ploeg, 2005), acid tolerance (Li *et al.*, 2001a), and bacteriocin-mediated self-lysis in a subfraction of the population (Perry *et al.*, 2009).

S. thermophilus

Before the current work, *S. thermophilus* had never been shown to be a naturally transformable species. In her master thesis, Ingunn Alne demonstrated that *S. thermophilus* encodes homologues of two late competence genes, *celA* and *celB* (also called *comEA* and *comEC* respectively) (Alne, 2001). Later, sequencing of the whole genome verified this finding and demonstrated that all essential late competence genes are present in the *S. thermophilus* genome (Bolotin *et al.*, 2004). Due to the long history of *S. thermophilus* as a starter strain, we speculated that it might have lost the ability to develop the competent state, as a result of evolving in a stabile habitat quite different from the natural habitats of other streptococci. Living in a stable controlled environment like milk in the dairy industry seems to be a rather protected milieu compared to those of other streptococci, whose habitats, as members of the commensal flora of human or animals, appears stressful and competitive. The question we sought to answer was therefore whether the competence regulon of *S. thermophilus* is still functional or whether it has degenerated as a result of the domestication of this bacterial species.

Genetic engineering in lactic acid bacteria

Opportunities and challenges of genetic engineering in lactic acid bacteria (LAB)

Research on many bacterial species is severely hampered by the lack of efficient methods for introducing foreign DNA into these bacteria. When an efficient procedure for transformation has been established, and the genome sequence is available, the possibilities for genetic engineering and molecular research are almost unlimited. In practice genetic engineering also represents a challenge. It can be difficult to choose from multiple options, and systems that work well in one species or strain will not necessarily work well in others. Different codon usage in the recipient and donor may result in unsuccessful or reduced expression of heterologous genes, and products of imported genes may be unstable or insoluble in the new host. Export of foreign proteins may cause problems as well, or the product may be subject to extracellular degradation. In some cases genes with high homology to each other may have different function in different species, or their regulation might be different.

In general, a selection marker is used to identify bacteria that have taken up foreign DNA in a heritable manner. Antibiotic resistance genes are commonly used for this purpose, but are not desirable in the dairy industry due to the risk of spreading these genes to other bacteria occupying the human alimentary tract. Alternative food grade markers, or other techniques for identifying positive transformants, are therefore required. Naturally occurring bacteriocin resistance genes or genes required for the metabolism of rare sugars like xylose, inulin and sucrose can be used as food grade selection markers (de Vos, 1999). Furthermore, cryptic plasmids from GRAS organisms may be used in the construction of food grade vectors (Shareck *et al.*, 2004). Vector pHRM1 (6,4 kb) (El Demerdash *et al.*, 2003), which is derived from the cryptic *S. thermophilus* plasmid pSt08, uses a small heat shock protein encoded by the *shsp* gene (from pSt04) as a selection marker. *S. thermophilus* transformants harbouring the pHRM1 vector have increased resistance to incubation at 60˚C or pH 3.5, and are able to grow at 52˚C. This marker was demonstrated to be as efficient as an erythromycin marker when selecting for transformants (El Demerdash *et al.*, 2003). It is very important for the dairy industry that food grade cloning systems are stable. Consequently, genetic manipulation of the chromosome will be preferred over more unstable plasmid systems. The loci of lactose

metabolizing genes have been used as a genomic integration target, taking advantage of the *lacZ* gene as a food grade marker (Mollet *et al.*, 1993).

For the dairy industry it is of outmost importance that their products are safe. All genetic manipulations of dairy organisms must therefore be made in a food grade manner, and according to strict rules defined by the European Food Safety Authority (EFSA) or the U.S. Food and Drug Administration. To the best of my knowledge, no genetically modified starter used by the dairy industry has so far been approved by the EFSA. In "Directive 2009/41/EC of the European Parliament and of the Council", dealing with the contained use of genetically modified microorganisms, a genetically modified organism (GMO) is defined as "*an organism in which the genetic material has been altered in a way that does not occur naturally through fertilisation and/or natural recombination*". According to Directive 2009/41/EC:

> **"i) genetic modification occurs at least through the use of the following techniques:**
>
> 1. Recombinant nucleic acid techniques involving the formation of new combinations of genetic material by the insertion of nucleic acid molecules produced by whatever means outside an organism, into any virus, bacterial plasmid or other vector system and their incorporation into a host organism in which they do not naturally occur but in which they are capable of continued propagation.
> 2. Techniques involving the direct introduction into a micro-organism of heritable material prepared outside the micro-organism including micro-injection, macro-injection and micro-encapsulation.
> 3. Cell fusion or hybridisation techniques where live cells with new combinations of heritable genetic material are formed through the fusion of two or more cells by means of methods that do not occur naturally
>
> **ii) the following techniques are not considered to result in genetic modification:**
> (1) *in vitro* fertilisation;
> (2) natural processes such as: conjugation, transduction, transformation;
> (3) polyploidy induction.
>
> on condition that they do not involve the use of recombinant-nucleic acid molecules or GMMs (genetically modified micro-organisms) made by techniques/methods other than:

1. Mutagenesis.

2. Cell fusion (including protoplast fusion) of prokaryotic species that exchange genetic material by known physiological processes.

3. Cell fusion (including protoplast fusion) of cells of any eukaryotic species, including production of hybridomas and plant cell fusions.

4. Self-cloning consisting in the removal of nucleic acid sequences from a cell of an organism which may or may not be followed by reinsertion of all or part of that nucleic acid (or a synthetic equivalent), with or without prior enzymic or mechanical steps, into cells of the same species or into cells of phylogenetically closely related species which can exchange genetic material by natural physiological processes where the resulting micro-organism is unlikely to cause disease to humans, animals or plants.

Self-cloning may include the use of recombinant vectors with an extended history of safe use in the particular micro-organisms"

According to above-mentioned definitions, EU legislation today seems more focused on the methodology used to make the mutation rather than the nature of the genetic modification. Consequently, organisms that have been subjected to random mutagenesis, by for example exposure to mutagenic radiation or chemicals, are not regarded as GMOs. However, if recombinant DNA techniques have been used to introduce mutations, even in the case of point mutations and small deletions which occur naturally in a growing population of bacteria, the resulting strains are regarded as GMOs.

The NICE system – an inducible food grade gene expression system

The nisin-controlled gene expression system (NICE) is widely used as a tool for induced gene expression in Gram-positive bacteria (Mierau & Kleerebezem, 2005). Nisin is produced by *Lactococcus lactis*, and its production is regulated by a classic two-component signal transduction system. It is a natural antimicrobial peptide that effectively inhibits growth of both Gram-positive and Gram-negative bacteria and the outgrowth of Bacilli and Clostridia spores (de Arauz *et al.*, 2009). Today nisin is widely used as a food preservative. Since nisin was first marketed in England in 1953, it has been approved for food applications in 48 countries (de Arauz *et al.*, 2009). Consequently nisin, and the two-component signal transduction pathway (NisR and NisK) regulating nisin expression, have been promoted as a food grade inducible expression system.

The *nis* cluster consist of 11 genes (Figure 6) (Kuipers *et al.*, 1995) and contains three promoters: nisA, nisR and nisF. The nisA promoter regulates the expression of the *nisABTCIP* operon, where *nisA* encodes the nisin A precursor. The gene products of *nisBTCIP* are involved in modification, translocation, immunity, and processing of nisin. The nisR promoter controls expression of the two-component signal transduction system NisRK. The HK is encoded by the *nisK* and the RR by *nisR*. The last promoter, nisF, controls expression of the genes *nisFEG*, encoding an additional immunity system protecting against nisin. NisA and nisF promoters are autoregulated by nisin in a concentration dependent manner, while the *nisRK* operon is constitutively expressed independently of nisin (de Ruyter *et al.*, 1996a).

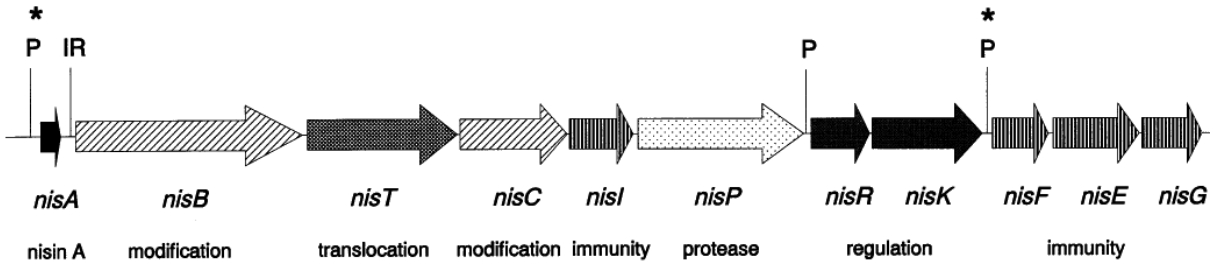

Figure 6. The nisin A gene cluster. Promoters are marked with flags, asterisk indicate that the promoter are regulated by the two-component signal transduction system, NisRK (response regulator and histidine kinase, respectively). Reprinted from (Kuipers *et al.*, 1998), with permission from Elsevier.

A β-glucuronidase gene from *E. coli* was the first heterologous gene expressed by nisin induction in *L. lactis* (Kuipers *et al.*, 1995). The nisA promoter was fused to the promoterless reporter gene *gusA* on plasmid pNZ273 resulting in a construct named pNZ8008. This construct was transformed into *L. lactis* NZ9800, a strain with a disrupted *nisA* gene (Δ*nisA*). The resulting strain produced β-glucuronidase when induced by nisin, and the level of production was directly related to the amount of inducer added (Kuipers *et al.*, 1995). When the *nisRK* operon was integrated in the chromosome of MG1614, a strain lacking the nisin A gene cluster, β-glucuronidase activity was observed after nisin induction. This demonstrated that in the presence of the *nisRK* operon, controlled expression of recombinant proteins is possible when the gene of interest is fused to the *nisA* promoter. The *nisRK* operon may be present on the chromosome naturally, or introduced by replacement recombination.

Alternatively the *nisRK* operon may be placed on a plasmid (Kleerebezem *et al.*, 1997). In order to control the expression of a target gene in a strain already containing the nisin gene cluster, the *nisA* gene (or genes involved in the export of nisin, e.g. *nisT*) must be deleted.

Expression systems based on NICE are versatile for several reasons. They originate from a food grade source (*L. lactis*), and have a broad host spectrum. The system has already been successfully used in *S. pneumoniae*, *S. agalactiae*, *S. pyogenes*, *Enterococcus faecalis* and *Bacillus subtilis* (Eichenbaum *et al.*, 1998). The *nisRK* are the only genes required for signal transduction, and only sub-lethal amounts of nisin are needed for effective induction (Kuipers *et al.*, 1995). Induction by nisin displays a linear dose-response curve, making it possible to produce toxic gene products in sub-inhibitory amounts (Kuipers *et al.*, 1995). The NICE system can also be used for overexpression of genes. High level production of aminopeptidase N, corresponding to 47 % of total intracellular protein content, has been achieved (de Ruyter *et al.*, 1996b). The inducer, nisin, is already used in food products, and the amount needed for induction is available at low cost. Despite the advantages, there might be species or strains where the NICE system does not work well. In these cases, similar systems may be identified and used, resulting in tailor-made expression systems consisting solely of genetic elements from the strain of interest.

Potential of genetically modified LAB

There are several benefits associated with the use of LAB in food production. The fermentation process causes lactic acid production and acidification of the product, thereby preventing development of spoilage microorganisms and possible pathogens. Consequently, the shelf-life of the fermented product is improved (Auclair & Accolas, 1983). Some strains will also produce bacteriocins which will further increase the antibacterial effect (Cotter *et al.*, 2005). The organoleptic qualities of the final product will also be influenced by LAB. In the production of yogurt both consistency and flavour depend of the metabolism of the lactic starter(s) (Auclair & Accolas, 1983). The rheological characteristic of dairy products will for example be influenced by the bacterial production of exopolysaccharides (EPS) (Purohit *et al.*, 2009), while hydrolysis of milk proteins by peptidases from LAB is important for flavour

formation in many types of cheeses (Smit *et al.*, 2005). Considering that the properties of the starter strain(s) have a strong impact on the quality of the fermented product, there should be a strong potential for genetic engineering in LAB. Commercial strains can be genetically modified to improve inherent properties, introduce desirable characteristics or remove unwanted traits. In *L. lactis,* for example, a food grade nisin-inducible expression system that was used to induce lysis of a starter strain, facilitated release of intracellular enzymes involved in flavour formation and accelerated cheese ripening (de Ruyter *et al.*, 1997).

Production of Mozzarella represents one example of how the dairy industry might take advantage of genetically modified lactic acid bacteria. *S. thermophilus* is used alone or together with lactobacilli in the production of Mozzarella. Residual galactose in Mozzarella cheese is associated with a browning defect in cooked products (Sanders, 1991). *S. thermophilus* is known to prefer lactose as a carbon source (van den Bogaard *et al.*, 2004), and the lactose enters the cells by a lactose/galactose antiporter system, the LacS permease (Gunnewijk *et al.*, 1999). When lactose is hydrolyzed, LacS will transport the galactose moiety out of the cell in exchange for lactose. So when *S. thermophilus* grows on lactose, galactose will be released to the environment, accumulating in the Mozzarella cheese. Most *S. thermophilus* strains are unable to metabolize galactose. The challenge is therefore to prevent accumulation of galactose in the cheese. Several studies have been published where different research groups have tried different strategies to select, or make, a stable galactose positive *S. thermophilus* strain (de Vin *et al.*, 2005; Robitaille *et al.*, 2007; Thomas & Crow, 1984; Vaillancourt *et al.*, 2004; Vaughan *et al.*, 2001). One very interesting industrial yogurt starter identified is the IMDOST40 strain, which does not excrete the galactose moiety during growth on lactose (de Vin *et al.*, 2005). In this strain the glucose and galactose moieties of lactose are both utilized. IMDOST40 may not be a good strain for Mozzarella production, but it may provide information on how strains already used in such production could become able to metabolize galactose through genetic engineering.

The potential use of recombinant LABs goes beyond the dairy industry. In the last years the interest in genetically modified LAB as vehicles for mucosal delivery of therapeutic and prophylactic compounds, as well as DNA vaccines, has increased substantially (Wells & Mercenier, 2008). So far, lactococci and lactobacilli have been most frequently used for these

purposes, probably due to better developed molecular toolboxes for members of these genera compared to other LAB with GRAS status. Lactococci have successfully been used to produce cancer antigens, tetanus antigens and cytokines (Wells & Mercenier, 2008; Diep *et al.*, 2009 and references therein). The cytokine murine interleukin-10, known to be useful for treatment of inflammatory bowel disease, was used in the first clinical trial with recombinant LAB as delivery vehicles (Braat *et al.*, 2006). Patients with Chron's disease were treated with living, genetically modified *L. lactis* in which the thymidylate synthase gene (*thyA*) on the genome was replaced with a synthetic sequence encoding the mature human interleukin-10 gene (*hIL10*). In this strain, the only foreign DNA was the human *hIL10* gene. By replacing the *thyA* gene with *hIL10* the strain became dependent on thymidine or thymine for growth and interleukin-10 production (Steidler *et al.*, 2003). This is a very elegant way to prevent GMO accumulation in the environment. Thymidine is not present outside the human body in amounts that would support growth of the GMO, implying that the genetically modified *L. lactis* strain will die after passage through the intestine. If the strain restores viability outside the human intestine through acquisition of an intact *thyA* gene, the genetic modification will be reversed during the recombination event resulting in a wild type *L. lactis* strain.

OBJECTIVES AND OUTLINE OF THE THESIS

Several different species belonging to the genus *Streptococcus* are competent for natural genetic transformation. However, most members of this genus, including *S. thermophilus*, have never been shown to possess this property. At the onset of this work the genome sequence of *S. thermophilus* had not been determined. We therefore did not know whether *S. thermophilus* possesses all of the genes required for natural transformation in *S. pneumoniae*. However, an earlier study by Ingunn Alne showed that *S. thermophilus* encodes homologues of the pneumococcal late competence genes *celA* and *celB* (Alne, 2001). For this reason we speculated that *S. thermophilus* might be competent for natural transformation, despite the fact that spontaneous competence development has never been observed for this bacterium. In addition to being very interesting it its own right, a way to induce the competent state in *S. thermophilus* would also be of great practical importance for the scientific community and the dairy industry. In contrast to existing molecular tools available for *S. thermophilus*, natural transformation represents a highly versatile and efficient mechanism that can easily be exploited for genetic manipulation of this bacterium. Natural transformation would be a very useful tool for studies aimed at improving our knowledge of the physiology and metabolism of *S. thermophilus*, and could also be used to generate mutants with novel and unique value-added properties. The major goal of the present work was therefore to determine whether or not *S. thermophilus* is a naturally transformable species. If the answer to this question turned out to be positive, an additional objective would be to investigate the potential of natural transformation as a tool for genetic engineering of *S. thermophilus.*

Paper **I.** "Pheromone-induced expression of recombinant proteins in *Streptococcus thermophilus*"; describes the development of an inducible expression system for this bacterium. The system, whose natural function is to control bacteriocin production in *S. thermophilus* LMG18311, consists of an unmodified peptide pheromone, its histidine kinase receptor, its cognate response regulator, and a pheromone-responsive bacteriocin promoter. To identify functional elements in the bacteriocin promoter, deletion mapping and site-directed mutagenesis were carried out. Our results showed that a degenerated direct-repeat motif is essential for pheromone-inducible transcriptional activation, and presumably represents the sequence motif recognized by the response regulator.

Paper **II.** In the article entitled; "Natural genetic transformation: a novel tool for efficient genetic engineering of the dairy bacterium *Streptococcus thermophilus*"; the activity of the promoters driving the expression of the gene encoding the competence specific sigma factor ComX and the late competence operon *comAE-comEC* were studied by fusing these promoters to a luciferase reporter gene. Under laboratory conditions low-level activity of the *comX* promoter was detected in early logarithmic growth phase, whereas the *comEC* promoter was virtually inactive. However, when using the system described in paper I to overexpress c*omX*, a strong activation of the *comAE-EC* promoter was observed. This result indicated that all late competence genes were expressed under these conditions, and that the *S. thermophilus* LMG18311 cells might have developed the competent state. Transformation experiments with naked extracellular DNA carrying an antibiotic resistance marker showed that this was indeed the case, demonstrating that *S. thermophilus* LMG18311 has preserved an intact set of late competence genes despite the fact that its genome has been found to contain a large number of pseudogenes. This new powerful technique makes genetic engineering in *S. thermophilus* simple and efficient, and makes it possible to construct starter strains with novel and/or improved properties.

Paper **III.** In the paper; "A food-grade site-directed mutagenesis system for *Streptococcus thermophilus* LMG 18311"; the genetic tools developed in paper II are used for genetic engineering of the LMG18311 strain. Due to the high efficiency of the natural transformation system, we were able to perform site-directed mutagenesis without the use of a selection marker. Our results demonstrate that the method described can be used to introduce specific changes anywhere in the genome of the *S. thermophilus* LMG 18311 strain without leaving behind any foreign DNA.

MAIN RESULTS AND DISCUSSION

S. thermophilus LMG18311 possesses a non-functional bacteriocin locus

Proteins encoded by the *stbABCHR* locus of the *S. thermophilus* LMG18311 strain are orthologues of the corresponding proteins encoded by the *blpABCHR* locus regulating bacteriocin production in *S. pneumoniae* (de Saizieu *et al.*, 2000; Reichmann & Hakenbeck, 2000). In *S. pneumoniae*, BlpC, the precursor of the mature signalling peptide, is processed and secreted by the ABC transporter BlpAB. The receptor of the signalling peptide is the membrane-bound histidine kinase BlpH, which together with its cognate response regulator BlpR constitute a two-component regulatory system. In *S. thermophilus* LMG18311 there is a frameshift mutation in the *stbB* gene that probably renders the transporter non-functional. This assumption has been confirmed by a later study showing that bacteriocin production can be observed for *S. thermophilus* strain LMD-9, which has an intact *stbB* gene, but is missing in two strains, LMG 18311 and CNRZ 1066, possessing truncated versions of this gene (Fontaine *et al.*, 2007). Comparison of the StbHR proteins from LMG18311 with the corresponding two-component systems from *S. pneumoniae* indicated that this system is still functional in strain LMG18311.

Construction of an inducible expression system

As the StbHR proteins of *S. thermophilus* LMG 18311 appeared to be functional, we set out to develop an inducible expression system based on this signal transduction pathway (Paper I). We assumed that the signalling peptide controlling the activity of the StbHR system is encoded by the *stbC* gene, even though the amino acid sequence of the *stbC* product is completely different from the product of the corresponding *blpC* gene. However, both proteins contain a so called double-glycine type leader peptide at their N-terminal ends (Håvarstein *et al.*, 1994). As the processing site of GG-leaders is known to be immediately downstream of the two glycine residues, a peptide corresponding to the predicted mature signalling peptide was synthesized chemically. To test whether this peptide, termed the STP-pheromone, is able to activate the StbHR system, the *gusA* reporter gene was cloned into the pTRKH2 vector behind the promoter of the putative bacteriocin gene *stbD*. Indeed, after electroporation of the resulting construct into the LMG18311 strain we were able to show that

the rate of transcription of *gusA* increased dramatically when cells harbouring the reporter plasmid were subjected to the STP-pheromone. The observed STP-activated *gusA* expression proved that the StbHR signal transduction system is functional and can be used for inducible expression of genes placed behind the *stbD* promoter. Unfortunately, the *stbD* promoter was not completely shut off in the absence of STP, giving rise to some background activity in uninduced cells (Paper I, Figure 3). This makes our system less useful in situations where a leaky promoter represents a problem, for instance when expressing toxic proteins. A mutational analysis of the *stbD* promoter was therefore carried out, with the aim to map the binding site of StbR, and to identify sequence elements responsible for the leakiness of this promoter.

Identification of a putative binding site for StbR in the promoter region of *stbD*

The response regulator, StbR, which activates transcription from the *stbD* promoter, has a DNA binding domain that belongs to the LytTR family (Galperin, 2006; Nikolskaya & Galperin, 2002). Previous studies have revealed that response regulators in this family bind as dimers to degenerated 9-10 bp direct repeat motifs in their target promoters (Risøen *et al.*, 1998; Ween *et al.*, 1999). The two repeated sequences are separated by a 12 bp spacer region, and the whole sequence motif is separated from the downstream Pribnow box by a stretch of 31-32 nucleotides. Through deletion mapping experiments we were able to identify a short region in the *stbD* promoter that was required for STP induced *gusA* expression. This region contained a direct repeat-like motif consisting of two 9 bp repeats located 31 bp upstream of a predicted Pribnow box (Paper I, Figure 2A). However, in contrast to previously studied direct repeat motifs (Risøen et al., 1998; Ween et al., 1999), the repeats in the *stbD* promoter are separated by eleven instead of twelve base pairs. Earlier studies have shown that the length of this spacer is crucial for promoter activity, and that removal of just one nucleotide renders the promoter non-functional (Knutsen *et al.*, 2004; Risøen *et al.*, 2001). Interestingly, introduction of an extra base pair between the left and right repeat (Mut 12, Paper I) essentially eliminated the inducibility of the *stbD* promoter. This result confirmed that the length of the spacer is very important. The reason for this is presumably that StbR binds as a dimer (Knutsen *et al.*, 2004; Risøen *et al.*, 1998; Ween *et al.*, 1999), and that the orientation of one monomer relative to the other is crucial for establishing intermolecular contact that stabilizes the dimer

complex. The angle of rotation between adjacent base pairs in double-stranded B-DNA is about 34°. Thus, addition or removal of a single base pair in the spacer region will rotate one StbR monomer by 34° with respect to the other around the DNA axis. It is therefore not surprising that changes in the length of the spacer region disrupt the functionality of the promoter.

Mutational analysis of the *stbD* promoter

As the *stbD* promoter unfortunately turned out to have some background activity in the absence of the STP-pheromone, it was of interest to identify the sequence element(s) responsible for the leakiness of this promoter. Sabelnikov *et al.* (1995) have found that so called extended –10 promoters, which contain a T-TG-extension of the -10 site (TNTGNTATAAT), sometimes function well without an accompanying -35 site in *S. pneumoniae*. A possible extended -10 promoter (TGCTATACT) was found to be present in the *stbD* promoter region (Paper I, Figure 2A). Using site directed mutagenesis this sequence was changed into AACTATACT to remove the conserved TG motif believed to be an essential functional element in the extended -10 promoter. Unfortunately, no reduction in background level transcription was observed for the resulting mutant (MutTG). However, the TG-mutant promoter displayed poor inducibility, strongly indicating that the TGCTATACT sequence constitutes the – 10 site of the inducible promoter controlled by StbR.

Recently, Sidote *et al.* (2008) determined the crystal structure of the DNA-binding domain of AgrA in complex with its target DNA. The response regulator AgrA together with the histidine kinase AgrC constitute a classical two-component regulatory system that controls the expression of virulence and other accessory genes in *Staphylococcus aureus* (Novick & Geisinger, 2008). The structure reported by Sidote and co-workers is the first to be solved for a DNA-binding domain of the LytTR family. The structure revealed that only three amino acids (H169, R233 and N201) form direct or water-mediated hydrogen bonds with bases in the 9 bp direct repeat sequence ($A_1C_2A_3G_4T_5T_6A_7A_8G_9$) specifically recognized by the LytTR domain of AgrA (Sidote *et al.*, 2008). Amino acid H169 interacts with G_9, whereas R233 interacts with A_1 and the complementary base to C_2 on the parallel strand. Amino acid N201 forms a water-mediated hydrogen bond with T_6, and also contacts the ribose of A_7. The latter

interaction is not base specific. Interestingly, the bases that are involved in hydrogen bonding with the H169, N201 and R233 residues of AgrA (A_1, indirectly C_2, T_6 and G_9) are also conserved in the direct-repeat motif recognized by the LytTR domain of StbR (ACCATTCGG - 11 bp - ACTTTTTGG). This indicates that these conserved bases, which play an essential role in AgrA-DNA interaction, perform the same function in the StbR system. However, the properties of the Mut11 and MutV mutants suggest that at least one of the bases in positions 3, 4 and 7 are important for StbR binding (Paper I, Figures 2B and 3). By exchanging the bases T_3, T_4 and T_7 in the right repeat into A_3, A_4 and C_7 (Paper I, Figure 2B, Mut11), increased expression of GusA was observed both in induced and uninduced cells. In contrast, substitution of the corresponding bases in the left repeat (C_3, A_4 and C_7) with T_3, T_4 and T_7 (Paper I, Figure 2B, MutV) resulted in considerable loss of inducibility. Alignment of the amino acid sequences of AgrA and StbR showed that only H169 is conserved in StbR (data not shown). Presumably, the N201 and R233 residues have been replaced by other amino acids, with slightly different positions in the StbR primary sequence, that make specific contacts with the critical bases in the direct repeat motif. Consequently, the bases that make direct contacts with the response regulators may not be identical in the AgrA and the StbR systems.

Our mutational analysis failed to identify the reason for the relatively high background activity of the *stbD* promoter. However, as a significant increase in *gusA* expression was observed in uninduced cultures of the Mut11 mutant, it is likely that the background activity arises as a result of StbR binding. It is known from the literature that in some cases unphosphorylated response regulators are able to activate low-level transcription from their target promoters (Ellison & McCleary, 2000; Geng *et al.*, 2004). It has also been reported that acetylphosphate, an intermediate in the pyruvate metabolism of *S. thermophilus*, might serve as a phosphoryl donor for response regulators (Wolfe, 2010). Thus it is not inconceivable that StbR could activate transcription from the *stbD* promoter even in the absence of the STP-pheromone. Other possible explanations for the STP-independent background activity of this promoter is cross-phosphorylation of StbR by other histidine kinases or transcriptional cross-activation by a related response regulator (Laub & Goulian, 2007). In *S. pneumoniae*, the response regulators BlpR and ComE regulates bacteriocin production and competence for natural transformation, respectively. BlpR and ComE both possess a LytTR-type DNA

binding domain (Galperin, 2006), and bind to similar but not identical direct repeat motifs. Interestingly, it has been demonstrated that BlpR as well as ComE are able to induce expression of the ABC transporter QsrAB. The reason for this is that the *qsrAB* promoter contains a direct repeat motif that is recognized by both response regulators (Knutsen *et al.*, 2004). The genome of *S. thermophilus* does not encode a response regulator related to ComE, but contains a gene (stu0543) encoding another response regulator with a LytTR-type DNA binding domain (Galperin, 2006). Together with its cognate histidine kinase (HK02), this response regulator (RR02) makes up a two-component regulatory system of unknown function. Thus, at least in theory, it is possible that the HK02/RR02 system might be responsible for the leaky behaviour of the *stbD* promoter.

Despite the fact that we were not able to prevent or explain the observed leakage from the *stbD* promoter, the expression system described in paper I still represents a valuable research tool for inducible expression of homologous and heterologous proteins in *S. thermophilus* LMG18311.

Is *S. thermophilus* competent for natural genetic transformation?

The ability to bring DNA into a living cell is a prerequisite for most studies involving gene manipulation in bacteria. Different types of artificial transformation, as for instance electroporation and heat shock/$CaCl_2$ treatment, are the most commonly used methods. In some cases, however, one of the natural systems, conjugation, transduction or natural transformation, will be a better choice. Many bacteria are poorly transformable or even impossible to transform with current artificial methods. In some of these cases natural transformation might represent an excellent alternative to artificial transformation. Genetic engineering based on natural transformation has the advantage that it does not depend on plasmids, but allows direct manipulation of the bacterial genome. Until now, 65 bacterial species have been demonstrated to be competent for natural genetic transformation (Johnsborg *et al.*, 2007). There is reason to believe, however, that the actual number is much higher. BLAST analysis of fully sequenced bacterial genomes available at the NCBI database (http://www.ncbi.nlm.nih.gov/) shows that competence genes are present in many bacteria that have never been observed to be naturally transformable in the laboratory. Most of the

species that belong to the genus *Streptococcus* possess so called late competence genes. These genes encode proteins involved in DNA binding, processing, uptake and recombination. The presence of these genes indicates that the bacterium carrying them is competent for natural transformation or that an ancestor of the bacterium had this property. At the onset of the present study it was known that the *S. thermophilus* genome encodes the late competence genes *comEA* and *comEC* (designated *celA* and *celB* in *S. pneumoniae*), and that the promoter region of these genes contains a ComX binding site (Alne, 2001). These findings strongly suggested that a ComX regulon that includes the late competence genes is present in *S. thermophilus.* Using degenerated primers based on sequence comparisons of streptococcal *comX* genes available in databases, we succeeded in amplifying and sequencing the complete *comX* gene from *S. thermophilus* by the end of 2002 (unpublished results). Approximately two years later the complete *S. thermophilus* genome was published (Bolotin *et al.*, 2004).

comX is expressed in early logarithmic phase

Knowing that *S. thermophilus* contains *comX* and the late competence genes *comEA* and *comEC* we wanted to determine whether these genes are transcribed under standard laboratory growth conditions. Two different reporter plasmids were therefore constructed by fusing the promoters of *comX* and *comEA* to a firefly luciferase gene (*luc*) in the shuttle vector pTRKH2 (O'Sullivan & Klaenhammer, 1993). The plasmids, pXP and pEAP, were introduced into *S. thermophilus* LMG 18311 by electroporation, giving raise to the strains XP and EAP. The amount of light produced was measured automatically while cultures of the XP and EAP strains were grown in an Anthos Lucy 1 luminometer at a preset temperature. Surprisingly, we discovered that both promoters were active during early to approximately mid-logarithmic phase (Paper II, Figure 1). However, the level of luciferase expression in the strain harbouring the pEAP plasmid was very low, indicating that too little ComX is produced under the conditions used to activate transcription of the late competence genes. In an attempt to identify conditions that would further increase the activity of the *comX* and *comEA* promoters we varied a number of different parameters, such as medium composition and temperature. We also tried to stress the *S. thermophilus* cells by adding mitomycin C (0.05 – 0.2 µg/ml), ethanol (1%) and NaCl (150 mM). The DNA damaging agent mitomycin C had previously been shown to induce the competent state in *S. pneumonia* (Prudhomme *et al.*, 2006). None of

these treatments dramatically increased the expression of the *luc* reporter gene. Among the conditions tested, cultivation of *S. thermophilus* at 37ºC in Todd-Hewitt medium supplemented with 0.8 % glucose (THG) gave the best overall results.

Overexpression of *comX* induces competence for natural transformation in *S. thermophilus* LMG 18311

As we did not succeed in identifying growth conditions or stress factors that induce competence for natural transformation in *S. thermophilus*, we considered that our best option would be to overexpress *comX* from an ectopic promoter. Luo *et al*. (2003) had already used this approach with success in *S. pneumonia.* Their strategy was to clone the *comX* gene into a multicopy plasmid behind a nisin-inducible promoter. Our attempt to use the same system in *S. thermophilus* was not successful, and we therefore had to rely on the pheromone induced expression system described above (Paper I). To determine the efficiency of the expression system in different media, a reporter plasmid was made in which the *luc* reporter gene was placed behind the *stbD* promoter. The highest level of light emission obtained was with STP-induced (250 ng/ml) cells grown at 37ºC in THG. To be able to monitor the effect of ectopic *comX* expression on the activity of late gene promoters, a combined expression and reporter plasmid was constructed. Two different DNA fragments were cloned in opposite directions into the polylinker of the pTRKH2 plasmid. One of the fragments consisted of the *stbD* promoter fused to the *comX* gene, while the other fragment consisted of the *luc* gene placed behind the *comEA* late gene promoter. Addition of the STP-pheromone to cultures of *S. thermophilus* LMG 18311 cells carrying this construct (pXL) resulted in an approximately 600-fold increase in luminescence compared to a corresponding culture of EAP. No increase in light emission was observed when cultures of the XP and EAP strains were subjected to STP. This demonstrates that the observed increase in luminescence was due to the STP-induced ectopic expression of *comX.* These very promising results suggested that the expression system might be powerful enough to induce the competent state, provided that all late genes required for transformation are still functional in LMG 18311. Indeed, about $3x10^3$ streptomycin resistant (Str^R) CFUs (Colony-Forming Units) per ml of culture were obtained when 1 µg/ml of genomic Str^R-DNA was added to cultures of the XL strain pretreated with the STP-pheromone. The total number of CFUs in the cultures was about $5x10^8$ per ml (Paper

II). To verify that the streptomycin resistant transformants resulted from uptake of Str^R-DNA by natural transformation, we disrupted the late competence gene *comEC* (also termed *celB*) known to be essential for natural transformation in *S. pneumoniae* and *Bacillus subtilis* (Berge *et al.*, 2002; Draskovic & Dubnau, 2005). Transformation experiments carried out with the XL Δ*comEC* strain gave no transformants, demonstrating that the positive results obtained with the XL strain must be due to natural genetic transformation.

After further optimization of the protocol, transformation efficiencies were approaching 1 % (CFU transformants x 100/CFU total) with the XL strain. Such results were achieved in experiments where a target gene was replaced by an antibiotic resistance marker or a modified version of the target gene by double-crossover recombination. The inserted sequence was flanked at both ends by approximately 1000 bp fragments that were homologous to the flanking regions of the target gene. Naked DNA taken up by a competent cell is nicked at a random site when binding to a receptor at the cell surface. Then one strand is degraded while entry of the other strand proceeds linearly from the newly formed 3'-end to the extremity of the donor fragment (Mejean & Claverys, 1988). To decrease the chance that nicking takes place within the fragment encompassing the marker gene and flanking regions, the entire fragment was cloned into the pCR2.1-TOPO plasmid before transformation. This strategy was later used in the work presented in paper III in order to obtain higher transformation efficiency.

In paper II we demonstrated for the first time that a member of the species *S. thermophilus* can be made competent for natural genetic transformation. This is remarkable considering that the LMG 18311 strain is a dairy starter derived from an ancestral strain that probably was domesticated several thousand years ago, and that adaptation to a life in milk has resulted in loss of functionality of about 10 % of the *S. thermophilus* genes (Bolotin *et al.*, 2004). The fact that the late competence genes involved in binding, uptake and recombination of DNA are still intact, strongly indicates that these genes perform an important function in the LMG 18311 starter strain. One possibility is that DNA released from bacteria that have died from natural causes during the fermentation process is taken up by competent sister cells as a source of nucleotides, nitrogen and/or energy. This DNA-for-food hypothesis is unlikely because only the DNA uptake machinery would have to be conserved to exploit extracellular

DNA as a source of nutrition. The high level of transformants obtained show that the competence proteins involved in protection and recombination of the incoming ssDNA is functional as well. Besides, if DNA was taken up as a source of food it would be expected that the presence of extracellular DNA would induce the competent state in *S. thermophilus.* Another possibility is that natural transformation serves as a mechanism for the acquisition of novel traits. It is for instance known from *in silico* analyses of *S. thermophilus* genomes that a 17 kb region of foreign DNA that is inserted into the *pepD* gene consists of a mosaic of fragments. Most of these fragments seem to have been acquired from *Lactobacillus bulgaricus* and *Lactococcus lactis* by lateral gene transfer. Since *L. bulgaricus* and *L. lactis* are used together with *S. thermophilus* in the production of yoghurt and so called hard-cooked cheeses, these lateral gene transfer events have probably taken place during the manufacturing process. Interestingly, the 17 kb region contains a unique gene, *metC*, which is 95 % identical to the *L. bulgaricus* version of this gene. The presence of *metC* allows *S. thermophilus* to synthesize methionine, a rare amino acid in milk. Even though useful genes appear to have been acquired by *S. thermophilus* through natural transformation after it was adapted to a life in milk, it is unlikely that such rare events will favour those bacteria that are naturally transformable over those that have lost this property over time. After all, there is a cost connected with the maintenance of a functional transformation machinery. The third and clearly most plausible possibility is that the transformation machinery has been conserved to enable *S. thermophilus* to cope with endogenous DNA damage. Such damage, which is an inadvertent side effect of normal metabolism, can be repaired by uptake and incorporation of homologous DNA released from dead sister cells during yoghurt and cheese fermentation.

Natural genetic transformation in *S. thermophilus* LMD-9

Recently *S. thermophilus* strain LMD-9 was demonstrated to turn on natural competence spontaneously in early exponential phase, when grown in peptide free chemically defined medium (CDM) (Gardan *et al.*, 2009). Curiously, identical conditions did not induce the competent state in *S. thermophilus* strain CNRZ 1066, and only a few transformants were obtained with the LMG 18311 strain. In the same paper it was also demonstrated that the Ami oligopeptide transporter (AmiCDEF) and the oligopeptide-binding protein AmiA3 are essential for positive transcriptional regulation of *comX* and the late competence genes. In

addition to playing a role in nutrition, oligopeptide transport systems are known in some cases to be involved in cell-cell communication and quorum sensing (Kozlowicz *et al.*, 2006; Pottathil & Lazazzera, 2003). In these cases the role of the oligopeptide transport system is to internalize a short signalling peptide, which acts inside the cell by directly or indirectly modulating the activity of a transcriptional regulator. A paper representing a breakthrough in our understanding of competence regulation in *S. thermophilus* was very recently published. In this study Fontaine and co-workers (2010) succeeded in identifying both the signalling peptide and its putative intracellular target, the transcriptional regulator ComR. Furthermore, by comparing the transcriptomes of a LMD-9 Δ*AmiA-F* mutant and a LMD-9 Δ*comX* mutant, they were able to identify the early competence genes. The signalling peptide, which is encoded by the *comS* gene, is synthesized as a precursor peptide consisting of 24 amino acids (MKTLKIFVLFSLLIAILPYFAGCL). The processing site and the protease involved in the production of mature signalling peptide remain unknown, but experiments with synthetic peptides have shown that only the C-terminal 7 amino acids (PYFAGCL) are required for full biological activity. ComS might be secreted by the Sec pathway, but it cannot be ruled out that a specific transporter is involved. Presumably, signalling peptides that have been imported by the Ami system interact directly with ComR, and provoke a conformational change that increases the regulator's affinity for an inverted repeat motif found to be present in the promoter regions of the *comR*, *comS* and *comX* genes. By stimulating transcription from its own promoter, and from the promoter of *comS*, an autocatalytic chain reaction takes place that amplifies the signal and coordinates competence development throughout the bacterial population. As mentioned above, the putative binding site of ComR is also present in the promoter region of the *comX* gene. It is therefore likely that activated ComR directly stimulates expression of *comX*, whose product then drives expression of the proteins involved in binding, processing, uptake and recombination of DNA.

Gardan *et al.* (2009) found that spontaneous competence development in CDM was very inefficient for *S. thermophilus* strain LMG 18311, and that strain CNRZ 1066 remained non-competent under these conditions. Interestingly, Fontain and colleagues later showed that by adding the synthetic pheromone PYFAGCL to cultures of LMG 18311 and CNRZ 1066, competence was induced in both strains with transformation efficiencies equal to that obtained for CDM-grown cultures of LMD-9 (Fontaine *et al.*, 2010). Since extracellularly added pheromone induced competence in LMG 18311 and CNRZ 1066, it must be the

production, export or maturation of ComS that are not functional in these strains. Alternatively, it is possible that while growth in CDM stimulates production of mature pheromone in strain LMD-9, pheromone production might require different conditions in strains LMG 18311 and CNRZ 1066.

In all naturally transformable streptococci studied prior to *S. thermophilus*, the quorum sensing system regulating competence development consists of a peptide pheromone (14-23 amino acids), a dedicated ABC-transporter and a two-component signal transduction pathway (Johnsborg *et al.*, 2006). It is therefore remarkable that the quorum sensing system regulating competence in *S. thermophilus* is completely different in all its components. It is tempting to speculate that this unusual feature represents a special adaptation to the milk habitat. However, as the same quorum sensing system is present in other non-dairy species belonging to the salivarius phylogenetic group (*Streptococcus salivarius* and *Streptococcus vestibularis*) this cannot be the case (Fontaine *et al.*, 2010).

Food-grade site directed mutagenesis in *S. thermophilus* LMG 18311

To date, no genetically modified yoghurt or cheese starter strains have been approved by The European Food Safety Authority (EFSA). Presumably consumers would be less sceptical towards the use of genetically modified organisms (GMOs) in the dairy industry if the strains used contained only minor genetic modifications. In light of this, we decided to make an attempt at developing a new genetic tool that would make it possible to introduce any desired mutation into the genome of *S. thermophilus* without the use of selection markers or other kinds of foreign DNA. In paper II we obtained results indicating that the relatively high efficiency of the transformation procedure developed for *S. thermophilus* LMG 18311 would make it possible to perform site-directed mutagenesis in this strain without having to rely on antibiotic resistance selection for the identification of desired mutants. Instead, we planned to screen for transformants without selection by performing colony hybridization with a digoxigenin-labelled oligonucleotide probe. This strategy would only be feasible if the fraction of mutants relative to wild type cells in the transformation mix turned out to be high enough to allow identification of positive colonies without having to screen a very large number of agar plates. To get a direct estimate of the transformation efficiency we decided to

take advantage of the fact that a simple blue/white screening procedure exists for the identification of Δ*lacZ* mutants. As a proof of principle we aimed at inactivating the *lacZ* gene by making only minor changes to this gene. We therefore planned to change G_{155} into A_{155} and insert an extra nucleotide (G) between A_{158} and A_{159}, alterations that would give rise to a nonsense and a frameshift mutation, respectively. Since *S. thermophilus* cells are linked together in chains, it would be necessary to separate transformants from non-transformants after transformation. A mechanical blender was used for this purpose. After seeding agar plates containing X-gal with appropriately diluted transformation mix, several white colonies could be observed on each plate (Paper III, Figure 2). This result established beyond doubt that the efficiency of our transformation procedure would allow direct identification of mutants by colony-lift hybridization. We therefore repeated the transformation experiment, and screened seeded agar plates without X-gal with a 30 bp DIG-labelled probe designed to specifically recognize Δ*lacZ* mutants. This generally applicable method worked very well, demonstrating that our site-directed mutagenesis system can be used to alter preselected sites in the genome of *S. thermophilus* LMG 18311 without introducing any foreign DNA. As the procedure relies on the pXL helper plasmid, a genetically engineered strain of *S. thermophilus* developed for the dairy industry must be cured before it can be used in fermentations. The helper plasmid is unstable, and it is therefore quickly lost from host cells during growth in the absence of erythromycin. There is a very small chance that the pXL plasmid or parts of it could integrate into the host cell genome by for example illegitimate recombination. As an extra precaution, a GMO strain intended for food production should therefore be examined by for instance Southern hybridization using the plasmid as a probe. Recent advances in sequencing technologies present the opportunity to perform large-scale sequencing rapidly and at a low cost. Consequently, whole genome sequencing might be used to detect unintended changes in a GMO constructed by the method outlined above.

CONCLUDING REMARKS

The aim of the present study was to develop new and more efficient genetic tools suitable for genetic engineering of *S. thermophilus.* This objective was fully realized for strain LMG 18311. Our contribution has opened up a new field of research in *S. thermophilus*, and has inspired other research groups to apply the novel tools and extend our work (Fontaine *et al.*, 2010; Gardan *et al.*, 2009). In addition, our contributions to the molecular toolbox of *S. thermophilus* include an inducible gene expression system that can be used for all purposes that do not require an absolutely tight promoter. These techniques will be of great benefit for researchers investigating the physiology and metabolism of *S. thermophilus*, and will hopefully also be used by the dairy industry to construct improved starter strains in the not too distant future. Consumer acceptance of dairy products made with starter strains classified as GMOs will require that the use of GMO strains significantly improves the product in one way or another. Such improvements might involve health or taste benefits, better keepability, or a substantially reduced price. In addition, the consumer must perceive the product as safe. This will require better communication between the consumers and the industry in a language that is understandable to the public in general.

Even though the signalling peptide and most of the cognate quorum sensing system required for competence induction in *S. thermophilus* have now been identified, the physico-chemical conditions stimulating expression, secretion and sensing of the peptide pheromone have not been sorted out. Future work should therefore focus on identifying the primary master switch, i.e. the environmental cues and corresponding cellular sensor systems that activates pheromone production and drives *S. thermophilus* into the competent state.

REFERENCES

Ahn, S. J., Wen, Z. Z. T. & Burne, R. A. (2006). Multilevel control of competence development and stress tolerance in *Streptococcus mutans* UA159. *Infect Immun* **74**, 1631-1642.

Alne, I. (2001). Studies of competence genes in *S. thermophilus*. In *Master thesis*, pp. 1-104: Agricultural University of Norway, Ås.

Auclair, J. & Accolas, J. P. (1983). Use of thermophilic lactic starters in the dairy-industry. *Antonie Van Leeuwenhoek Journal of Microbiology* **49**, 313-326.

Banas, J. A. (2004). Virulence properties of *Streptococcus mutans*. *Front Biosci* **9**, 1267-1277.

Behare, P. V., Singh, R., Kumar, M., Prajapati, J. B. & Singh, R. P. (2009). Exopolysaccharides of lactic acid bacteria: A review. *Journal of Food Science and Technology-Mysore* **46**, 1-11.

Berge, M., Moscoso, M., Prudhomme, M., Martin, B. & Claverys, J. P. (2002). Uptake of transforming DNA in Gram-positive bacteria: a view from *Streptococcus pneumoniae*. *Mol Microbiol* **45**, 411-421.

Bilwes, A. M., Quezada, C. M., Croal, L. R., Crane, B. R. & Simon, M. I. (2001). Nucleotide binding by the histidine kinase CheA. *Nature Structural Biology* **8**, 353-360.

Blomqvist, T., Steinmoen, H. & Håvarstein, L. S. (2006a). Natural genetic transformation: A novel tool for efficient genetic engineering of the dairy bacterium *Streptococcus thermophilus*. *Appl Environ Microbiol* **72**, 6751-6756.

Blomqvist, T., Steinmoen, H. & Håvarstein, L. S. (2006b). Pheromone-induced expression of recombinant proteins in *Streptococcus thermophilus*. *Arch Microbiol* **186**, 465-473.

Blomqvist, T., Steinmoen, H. & Håvarstein, L. S. (2010). A food-grade site-directed mutagenesis system for *Streptococcus thermophilus* LMG 18311. *Lett Appl Microbiol* **50**, 314-319.

Bolotin, A., Quinquis, B., Renault, P. & other authors (2004). Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nature Biotechnology*.

Braat, H., Rottiers, P., Hommes, D. W. & other authors (2006). A phase I trial with transgenic bacteria expressing interleukin-10 in Crohn's disease. *Clin Gastroenterol Hepatol* **4**, 754-759.

Campbell, E. A., Choi, S. Y. & Masure, H. R. (1998). A competence regulon in *Streptococcus pneumoniae* revealed by genomic analysis. *Mol Microbiol* **27**, 929-939.

Chastanet, A., Prudhomme, M., Claverys, J. P. & Msadek, T. (2001). Regulation of *Streptococcus pneumoniae clp* genes and their role in competence development and stress survival. *J Bacteriol* **183**, 7295-7307.

Cotter, P. D., Hill, C. & Ross, R. P. (2005). Bacteriocins: Developing innate immunity for food. *Nat Rev Microbiol* **3**, 777-788.

Courtin, P., Monnet, V. & Rul, F. (2002). Cell-wall proteinases PrtS and PrtB have a different role in *Streptococcus thermophilus*/*Lactobacillus bulgaricus* mixed cultures in milk. *Microbiology-Sgm* **148**, 3413-3421.

de Arauz, L. J., Jozala, A. F., Mazzola, P. G. & Penna, T. C. V. (2009). Nisin biotechnological production and application: a review. *Trends in Food Science & Technology* **20**, 146-154.

de Ruyter, P. G. G. A., Kuipers, O. P., Beerthuyzen, M. M., van Alen Boerrigter, I. & de Vos, W. M. (1996a). Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*. *J Bacteriol* **178**, 3434-3439.

de Ruyter, P. G. G. A., Kuipers, O. P. & de Vos, W. M. (1996b). Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. *Appl Environ Microb* **62**, 3662-3667.

de Ruyter, P. G. G. A., Kuipers, O. P., Meijer, W. C. & deVos, W. M. (1997). Food-grade controlled lysis of *Lactococcus lactis* for accelerated cheese ripening. *Nature Biotechnology* **15**, 976-979.

de Saizieu, A., Gardes, C., Flint, N., Wagner, C., Kamber, M., Mitchell, T. J., Keck, W., Amrein, K. E. & Lange, R. (2000). Microarray-based identification of a novel *Streptococcus pneumoniae* regulon controlled by an autoinduced peptide. *J Bacteriol* **182**, 4696-4703.

de Vin, F., Rådström, P., Herman, L. & De Vuyst, L. (2005). Molecular and biochemical analysis of the galactose phenotype of dairy *Streptococcus thermophilus* strains reveals four different fermentation profiles. *Appl Environ Microb* **71**, 3659-3667.

de Vos, W. M. (1999). Safe and sustainable systems for food-grade fermentations by genetically modified lactic acid bacteria. *Int Dairy J* **9**, 3-10.

Delorme, C., Bartholini, C., Bolotine, A., Ehrlich, S. D. & Renault, P. (2010). Emergence of a cell wall protease in the *Streptococcus thermophilus* population. *Appl Environ Microb* **76**, 451-460.

Diep, D. B., Håvarstein, L. S. & Nes, I. F. (1995). A bacteriocin-like peptide induces bacteriocin synthesis in *Lactobacillus plantarum* C11. *Mol Microbiol* **18**, 631-639.

Diep, D. B., Mathiesen, G., Eijsink, V. G. H. & Nes, I. F. (2009). Use of Lactobacilli and their pheromone-based regulatory mechanism in gene expression and drug delivery. *Current Pharmaceutical Biotechnology* **10**, 62-73.

DIRECTIVE 2009/41/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 6 May 2009 on the contained use of genetically modified micro-organisms
(http://www.bmwf.gv.at/fileadmin/user_upload/forschung/gentechnik/2009-41-EC.pdf)

Draskovic, I. & Dubnau, D. (2005). Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulphide bonds. *Mol Microbiol* **55**, 881-896.

Eichenbaum, Z., Federle, M. J., Marra, D., de Vos, W. M., Kuipers, O. P., Kleerebezem, M. & Scott, J. R. (1998). Use of the lactococcal *nisA* promoter to regulate gene expression in gram-positive bacteria: Comparison of induction level and promoter strength. *Appl Environ Microb* **64**, 2763-2769.

El Demerdash, H. A., Heller, K. J. & Geis, A. (2003). Application of the *shsp* gene, encoding a small heat shock protein, as a food-grade selection marker for lactic acid bacteria. *Appl Environ Microbiol* **69**, 4408-4412.

Ellison, D. W. & McCleary, W. R. (2000). The unphosphorylated receiver domain of PhoB silences the activity of its output domain. *J Bacteriol* **182**, 6592-6597.

Facklam, R. (2002). What happened to the streptococci: Overview of taxonomic and nomenclature changes. *Clin Microbiol Rev* **15**, 613-630.

Fernandez-Espla, M. D., Garault, P., Monnet, V. & Rul, F. (2000). *Streptococcus thermophilus* cell wall-anchored proteinase: release, purification, and biochemical and genetic characterization. *Appl Environ Microbiol* **66**, 4772-4778.

Fontaine, L., Boutry, C., Guedon, E., Guillot, A., Ibrahim, M., Grossiord, B. & Hols, P. (2007). Quorum-sensing regulation of the production of Blp bacteriocins in *Streptococcus thermophilus*. *J Bacteriol* **189**, 7195-7205.

Fontaine, L., Boutry, C., de Frahan, M. H., Delplace, B., Fremaux, C., Horvath, P., Boyaval, P. & Hols, P. (2010). A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J Bacteriol* **192**, 1444-1454.

Galperin, M. Y. (2006). Structural classification of bacterial response regulators: Diversity of output domains and domain combinations. *J Bacteriol* **188**, 4169-4182.

Gao, R., Mack, T. R. & Stock, A. M. (2007). Bacterial response regulators: versatile regulatory strategies from common domains. *Trends in Biochemical Sciences* **32**, 225-234. Copyright (2010).

Gao, R. & Stock, A. M. (2009). Biological insights from structures of two-component proteins. *Annual Review of Microbiology* **63**, 133-154.

Garault, P., Letort, C., Juillard, V. & Monnet, V. (2000). Branched-chain amino acid biosynthesis is essential for optimal growth of *Streptococcus thermophilus* in milk. *Appl Environ Microb* **66**, 5128-5133.

Garault, P., Le Bars, D., Besset, C. & Monnet, V. (2002). Three oligopeptide-binding proteins are involved in the oligopeptide transport of *Streptococcus thermophilus*. *J Biol Chem* **277**, 32-39.

Garcia-Suarez, M. D. M., Vazquez, F. & Mendez, F. J. (2006). *Streptococcus pneumoniae* virulence factors and their clinical impact: an update. *Enfermedades Infecciosas Y Microbiologia Clinica* **24**, 512-517.

Gardan, R., Besset, C., Guillot, A., Gitton, C. & Monnet, V. (2009). The oligopeptide transport system is essential for the development of natural competence in *Streptococcus thermophilus* strain LMD-9. *J Bacteriol* **191**, 4647-4655.

Geng, H., Nakano, S. & Nakano, M. M. (2004). Transcriptional activation by *Bacillus subtilis* ResD: Tandem binding to target elements and phosphorylation-dependent and -independent transcriptional activation. *J Bacteriol* **186**, 2028-2037.

Grebe, T. W. & Stock, J. B. (1999). The histidine protein kinase superfamily. *Adv Microb Physiol* **41**, 139-227.

Griffith, F. (1928). The significance of pneumococcal types. *Journal of Hygiene* **27**, 113-159.

Gunnewijk, M. G., Postma, P. W. & Poolman, B. (1999). Phosphorylation and functional properties of the IIA domain of the lactose transport protein of *Streptococcus thermophilus*. *J Bacteriol* **181**, 632-641.

Herbert, M. A., Beveridge, C. J. E. & Saunders, N. J. (2004). Bacterial virulence factors in neonatal sepsis: group B streptococcus. *Current Opinion in Infectious Diseases* **17**, 225-229.

Hols, P., Hancy, F., Fontaine, L. & other authors (2005). New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed by comparative genomics. *Fems Microbiol Rev* **29**, 435-463.

Horvath, P. & Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167-170.

Hui, F. M. & Morrison, D. A. (1991). Genetic transformation in *Streptococcus pneumoniae*: nucleotide sequence analysis shows *comA*, a gene required for competence induction, to be a member of the bacterial ATP-dependent transport proteinfFamily. *J Bacteriol* **173**, 372-381.

Hui, F. M., Zhou, L. X. & Morrison, D. A. (1995). Competence for genetic transformation in S*treptococcus pneumoniae*: organization of a regulatory locus with homology to two lactococcin A secretion genes. *Gene* **153**, 25-31.

Håvarstein, L. S., Holo, H. & Nes, I. F. (1994). The leader peptide of colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by Gram-positive bacteria. *Microbiology* **140 (Pt 9)**, 2383-2389.

Håvarstein, L. S., Coomaraswamy, G. & Morrison, D. A. (1995a). An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae*. *Proc Natl Acad Sci U S A* **92**, 11140-11144.

Håvarstein, L. S., Diep, D. B. & Nes, I. F. (1995b). A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. *Mol Microbiol* **16**, 229-240.

Håvarstein, L. S., Gaustad, P., Nes, I. F. & Morrison, D. A. (1996). Identification of the streptococcal competence-pheromone receptor. *Mol Microbiol* **21**, 863-869.

Håvarstein, L. S. (2003). Intercellular communication in Gram-positive bacteria depends on peptide pheromones and their histidine kinase receptors. *In: Inouye M, and Dutta R (eds) Histidine Kinases in Signal Transduction New York, Academic Press, pp 341-363*.

Ishii, S., Yano, T. & Hayashi, H. (2006). Expression and characterization of the peptidase domain of *Streptococcus pneumoniae* ComA, a bifunctional ATP-binding cassette transporter involved in quorum sensing pathway. *J Biol Chem* **281**, 4726-4731.

Johnsborg, O., Blomqvist, T., Kilian, M. & Håvarstein, L. S. (2006). Biologically active peptides in streptococci. In *The Molecular Biology of Streptococci*. Edited by R. a. C. Eds. Hakenbeck, S.,: Horizon Scientific Press.

Johnsborg, O., Eldholm, V. & Håvarstein, L. S. (2007). Natural genetic transformation: prevalence, mechanisms and function. *Res Microbiol* **158**, 767-778.

Johnsborg, O. & Håvarstein, L. S. (2009). Regulation of natural genetic transformation and acquisition of transforming DNA in *Streptococcus pneumoniae*. *Fems Microbiol Rev* **33**, 627-642.

Kilian, M. (1998). *Streptococcus* and *Lactobasillus*. In *Topley and Wilson's Microbiology and microbial infections 9th edition*, pp. 633-667. Edited by A. B. a. B. I. Duerden. London: Arnold.

Kleerebezem, M., Beerthuyzen, M. M., Vaughan, E. E., de Vos, W. M. & Kuipers, O. P. (1997). Controlled gene expression systems for lactic acid bacteria: Transferable nisin-inducible expression cassettes for *Lactococcus*, *Leuconostoc*, and *Lactobacillu*s spp. *Appl Environ Microb* **63**, 4581-4584.

Knutsen, E., Ween, O. & Håvarstein, L. S. (2004). Two separate quorum-sensing systems upregulate transcription of the same ABC transporter in *Streptococcus pneumoniae*. *J Bacteriol* **186**, 3078-3085.

Kotake, Y., Ishii, S., Yano, T., Katsuoka, Y. & Hayashi, H. (2008). Substrate recognition mechanism of the peptidase domain of the quorum-sensing-signal-producing ABC transporter ComA from Streptococcus. *Biochemistry* **47**, 2531-2538.

Kozlowicz, B. K., Dworkin, M. & Dunny, G. M. (2006). Pheromone-inducible conjugation in *Enterococcus faecalis*: A model for the evolution of biological complexity? *International Journal of Medical Microbiology* **296**, 141-147.

Kreth, J., Merritt, J., Shi, W. & Qi, F. (2005). Co-ordinated bacteriocin production and competence development: a possible mechanism for taking up DNA from neighbouring species. *Mol Microbiol* **57**, 392-404.

Kreth, J., Merritt, J., Zhu, L., Shi, W. Y. & Qi, F. X. (2006). Cell density- and ComE-dependent expression of a group of mutacin and mutacin-like genes in *Streptococcus mutans*. *Fems Microbiol Lett* **265**, 11-17.

Kuipers, O. P., Beerthuyzen, M. M., De ruyter, P. G. G. A., Luesink, E. J. & De vos, W. M. (1995). Autoregulation of nisin biosynthesis in *Lactococcus lactis* by signal-transduction. *J Biol Chem* **270**, 27299-27304.

Kuipers, O. P., de Ruyter, P. G. G. A., Kleerebezem, M. & de Vos, W. M. (1998). Quorum sensing-controlled gene expression in lactic acid bacteria. *J Biotechnol* **64**, 15-21. Copyright (2010).

Köhler, W. (2007). The present state of species within the genera *Streptococcus* and *Enterococcus*. *Int J Med Microbiol* **297**, 133-150.

Laub, M. T. & Goulian, M. (2007). Specificity in two-component signal transduction pathways. *Annu Rev Genet* **41**, 121-145.

Lee, M. S. & Morrison, D. A. (1999). Identification of a new regulator in *Streptococcus pneumoniae* linking quorum sensing to competence for genetic transformation. *J Bacteriol* **181**, 5004-5016.

Letort, C. & Juillard, V. (2001). Development of a minimal chemically-defined medium for the exponential growth of *Streptococcus thermophilus*. *J Appl Microbiol* **91**, 1023-1029.

Li, Y. H., Hanna, M. N., Svensater, G., Ellen, R. P. & Cvitkovitch, D. G. (2001a). Cell density modulates acid adaptation in *Streptococcus mutans*: Implications for survival in biofilms. *J Bacteriol* **183**, 6875-6884.

Li, Y. H., Lau, P. C. Y., Lee, J. H., Ellen, R. P. & Cvitkovitch, D. G. (2001b). Natural genetic transformation of *Streptococcus mutans* growing in biofilms. *J Bacteriol* **183**, 897-908.

Li, Y. H., Tang, N., Aspiras, M. B., Lau, P. C. Y., Lee, J. H., Ellen, R. P. & Cvitkovitch, D. G. (2002). A quorum-sensing signaling system essential for genetic competence in *Streptococcus mutans* is involved in biofilm formation. *J Bacteriol* **184**, 2699-2708.

Makarova, K., Slesarev, A., Wolf, Y. & other authors (2006). Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A* **103**, 15611-15616.

Marina, A., Mott, C., Auyzenberg, A., Hendrickson, W. A. & Waldburger, C. D. (2001). Structural and mutational analysis of the PhoQ histidine kinase catalytic domain. Insight into the reaction mechanism. *J Biol Chem* **276**, 41182-41190.

Martin, B., Quentin, Y., Fichant, G. & Claverys, J. P. (2006). Independent evolution of competence regulatory cascades in streptococci? *Trends Microbiol* **14**, 339-345.

Mascher, T., Helmann, J. D. & Unden, G. (2006). Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol Mol Biol Rev* **70**, 910-938.

Mejean, V. & Claverys, J. P. (1988). Polarity of DNA entry in transformation of *Streptococcus pneumoniae*. *Molecular & General Genetics* **213**, 444-448.

Mierau, I. & Kleerebezem, M. (2005). 10 years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*. *Appl Microbiol Biotechnol*, 1-13.

Mollet, B., Knol, J., Poolman, B., Marciset, O. & Delley, M. (1993). Directed genomic integration, gene replacement, and integrative gene expression in *Streptococcus thermophilus*. *J Bacteriol* **175**, 4315-4324.

Morrison, D. A. (1997). Streptococcal competence for genetic transformation: Regulation by peptide pheromones. *Microbial Drug Resistance-Mechanisms Epidemiology and Disease* **3**, 27-37.

Mulholland, K. (2007). Childhood pneumonia mortality--a permanent global emergency. *Lancet* **370**, 285-289.

Nikolskaya, A. N. & Galperin, M. Y. (2002). A novel type of conserved DNA-binding domain in the transcriptional regulators of the AlgR/AgrA/LytR family. *Nucleic Acids Res* **30**, 2453-2459.

Novick, R. P. & Geisinger, E. (2008). Quorum sensing in staphylococci. *Annu Rev Genet* **42**, 541-564.

O'Sullivan, D. J. & Klaenhammer, T. R. (1993). High-copy-number and low-copy-number *Lactococcus* shuttle cloning vectors with features for clone screening. *Gene* **137**, 227-231.

Pastink, M. I., Sieuwerts, S., de Bok, F. A. M., Janssen, P. W. M., Teusink, B., Vlieg, J. E. T. V. & Hugenholtz, J. (2008). Genomics and high-throughput screening approaches for optimal flavour production in dairy fermentation. *Int Dairy J* **18**, 781-789.

Pastink, M. I., Teusink, B., Hols, P., Visser, S., de Vos, W. M. & Hugenholtz, J. (2009). Genome-scale model of *Streptococcus thermophilus* LMG18311 for metabolic comparison of lactic acid bacteria. *Appl Environ Microbiol* **75**, 3627-3633.

Perry, J. A., Jones, M. B., Peterson, S. N., Cvitkovitch, D. G. & Levesque, C. M. (2009). Peptide alarmone signalling triggers an auto-active bacteriocin necessary for genetic competence. *Mol Microbiol* **72**, 905-917.

Pestova, E. V., Håvarstein, L. S. & Morrison, D. A. (1996). Regulation of competence for genetic transformation in *Streptococcus pneumoniae* by an auto-induced peptide pheromone and a two-component regulatory system. *Mol Microbiol* **21**, 853-862.

Peterson, S. N., Sung, C. K., Cline, R. & other authors (2004). Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. *Mol Microbiol* **51**, 1051-1070.

Pfeiler, E. A. & Klaenhammer, T. R. (2007). The genomics of lactic acid bacteria. *Trends Microbiol* **15**, 546-553.

Pottathil, M. & Lazazzera, B. A. (2003). The extracellular Phr peptide-rap phosphatase signaling circuit of *Bacillus subtilis*. *Front Biosci* **8**, D32-D45.

Prudhomme, M., Attaiech, L., Sanchez, G., Martin, B. & Claverys, J. P. (2006). Antibiotic stress induces genetic transformability in the human pathogen *Streptococcus pneumoniae*. *Science* **313**, 89-92.

Purohit, D. H., Hassan, A. N., Bhatia, E., Zhang, X. & Dwivedi, C. (2009). Rheological, sensorial, and chemopreventive properties of milk fermented with exopolysaccharide-producing lactic cultures. *J Dairy Sci* **92**, 847-856.

Rasmussen, T. B., Danielsen, M., Valina, O., Garrigues, C., Johansen, E. & Pedersen, M. B. (2008). *Streptococcus thermophilus* core genome: Comparative genome hybridization study of 47 strains. *Appl Environ Microb* **74**, 4703-4710.

Reichmann, P. & Hakenbeck, R. (2000). Allelic variation in a peptide-inducible two-component system of *Streptococcus pneumoniae*. *Fems Microbiol Lett* **190**, 231-236.

Risøen, P. A., Håvarstein, L. S., Diep, D. B. & Nes, I. F. (1998). Identification of the DNA-binding sites for two response regulators involved in control of bacteriocin synthesis in *Lactobacillus plantarum* C11. *Mol Gen Genet* **259**, 224-232.

Risøen, P. A., Johnsborg, O., Diep, D. B., Hamoen, L., Venema, G. & Nes, I. F. (2001). Regulation of bacteriocin production in *Lactobacillus plantarum* depends on a conserved promoter arrangement with consensus binding sequence. *Mol Genet Genomics* **265**, 198-206.

Robertson, G. T., Ng, W. L., Foley, J., Gilmour, R. & Winkler, M. E. (2002). Global transcriptional analysis of *clpP* mutations of type 2 *Streptococcus pneumoniae* and their effects on physiology and virulence. *J Bacteriol* **184**, 3508-3520.

Robinson, V. L., Buckler, D. R. & Stock, A. M. (2000). A tale of two components: a novel kinase and a regulatory switch. *Nature Structural Biology* **7**, 626-633.

Robitaille, G., Moineau, S., St-Gelais, D., Vadeboncoeur, C. & Britten, M. (2007). Galactose metabolism and capsule formation in a recombinant strain of *Streptococcus thermophilus* with a galactose-fermenting phenotype. *J Dairy Sci* **90**, 4051-4057.

Sanders, M. E. (1991). Genetic approaches for the improvement of strains for Italian cheese manufacture - a review. *J Dairy Sci* **74**, 3647-3655.

Shareck, J., Choi, Y., Lee, B. & Miguez, C. B. (2004). Cloning vectors based on cryptic plasmids isolated from lactic acid bacteria: their characteristics and potential applications in biotechnology. *Crit Rev Biotechnol* **24**, 155-208.

Sherman, J. M. (1937). The Streptococci. *Bacteriol Rev* **1**, 3-97.

Sidote, D. J., Barbieri, C. M., Wu, T. & Stock, A. M. (2008). Structure of the *Staphylococcus aureus* AgrA LytTR domain bound to DNA reveals a beta fold with an unusual mode of binding. *Structure* **16**, 727-735.

Smit, G., Smit, B. A. & Engels, W. J. M. (2005). Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. *Fems Microbiol Rev* **29**, 591-610.

Sorek, R., Kunin, V. & Hugenholtz, P. (2008). CRISPR - a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol* **6**, 181-186.

Steidler, L., Neirynck, S., Huyghebaert, N., Snoeck, V., Vermeire, A., Goddeeris, B., Cox, E., Remon, J. P. & Remaut, E. (2003). Biological containment of genetically modified *Lactococcus lactis* for intestinal delivery of human interleukin 10. *Nature Biotechnology* **21**, 785-789.

Sung, C. K. & Morrison, D. A. (2005). Two distinct functions of ComW in stabilization and activation of the alternative sigma factor ComX in *Streptococcus pneumoniae*. *J Bacteriol* **187**, 3052-3061.

Thomas, T. D. & Crow, V. L. (1984). Selection of galactose-fermenting *Streptococcus thermophilus* in lactose limited chemostat cultures. *Appl Environ Microb* **48**, 186-191.

Vaillancourt, K., LeMay, J. D., Lamoureux, M., Frenette, M., Moineau, S. & Vadeboncoeur, C. (2004). Characterization of a galactokinase-positive recombinant strain of *Streptococcus thermophilus*. *Appl Environ Microb* **70**, 4596-4603.

van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P. & de Vos, W. M. (2000). Control of lactose transport, beta-galactosidase activity, and glycolysis by CcpA in *Streptococcus thermophilus*: Evidence for carbon catabolite repression by a non-phosphoenolpyruvate-dependent phosphotransferase system sugar. *J Bacteriol* **182**, 5982-5989.

van den Bogaard, P. T. C., Hols, P., Kuipers, O. P., Kleerebezem, M. & de Vos, W. M. (2004). Sugar utilisation and conservation of the *gal-lac* gene cluster in *Streptococcus thermophilus*. *Syst Appl Microbiol* **27**, 10-17.

van der Ploeg, J. R. (2005). Regulation of bacteriocin production in *Streptococcus mutans* by the quorum-sensing system required for development of genetic competence. *J Bacteriol* **187**, 3980-3989.

Vaughan, E. E., van den Bogaard, P. T. C., Catzeddu, P., Kuipers, O. P. & de Vos, W. M. (2001). Activation of silent *gal* genes in the *lac-gal* regulon of *Streptococcus thermophilus*. *J Bacteriol* **183**, 1184-1194.

Waters, C. M. & Bassler, B. L. (2005). Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol* **21**, 319-346.

Waters, L. S. & Storz, G. (2009). Regulatory RNAs in Bacteria. *Cell* **136**, 615-628.

Ween, O., Gaustad, P. & Håvarstein, L. S. (1999). Identification of DNA binding sites for ComE, a key regulator of natural competence in *Streptococcus pneumoniae*. *Mol Microbiol* **33**, 817-827.

Wells, J. M. & Mercenier, A. (2008). Mucosal delivery of therapeutic and prophylactic molecules using lactic acid bacteria. *Nat Rev Microbiol* **6**, 349-362.

Wolfe, A. J. (2010). Physiologically relevant small phosphodonors link metabolism to signal transduction. *Curr Opin Microbiol* **13**, 204-209.

Zhang, K., Ou, M. Z., Wang, W. L. & Ling, J. Q. (2009). Effects of quorum sensing on cell viability in *Streptococcus mutans* biofilm formation. *Biochemical and Biophysical Research Communications* **379**, 933-938.

Paper I

Arch Microbiol (2006) 186:465–473
DOI 10.1007/s00203-006-0162-0

ORIGINAL PAPER

Pheromone-induced expression of recombinant proteins in *Streptococcus thermophilus*

Trinelise Blomqvist · Hilde Steinmoen · Leiv Sigve Håvarstein

Received: 30 January 2006 / Revised: 16 May 2006 / Accepted: 28 July 2006 / Published online: 24 August 2006

Abstract A locus encoding proteins with high homology to the pneumococcal BlpABCHR quorum-sensing system was identified in *Streptococcus thermophilus* LMG 18311. The BlpABCHR system regulates bacteriocin production in *Streptococcus pneumoniae* by monitoring the extracellular concentration of a peptide-pheromone encoded by *blpC*. The homologous system in *S. thermophilus*, termed StbABCHR, contains a corresponding gene (*stbC*) encoding a possible peptide-pheromone (STP) that presumably controls bacteriocin production in *S. thermophilus*. We synthesized this peptide and found that it activates transcription of a *gusA* reporter gene placed behind the promoter of the bacteriocin-like gene *stbD*. Furthermore, deletion mapping and mutational analysis of the *stbD* promoter region were used to identify a degenerated direct repeat motif required for STP induced GusA expression. Our findings provide strong evidence that STP regulates bacteriocin production in *S. thermophilus* LMG 18311, and show that the StbABCHR quorum-sensing system can be exploited for inducible expression of recombinant proteins in this bacterial species.

Keywords *Streptococcus thermophilus* · Inducible expression of proteins · Peptide pheromone · Bacteriocin

T. Blomqvist · H. Steinmoen · L. S. Håvarstein (✉)
Department of Chemistry,
Biotechnology and Food Science,
Norwegian University of Life Sciences,
P.O.Box 5003, 1432 Ås, Norway
e-mail: sigve.havarstein@umb.no

Introduction

Streptococcus thermophilus is used as a starter in the production of yoghurt and certain cheeses. Even though this makes it one of the economically most important bacteria in the dairy industry, the genetic tools available for researchers studying *S. thermophilus* is limited compared to those developed for other widely used lactic acid bacteria (LAB). In particular, there is a need for better tools that will make it possible to express foreign or endogenous proteins in a controlled manner. We therefore set out to develop an inducible system for expression of recombinant proteins in *S. thermophilus*. Since this bacterium is used in the fermentation of various foods, we wanted to use genetic elements originating from the bacterium itself in order to introduce as little foreign DNA as possible.

Cell–cell communication by secreted peptide-pheromones is a common phenomenon in Gram-positive bacteria (Eijsink et al. 2002; Kleerebezem 2004; Waters and Bassler 2005). A subgroup of these pheromones consists of unmodified peptides (Håvarstein 2003). In the genus *Streptococcus* pheromones belonging to this subgroup regulate production of bacteriocins, competence for natural genetic transformation, virulence and biofilm formation (Diep et al. 1995; Håvarstein et al. 1995a; Hava and Camilli 2002; Li et al. 2002; Hidalgo-Grass et al. 2004). They are ribosomally synthesized as precursor peptides containing N-terminal double-glycine leader sequences that are removed concomitant with transport over the cytoplasmic membrane (Håvarstein et al. 1995b). Their dedicated secretion apparatuses consist of an ABC-transporter and an accessory protein. Once outside the bacterial cell the mature peptide-pheromone

is sensed by a specific histidine kinase residing in the cytoplasmic membrane of the producer cell (Håvarstein et al. 1996; Pestova et al. 1996). Binding of the peptide ligand presumably results in autophosphorylation and subsequent transfer of the phosphoryl group from the histidine kinase to its cognate response regulator. In the phosphorylated state the response regulator will initiate transcription of target genes, i.e. genes with promoters containing the appropriate binding site (Risøen et al. 1998; Ween et al. 1999). As mentioned above, production of class II bacteriocins in streptococci and other Gram-positive bacteria is often regulated by this kind of quorum-sensing system (Saucier et al. 1997; de Saizieu et al. 2000; Eijsink et al. 2002; van der Ploeg 2005; Mierau and Kleerebezem 2005). Production of bacteriocin is turned on when the extracellular concentration of a particular peptide-pheromone reaches a critical level, reflecting that the bacterial population responding to that peptide-pheromone is dense enough to attack competing bacteria sharing the same niche (Eijsink et al. 2002). Another way to turn on bacteriocin production is to add natural or synthetic peptide-pheromone to the bacterial culture. Consequently, if a recombinant gene is placed behind a bacteriocin promoter, it is possible to induce expression of this gene by adding the appropriate pheromone. This principle has already been exploited to construct efficient systems for regulated gene expression in other LAB (Axelsson et al. 2003; Mathiesen et al. 2004; Sørvig et al. 2005; Mierau and Kleerebezem 2005). In the present work we took advantage of a quorum-sensing system identified in the genome of *S. thermophilus* (Hols et al. 2005) to develop a peptide-pheromone based system specifically designed for inducible expression of recombinant proteins in this important dairy starter.

Materials and methods

Bacterial strains and growth conditions

Streptococcus thermophilus strain LMG 18311 (ATCC no.BAA-250) (Bolotin et al. 2004) was grown in Hogg-Jago glucose broth (HJG) (3% tryptone; 1% yeast extract; 0.2% beef extract; 0.5% KH_2PO_4; 0.5% glucose). Electroporated cells were grown in HJG supplemented with 0.5% lactose (HJGL), and 0.4 M D-sorbitol (HJGLS). *Escherichia coli* were grown in Luria-Bertani (LB) broth (1% tryptone; 0.5% yeast extract; 1% $NaCl_2$). Agar plates were prepared by adding 1.5% (w/v) agar to the media. All incubations were carried out at 37°C, and all optical density measurements of bacterial cultures were performed spectrophotometrically at 660 nm.

Electroporation

Preparation of electrocompetent *S. thermophilus* LMG 18311 cells were performed essentially as described by Buckley et al. (1999). Competent 18311 cells were thawed on ice, and 1 μg plasmid was added to 80 μl cells. After 30-min incubation on ice, the cells were transferred to a 1 mm electroporation-cuvette, and a single electric pulse was delivered (settings at 1.6 kV and 2.5 ms) using a Bio-Rad MicroPulser unit. The electroporated cells were immediately resuspended in 1 ml ice-cold HJGLS and incubated for 3 h at 37°C, before spreading on HJGL plates containing 2 μg/ml erythromycin. Transformants were picked after 24–48 h incubation at 37°C, and verified by PCR using the primers M13F and M13R (Table 1) complementary to sequences flanking the cloning site in the pTRKH2 vector (O'Sullivan and Klaenhammer 1993).

Construction of the inducible expression vector pTB18

A ~250 bp PCR fragment, corresponding to the promoter region of the *stbD* gene (see Figs. 1, 2a), was amplified using the primers Prim 1 and Prim 2 (Table 1) and genomic DNA from *S. thermophilus* strain LMG 18311 as template. Thermal cycle conditions were as follows: 2 min at 97°C followed by 30 cycles of 1 min at 94°C, 30 s at 60°C (annealing) and 1 min at 72°C (extension). The PCR reaction was analyzed by agarose gel electrophoresis, and a fragment of the expected size was detected. The PCR fragment was ligated into the pDRIVE-vector as described by the manufacturer (Qiagen), and transformed into chemically competent *E. coli* TOP10 cells (Invitrogen). Transformants were selected on LB plates containing 50 μg/ml of kanamycin. After purification of the recombinant plasmid, the cloned PCR fragment was excised from the pDRIVE vector using the restriction enzymes *Pst*I and *Sac*I (Fermentas). The promoter-fragment was then purified by agarose gel electrophoresis followed by gel extraction (Jet Quick, Saveen), ligated in to the plasmid pTRKH2 (O'Sullivan and Klaenhammer 1993) precleaved with *Pst*I and *Sac*I, and transformed into chemically competent *E. coli* TOP10 cells (Invitrogen). Finally, transformants containing pTRKH2 plasmid harboring the *stbD* promoter-fragment were selected on LB plates containing 500 μg/ml erythromycin. The resulting expression vector was termed pTB18.

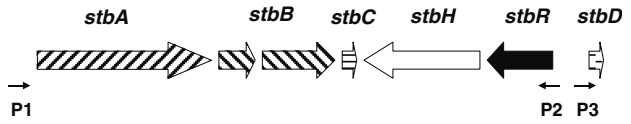

Fig. 1 The genomic organization of the *stb* locus in *Streptococcus thermophilus* LMG 18311. StbA: ABC-transporter. StbB: the accessory protein of StbA. A frameshift mutation in *stbB* probably makes the StbAB transporter non-functional. StbC: the precursor of the STP peptide pheromone. StbH: histidine kinase assumed to be the receptor of STP. StbR: the cognate response regulator of StbH. StbD: a bacteriocin-like peptide. P1-3 indicates the promoter regions of the three different transcription units. The *stb* locus is situated between nucleotide positions 1490249 and 1496546 in the genome of *Streptococcus thermophilus* LMG 18311

A

Mut 1

CTTCAAGGTCTAGTCCTCTCTTTTATGACGAATACTGTTTATTGAAAAATTTG

TAACATAAAGAAAACGGTTTTTCATTTTTTTATGAGTATAAAATGAGATTTTT

Mut 2 Left D.R.

TTCTGAATTTTAGAAATAAATATACATTAGGAATTACCATTCGGGACATATAG

Right D.R. Mut 3 P.B.

CCACTTTTTGGGTACGCTAGCTCTGATAGAGACAATTGAA**TGCTATACT**AAAG

ATGTGATTGAGAGATCACACGATAAAAATTTTAGGAGGTAGTTGCC*ATG - gusA*

B

TB18G	ACCATTCGG - 11 bp - ACTTTTTGG - 28 bp - TGCTATACT
Mut 11	ACCATTCGG - 11 bp - AC**AA**TT**C**GG - 28 bp - TGCTATACT
Mut 11c	ACCATTCGG - 11 bp - AC**AA**TT**C**GG - **29** bp - TGCTATACT
Mut 12	ACCATTCGG - **12** bp - **T**AC**AA**TT**C**GG - 28 bp - TGCTATACT
Mut V	AC**TT**TT**T**GG - 11 bp - ACTTTTTGG - 28 bp - TGCTATACT
Mut TG	ACCATTCGG - 11 bp - ACTTTTTGG - 28 bp - **AA**CTATACT

Fig. 2 **a** The sequence of the *stbD*-promoter (P3). The Pribnow-box (TATACT) and extended Pribnow-box (TGNTATACT) are indicated in *bold*. The direct repeat motifs (D.R.) are *underlined*, and the start codon of the *gusA* reporter gene is indicated in *italics*. The endpoints of the deletions giving rise to Mut 1, Mut 2 and Mut 3 are indicated by *vertical bars*. **b** Sequences of direct repeat motifs and extended Pribnow-box of various *stbD* promoter mutants. Introduced base substitutions are indicated in *bold*. TB18G represents the wild-type sequence

Cloning of the *gusA* reporter gene in pTB18

The *gusA* reporter gene (A-S69414) was excised from the pSIP403 vector (Sørvig et al. 2003) with the restriction enzymes *Nco*I and *Xho*I (Promega), and purified by agarose gel extraction (Jet Quick, Saveen). As the start codon of the *stbD* gene is naturally part of a *Nco*I site (see Prim 1, Table 1), the purified *gusA* fragment was ligated directly into the pTB18 expression vector precleaved with *Nco*I and *Xho*I. The resulting construct, pTB18G, was subsequently transformed into chemically competent *E. coli* TOP10 cells (Invitrogen), and positive clones were selected on LB plates containing 500 μg/ml of erythromycin. The presence of the *gusA* gene, correctly inserted behind the *stbD* promoter, was verified by DNA sequencing, using a combination of the primers M13F, M13R, Gus5F, Gus1R, Gus2R, and Gus3R (Table 1). To ensure that the cloned promoter/*gusA* fragment was inserted in the opposite direction of the *lacZ* promoter, restriction mapping of pTB18G with the restriction enzymes *Nco*I, *Sma*I and *Eco*RV was carried out. Finally, the pTB18G construct was transformed into *S. thermophilus* strain LMG 18311 by electroporation as described above, resulting in the TB18G strain.

The Arg-tRNA transcriptional terminator from *S. pneumoniae* (Håvarstein et al. 1995a) was inserted upstream of the *stbD* promoter to prevent read-through of possible transcripts originating in the pTRKH2 vector. Using pTB18G as template, and the primers TB18G-term and GusR (Table 1), a DNA fragment encompassing the *stbD* promoter and the *gusA* reporter gene was amplified. The TB18G-term primer contains the Arg-tRNA terminator, and its 3′-end is complementary to the vector sequence immediately upstream of the *stbD* promoter sequence. The GusR is complementary to the region immediately downstream of the *gusA* gene. The resulting PCR fragment was first cloned into the pCR 2.1-TOPO vector (Invitrogene), excised by *Bam*HI and *Xho*I, and ligated into the corresponding sites of the pTRKH2 vector. After sequencing the complete insert using the primers described above, the construct was transformed into the *S. thermophilus* strain LMG 18311 by electroporation.

GUS assay

To determine whether transcription from the *stbD*-promoter was induced by synthetic STP added to *S. thermophilus* LMG 18311 cells harboring the pTB18G plasmid, cultures subjected to the peptide-pheromone were assayed for β-glucuronidase activity. Synthetic STP (NH_2-SGWMDYINGFLKGFGGQRTLPTKDY NIPQV-COOH) was supplied by Genosphere Biotechnologies (Paris, France). Uninduced cultures, receiving no peptide-pheromone, were run in parallel as negative controls. An overnight culture of the TB18G strain grown in HJG was diluted to an optical density of 0.05 in preheated HJG (37°C), and incubated until it reached an optical density of 0.15. STP was then added to the final concentration of 250 ng/ml. After 2 h of incubation at 37°C, 1 ml samples of STP-induced and uninduced cultures were withdrawn and centrifuged at 16,000×*g* for 1 min. Following removal

Table 1 Primers used in this study

Primer	Sequence
Prim 1	5′-GTTTGAGTTG CCATGGCAACTACCTCC−3′
Prim 2	5′-ATTAGGATCCTTCAAGGTCTAGTCCTCTCTTTTATGACG-3′
M13F	5′-GTAAAACGACGGCCAG-3′
M13R	5′-CAGGAAACAGCTATGAC-3′
Gus5F	5′-GTCACGCCGTATGTTATTGCC-3′
Gus1R	5′-GGACATACCATCCGTAATAACGG-3′
Gus2R	5′-CTGCCCAGTCGAGCATCTCTT-3′
Gus3R	5′-CCTTGTCCAGTTGCAACCACCT-3′
TB18G-Term	5′-ACTGGGATCCAAAGCCGGGAATTTTCCCGGCTTTTTTTTTAAAAAACCTGCAGAC GCGTTACGTATC-3′
GusR	5′-TACCGAATTCCTCGAGTCTAGATCATTGTT-3′
Mut 1	5′-AAATGGATCCTGTAACATAAAGAAAACGGTTTT-3′
Mut 2	5′-AATAAGGATCCTATACATTAGGAATTACCATTCGGGAC-3′
Mut 3	5′-TTTGGGATCCACGCTAGCTCTGAT-3′
MutST.1	5′-GGAATTACCATTCGGGACATATAGCCACAATTCGGGACGCTAGCTCTGATAGAG-3′
MutST.2	5′-CTCTATCAGAGCTAGCGTCCCGAATTGTGGCTATATGTC CCGAATGGTAATTCC-3′
MutST.3	5′-GGAATTACCATTCGGGACATATAGCCTACAATTCGGGACGCTAGCTCTGATAGAG-3′
MutST.4	5′-CTCTATCAGAGCTAGCGTCCCGAATTGTAGGCTATATGTCCCGAATGGTAATTCC-3′
MutST.5	5′-CGCTAGCTCTGATAGAGACAATTGAAAACTATACTAAAGATGTGATTGAGAG-3′
MutST.6	5′-CTCTCAATCACATCTTTAGTATAGTTTTCAATTGTCTCTATCAGAGCTAGCG-3′
MutST.7	5′-TAATATACATTAGGAATTACTTTTTGGGACATATAGCCACTTTTTGGGACG-3′
MutST.8	5′-CGTCCCAAAAAGTGGCTATATGTCCCAAAAAGTAATTC CTAATGTATATTA-3′

of the supernatants, the pellets were resuspended in 400 µl ice-cold NaPi buffer (50 mM NaPi, pH 7.0). The resuspended cells were lysed by adding 20 µl acetone:toluene (9:1 vol/vol) followed by 10-min incubation at 37°C and subsequent storage on ice. Then 10 µl aliquots of cell extracts were mixed with 400 µl Gus buffer (50 mM NaPi, pH 7.0; 10 mM β-mercaptoethanol; 1 mM EDTA; 0.1% TritonX-100), before 40 µl of a 16 mg/ml solution of *p*-nitrophenyl-β-D-glucuronide dissolved in 50 mM NaPi, pH 7.0, was added. Samples were protected from light and incubated for 20 min at 37°C. The reactions were stopped by adding 400 µl 0.2 M Na_2CO_3. The increase in A_{405} (ΔA_{405}/min) was used to calculate nanomoles of β-glucuronidase produced per minute and OD at 660 nm (nanomoles/minute/optical density). When the molar adsorption coefficient ε is used (ε for para-nitrophenyl-β-D-glucuronic acid = 18.000 l/mol/cm), the β-glucuronidase activity can be calculated in nanomoles/minute/optical density using the relationship $A = \varepsilon bc$ [A (absorbance) = ε (l/mol/cm) × b (cm pathway through the solution) × c (molar concentration of the absorbing species in M)] (de Ruyter et al. 1996).

Construction of promoter deletion mutants

In order to identify the region of the *stbD* promoter required for pheromone-driven transcriptional activation, three promoter deletion mutants were constructed (Mut 1, Mut 2, and Mut 3; Fig. 2a). The deletion mutants were made by amplifying DNA fragments containing the desired promoter deletion plus the complete downstream *gusA* gene. Each PCR reaction was set up using one promoter-primer (Mut 1, Mut 2, or Mut 3; Table 1) in conjunction with a primer (GusR, Table 1) complementary to the region immediately downstream of the *gusA* gene. The pTB18G plasmid was used as template. To avoid introducing errors during amplification, the Phusion high-fidelity polymerase (Finnzymes) was employed. The annealing temperature was set to 50°C. After 30 cycles, Taq polymerase was added to the samples, and the PCR reaction was incubated for 9 min at 78°C in order to generate 3′-end adenine overhangs for the subsequent TOPO cloning. The products were analyzed by agarose gel electrophoresis, and three PCR-fragments of the expected size were obtained. Then, the PCR reactions were used in TOPO TA-Cloning (Invitrogen) as described by the manufacturer, and transformed into chemically competent *E. coli* TOP10 cells (Invitrogen). Positive clones were selected on LB plates containing 50 µg/ml of kanamycin. The plasmid constructs were purified, and the cloned PCR fragment excised from the TOPO vector using the restriction enzymes *Bam*HI and *Xho*I (Fermentas). After purification by gel extraction (Jet Quick, Saveen) the fragments were ligated into the pTRKH2 plasmid precleaved with *Bam*HI and *Xho*I, and transformed into chemically competent *E. coli* TOP10 cells (Invitrogen). Transformants were selected on LB plates containing 500 µg/ml erythromycin. Before the purified plasmid constructs were transformed into the *S. thermophilus* LMG 18311 strain by

electroporation, the cloned inserts were sequenced to ensure that no undesired mutations had been introduced during the PCR amplifications. Finally, the resulting promoter deletion strains, Mut1, Mut 2, and Mut 3 (see Fig. 2a), were tested in GUS assays to determine whether their shortened promoters still responded to synthetic STP pheromone.

Introduction of point mutations in the putative StbR binding site

The Quick-Change Site-Directed Mutagenesis Kit (Stratagene) was used as described by the manufacturer to make the promoter mutants, Mut 11, 12, V and TG, depicted in Fig. 2b. The pTB18G plasmid was used as template in the PCR reactions together with four pairs of complementary primers (MutST.1–8, see Table 1) designed to construct each of the mutants listed above. Transformants containing the various mutant plasmids were isolated by plating on LB plates containing 500 μg/ml erythromycin. Then, the inserts of the four purified mutant plasmids were sequenced to verify that the correct base substitutions had been introduced. During this work it was discovered that one of the sequenced Mut 11 plasmids contained an extra cytosine 6 bp downstream from the right direct repeat (Fig. 2a). This mutant was called Mut 11c (Fig. 2b). All five mutant plasmids were subsequently transformed into the *S. thermophilus* LMG 18311 strain as described above, giving rise to the Mut 11, 11c, 12, V and TG mutant strains.

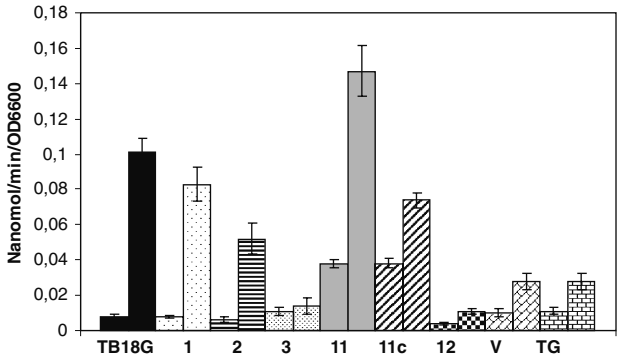

Fig. 3 Effect of various mutations in *stbD* promoter on its responsiveness to exogenously add STP pheromone. The *gusA* gene was used as a reporter. TB18G: wild-type promoter. *1*, *2* and *3*: *stbD* promoter deletion mutants Mut 1, Mut 2 and Mut 3 (see Fig. 2a). *11*, *11c*, *12*, *V* and *TG*: stbD promoter mutants Mut 11, Mut 11c, Mut 12, Mut V, and Mut TG, containing various base substitutions (see Fig. 2b). *Left column* represents uninduced culture, whereas the *right column* represents cultures subjected to 250 ng/ml of STP. Hydrolysis by GusA of the substrate *p*-nitrophenyl-β-D-glucuronide is given in nanomoles/minute/OD_{600}, and is the mean of at least three experiments ± SEM

Results

The genome of *S. thermophilus* contains a quorum-sensing system

When searching the genome of *S. thermophilus* strain LMG 18311 for possible peptide-pheromone receptors, a locus encoding a complete quorum-sensing system with high homology to the corresponding components of the pneumococcal BlpABCHR system was identified (de Saizieu et al. 2000; Hols et al. 2005). This locus consists of at least six genes, StbABCDHR, apparently situated in three different transcription units (Fig. 1). The largest operon encodes a small peptide, StbC, and its putative secretion apparatus StbAB. This operon is organized in the same way as the *blpABC* operon of *S. pneumoniae*. It has previously been shown that *blpC* encodes a secreted peptide pheromone that regulates the transcription of a number of pneumococcal bacteriocin-like genes (de Saizieu et al. 2000; Reichmann and Hakenbeck 2000). Just as BlpC, StbC appears to be synthesized as a precursor peptide containing a double-glycine leader sequence at its N-terminal end. Consequently, it was reasonable to assume that the processed form of StbC, termed STP, acts as a signaling peptide that together with StbAB and the two-component system StbHR regulate bacteriocin production in *S. thermophilus* LMG 18311 by a quorum-sensing mechanism. A closer examination of the *stbB* gene revealed that it contains a frameshift mutation which presumably makes the transporter non-functional and for this reason short-circuits the whole quorum-sensing system. Even though this would destroy bacteriocin production in the LMG 18311 strain, the histidine kinase (StbH) and its cognate response regulator (StbR) appeared to be fully functional. Thus, synthetic STP added to the growth medium should induce transcription of target genes placed behind a bacteriocin promoter. We therefore decided to take advantage of this natural signal transduction pathway by constructing an expression system for recombinant proteins that can be induced by the STP-pheromone.

Construction of an expression vector with a STP-responsive promoter

The *stbD* gene encodes a small 76 aa protein containing an N-terminal double glycine leader sequence. Interestingly, the mature part of StbD, corresponding to the 53 C-terminal amino acids, is 71% identical to a putative bacteriocin encoded by the pneumococcal *blpK* gene. Using DNA microarray analysis, de Saizieu et al. (2000) have shown that expression of BlpK is

regulated by the BlpABCHR quorum-sensing system. In other words, BlpK expression is induced by the peptide-pheromone encoded by the *blpC* gene. On the basis of these results we decided to investigate whether it is possible to exploit the promoter in front of the putative bacteriocin gene *stbD* to drive expression of recombinant proteins in *S. thermophilus*. A ~250 bp DNA fragment, corresponding to the promoter region of the *stbD* gene, was therefore transcriptionally fused to a *gusA* reporter gene (see Materials and methods section for details). β-Glucuronidase activity was then determined in induced and non-induced broth cultures following a 2-h incubation. Interestingly, the results showed that the GUS activity was more than tenfold higher in the induced cells compared to the non-induced cells (Fig. 3). This finding demonstrates that the STP-peptide acts as a pheromone whose presence in the environment is sensed by the LMG 18311 cells, and that this system is suitable for expression of recombinant proteins in *S. thermophilus*. However, for some applications it would be desirable to have a lower background level (non-induced) expression than obtained with the wild-type *stbD* promoter. In order to rule out transcriptional read-through from possible promoters located in the pTRKH2 vector, the terminator of the Arg-tRNA gene of *S. pneumoniae* was inserted upstream of the *stbD* promoter. The properties of this mutant was indistinguishable from wild type (TB18G), demonstrating that the observed background level expression of GusA originates from within the ~250 bp *stbD* promoter region (results not shown).

Deletion mapping of the *stbD* promoter region

By analogy with other pheromone-induced two-component regulatory systems, it is reasonable to assume that binding of the STP-pheromone to its histidine kinase receptor StbH results in autophosphorylation and subsequent transfer of the phosphoryl group to the cognate response regulator StbR. Phosphorylation of StbR will increase its affinity for specific DNA-binding sites located in the promoter regions of target genes such as *stbD*. To identify the binding site of StbR, we made deletions in the *stbD* promoter which progressively removed sequences from the 5′-end of the cloned 250 bp fragment (Fig. 2a). The results clearly showed that removal of the first 50 bp (Mut 1) did not significantly affect the functionality of the promoter (Figs. 2a, 3). Removal of the next 74 bp (Mut 2) almost halved the β-glucuronidase production, but the promoter still responded strongly to STP. In contrast, deletion of the next 45 bp (Mut 3) completely abolished the effect of the peptide-pheromone, demonstrating that the StbR binding site must be located within this fragment. Previous studies on the related quorum-sensing systems ComCDE (Ween et al. 1999) and PlnABCD (Risøen et al. 1998) have shown that the response regulators of these systems (ComE and PlnC) bind as dimers to degenerated 9–10 bp direct repeats separated by a 11–12 bp spacer region. Furthermore, in promoters containing ComE or PlnC binding sites, the direct repeat motifs was found to be separated from the Pribnow box by a spacer of about 30 bp. A close inspection of the 45 bp sequence believed to contain the StbR binding site revealed a possible direct repeat motif ACCATTCGG-ACTTTTTGG separated by 11 bp. In addition, the fact that this motif is separated from a putative downstream Pribnow box (TATACT) by 31 bp makes it a likely candidate for the StbR binding site (Fig. 2a).

Mutational analysis of the putative StbR binding site

Several recent studies have demonstrated that the 9-bp direct repeats are directly involved in binding of the cognate response regulator (Ween et al. 1999; Risøen et al. 2001). To determine whether mutations within these repeats would influence pheromone-induced GusA expression, we made two mutants, Mut 11 and Mut V, with several point mutations in the right and left repeat, respectively (Fig. 2b). Mut V displayed considerable loss of inducibility, showing that the three base substitutions had a clear effect on the functionality of the promoter. In contrast, the Mut 11 mutant displayed increased expression of GusA in STP-induced as well as uninduced cultures. Previous studies have shown that the length of the spacer between the direct repeats is crucial for promoter activity, and that removal of just a single nucleotide renders the promoter non-functional (Risøen et al. 2001; Knutsen et al. 2004). In order to activate transcription from the promoter, the response regulator must bind as a dimer, where the two monomers presumably interact with each other to stabilize the dimeric complex. To obtain further evidence that the identified direct repeat motif constitutes the binding site of StbR, we introduced an extra basepair in the spacer region between the left and right repeat (Fig. 2b). This mutant, termed Mut 12, essentially lost inducibility (Fig. 3), indicating that the extra basepair was introduced between the actual binding sites of StbR, and that the resulting extension of the spacer region strongly impairs formation of the StbR dimeric complex. Extension of the spacer region between the right repeat and the downstream Pribnow box by 1 bp in Mut 11 reduced the inducibility of the resulting mutant (Mut 11c), but did not affect GusA

expression in uninduced Mut 11c cells compared to the Mut 11 parental strain (Figs. 2b, 3). Compared to the other mutants constructed in this study, the background expression of GusA in Mut 11 and Mut 11c is significantly elevated. This must be due to the three base substitutions introduced into the right repeat of these strains.

A drawback of the inducible expression system described above is that the *stbD* promoter is somewhat leaky. When scrutinizing the promoter sequence for possible sigmaA binding sites, we noticed that the Pribnow box is similar to so-called extended -10 promoter sites (Fig. 2a). Such sites, which are common in *S. pneumoniae* and presumably other streptococci, can function naturally without a -35 site (Sabelnikov et al. 1995). We therefore decided to change the putative extended −10 site (TGNTATACT) into an ordinary Pribnow box (AANTATACT) (Fig. 2b), with the intention of reducing the level of GusA expression in uninduced cells. Regrettably, the resulting mutant (TG) displayed the same GusA background level as strain (TB18G) carrying the wild-type promoter (Fig. 3). In addition, the TG mutant promoter turned out to have low inducibility, making this construct unsuitable for expression of recombinant proteins in *S. thermophilus*.

Discussion

The inducible expression system presented in this work was specifically developed for use in *S. thermophilus.* A substantial part of the system is provided by the bacterium's own genetic elements and gene products, and only the pTRKH2 vector represents foreign DNA. However, to develop a complete food-grade inducible expression system, the pTRKH2 shuttle vector will have to be exchanged with a plasmid and selection marker originating from *S. thermophilus*. According to the strict definition of food-grade vectors they must contain only DNA from the same genus or preferentially species as the host bacterium, and antibiotic resistance genes should not be used as selectable markers (Johansen 1999). Several expression systems exploiting the same type of regulation mechanism have been developed previously (Axelsson et al. 2003; Mathiesen et al. 2004; Mierau and Kleerebezem 2005), but none of them was tailor-made for *S. thermophilus* and can therefore not be regarded as food-grade in this organism. Even though the current system is not suitable for use in food production, it represents a valuable research tool for inducible expression of homologous and heterologous proteins in various strains of *S. thermophilus.* The recent sequencing of the complete genomes of three different strains of *S. thermophilus* has opened up new possibilities in the study of the molecular biology and metabolism of this species (Hols et al. 2005). In this context the development of better molecular tools is of considerable importance.

Deletion mapping of the *stbD* promoter region showed that the StbR binding site resides on a 45 bp fragment containing a sequence motif reminiscent of previously characterized binding sites for response regulators showing high homology to StbR (Ween et al. 1999). Mutational analysis supported that the identified degenerated direct repeat motif is important for promoter function, as base substitutions made within the repeats strongly affected STP-induced GusA expression. In addition, extension of the length of the spacer region between the degenerated direct repeats by only 1 bp virtually abolished the promoter's response to STP. Previous research on homologous systems has demonstrated that each repeat binds a response regulator, and that correct spacing between the repeats is crucial for the formation of a functional dimeric complex (Risøen et al. 2001; Knutsen et al. 2004). Together, our results indicate that the identified direct repeat motif represents the StbR binding-site. However, further studies using gel shift and footprinting analyses will be required to confirm our findings, and map the interaction between the repeats and the StbR protein in more detail.

As mentioned above, the *stbD* promoter is somewhat leaky, and for certain purposes it would be desirable to have a more tightly regulated system. We suspected that the observed background level expression was caused by the activity of a weak constitutive sigmaA dependent promoter operating independently of the regulatory elements involved in STP-induced transcription. This hypothesis is supported by the fact that whereas Mut 3 does not respond to STP, uninduced cultures of this mutant display the same level of GusA expression as Mut 1 and 2 (Fig. 3). Transcriptional read-through from the pTRKH2 vector seems unlikely, as the introduction of a rho-independent terminator immediately upstream of the *stbD* promoter did not reduce background expression. Consequently, the constitutive promoter most likely resides between the right repeat and the start codon of the *gusA* gene (Fig. 2a). The only promoter-like element detected in this region was a possible extended −10 site (Fig. 2a). Such sites are quite common in Gram-positive bacteria, and have been shown to be able to function without an accompanying −35 region (Sabelnikov et al. 1995). Unexpectedly, point mutations changing this putative extended −10 site into an ordinary Pribnow

box did not reduce GusA expression in uninduced cells, but made the promoter less responsive to STP (Fig. 3). Thus, at present we are not able to provide an explanation for the observed leakiness of the *stbD* promoter.

Another puzzling observation was the significant increase in background level expression caused by three base substitutions introduced in the right repeat of Mut 11 and Mut 11c (Fig. 3). It has been well documented that phosphorylation of response regulators by their cognate kinases greatly enhances transcription from their target promoters. Interestingly, it has also been reported that at least some unphosphorylated response regulators are able to activate transcription, although at a low level (Ellison and McCleary 2000; Geng et al. 2004). A possible explanation for the elevated background expression of GusA in Mut 11 and Mut 11c is therefore that the introduced base substitutions increase the affinity of the right repeat for unphosphorylated StbR, leading to enhanced transcription of the *gusA* gene in uninduced cells.

Acknowledgments This work was supported by a grant from the Research Council of Norway.

References

Axelsson L, Lindstad G, Naterstad K (2003) Development of an inducible gene expression system for *Lactobacillus sakei*. Lett Appl Microbiol 37:115–120

Bolotin A, Quinquis B, Renault P, Sorokin A, Ehrlich SD, Kulakauskas S, Lapidus A, Goltsman E, Mazur M, Pusch GD, Fonstein M, Overbeek R, Kyprides N, Purnelle B, Prozzi D, Ngui K, Masuy D, Hancy F, Burteau S, Boutry M, Delcour J, Goffeau A, Hols P (2004) Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus.* Nat Biotechnol 22:1554–1558

Buckley ND, Vadeboncoeur C, Leblanc DJ, Lee LN, Frenette M (1999) An effective strategy, applicable to *Streptococcus salivarius* and related bacteria, to enhance or confer electroporation competence. Appl Environ Microbiol 65:3800–3804

Diep DB, Håvarstein LS, Nes IF (1995) A bacteriocin-like peptide induces bacteriocin synthesis in *Lactobacillus plantarum* C11. Mol Microbiol 18:631–639

Eijsink VG, Axelsson L, Diep DB, Håvarstein LS, Holo H, Nes IF (2002) Production of class II bacteriocins by lactic acid bacteria; an example of biological warfare and communication. Antonie van Leeuwenhoek 81:639–654

Ellison DW, McCleary WR (2000) The unphosphorylated receiver domain of PhoB silences the activity of its output domain. J Bacteriol 182:6592–6597

Geng H, Nakano S, Nakano MM (2004) Transcriptional activation by Bacillus subtilis ResD: tandem binding to target elements and phosphorylation-dependent and -independent transcriptional activation. J Bacteriol 186:2028–2037

Hava D, Camilli A (2002) Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Mol Microbiol 45:1389–1405

Håvarstein LS (2003) Intercellular communication in Gram-positive bacteria depends on peptide pheromones and their histidine kinase receptors. In: Inouye M, Dutta R (eds) Histidine kinases in signal transduction. Academic, New York, pp 341–363

Håvarstein LS, Coomaraswamy G, Morrison DA (1995a) An unmodified heptadecapeptide pheromone induces competence for genetic-transformation in *Streptococcus pneumoniae*. Proc Natl Acad Sci USA 92:11140–11144

Håvarstein LS, Diep DB, Nes IF (1995b) A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. Mol Microbiol 16:229–240

Håvarstein LS, Gaustad P, Nes IF, Morrison DA (1996) Identification of the streptococcal competence-pheromone receptor. Mol Microbiol 21:863–869

Hidalgo-Grass C, Dan-Goor M, Maly A, Evan Y, Kwinn LA, Nizet V, Ravins M, Jaffe J, Peyser A, Moses AE, Hanski E (2004) Effect of a bacterial pheromone peptide on host chemokine degradation in group A streptococcal necrotising soft-tissue infections. Lancet 363:696–703

Hols P, Hancy F, Fontaine L, Grossiord B, Prozzi D, Leblond-Bourget N, Decaris B, Bolotin A, Delorme C, Ehrlich SD, Guédon E, Monnet V, Renault P, Kleerebezem M (2005) New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed by comparative genomics. FEMS Microbiol Rev 29:435–463

Johansen E (1999) Genetic engineering (b) Modification of bacteria. In: Robinson R, Batt C, Patel P (eds) Encyclopedia of food microbiology. Academic, London, pp 917–921

Kleerebezem M (2004) Quorum sensing control of antibiotic production; nisin and subtilin autoregulate their own biosynthesis. Peptides 25:1405–1414

Knutsen E, Ween O, Havarstein LS (2004) Two separate quorum-sensing systems upregulate transcription of the same ABC transporter in *Streptococcus pneumoniae*. J Bacteriol 186:3078–3085

Li YH, Aspiras MB, Lau PC, Lee JH, Ellen RP, Cvitkovitch DG (2002) A quorum-sensing signaling system essential for genetic competence in *Streptococcus mutans* is involved in biofilm formation. J Bacteriol 184:2699–2708

Mathiesen G, Sørvig E, Blatny J, Naterstad K, Axelsson L, Eijsink VGH (2004) High-level gene expression in *Lactobacillus plantarum* using a pheromone-regulated bacteriocin promoter. Lett Appl Microbiol 39:137–143

Mierau I, Kleerebezem M (2005) 10 years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*. Appl Microbiol Biotechnol 68:705–717

O'Sullivan DJ, Klaenhammer TR (1993) High-copy-number and low-copy-number *Lactococcus* shuttle cloning vectors with features for clone screening. Gene 137:227–231

Pestova EV, Håvarstein LS, Morrison DA (1996) Regulation of competence for genetic transformation in *Streptococcus pneumoniae* by an auto-induced peptide pheromone and a two-component regulatory system. Mol Microbiol 21:853–862

van der Ploeg JR (2005) Regulation of bacteriocin production in *Streptococcus mutans* by the quorum-sensing system required for development of genetic competence. J Bacteriol 187:3980–3989

Reichmann P, Hakenbeck R (2000) Allelic variation in a peptide-inducible two-component system of *Streptococcus pneumoniae*. FEMS Microbiol Lett 190:231–236

Risøen PA, Havarstein LS, Diep DB, Nes IF (1998) Identification of the DNA-binding sites for two response regulators

involved in control of bacteriocin synthesis in *Lactobacillus plantarum* C11. Mol Gen Genet 259:224–232

Risøen PA, Johnsborg O, Diep DB, Hamoen L, Venema G, Nes IF (2001) Regulation of bacteriocin production in *Lactobacillus plantarum* depends on a conserved promoter arrangement with consensus binding sequence. Mol Genet Genomi 265:198–206

de Ruyter PGGA, Kuipers OP, Beerthuyzen MM, van Alen Boerrigter I, de Vos WM (1996) Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*. J Bacteriol 178:3434–3439

Sabelnikov AG, Greenberg B, Lacks SA (1995) An extended −10-promoter alone directs transcription of the *Dpn*II *o*peron of *Streptococcus pneumoniae*. J Mol Biol 250:144–155

de Saizieu A, Gardes C, Flint N, Wagner C, Kamber M, Mitchell TJ, Keck W, Amrein KE, Lange R (2000) Microarray-based identification of a novel *Streptococcus pneumoniae* regulon controlled by an autoinduced peptide. J Bacteriol 182:4696–4703

Saucier L, Paradkar AS, Frost LS, Jensen SE, Stiles ME (1997) Transcriptional analysis and regulation of carnobacteriocin production in *Carnobacterium piscicola* LV17. Gene 188:271–277

Sørvig E, Grønqvist S, Naterstad K, Mathiesen G, Eijsink VGH, Axelsson L (2003) Construction of vectors for inducible gene expression in *Lactobacillus sakei* and *L. plantarum*. FEMS Microbiol Lett 229:119–126

Sørvig E, Mathiesen G, Naterstad K, Eijsink VG, Axelsson L (2005) High-level, inducible gene expression in Lactobacillus sakei and Lactobacillus plantarum using versatile expression vectors. Microbiology 151:2439–2449

Waters CM, Bassler BL (2005) Quorum sensing: cell-to-cell communication in bacteria. Annu Rev Cell Dev Biol 21:319–346

Ween O, Gaustad P, Havarstein LS (1999) Identification of DNA binding sites for ComE, a key regulator of natural competence in *Streptococcus pneumoniae*. Mol Microbiol 33:817–827

Paper II

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Oct. 2006, p. 6751–6756 Vol. 72, No. 10
0099-2240/06/$08.00+0 doi:10.1128/AEM.01156-06

Natural Genetic Transformation: a Novel Tool for Efficient Genetic Engineering of the Dairy Bacterium *Streptococcus thermophilus*

Trinelise Blomqvist, Hilde Steinmoen, and Leiv Sigve Håvarstein*

Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, N-1432 Ås, Norway

Received 19 May 2006/Accepted 16 July 2006

Streptococcus thermophilus **is widely used for the manufacture of yoghurt and Swiss or Italian-type cheeses. These products have a market value of approximately $40 billion per year, making *S. thermophilus* a species that has major economic importance. Even though the fermentation properties of this bacterium have been gradually improved by classical methods, there is great potential for further improvement through genetic engineering. Due to the recent publication of three complete genome sequences, it is now possible to use a rational approach for designing *S. thermophilus* starter strains with improved properties. Progress in this field, however, is hampered by a lack of genetic tools. Therefore, we developed a system, based on natural transformation, which makes genetic manipulations in *S. thermophilus* easy, rapid, and highly efficient. The efficiency of this novel tool should make it possible to construct food-grade mutants of *S. thermophilus*, opening up exciting new possibilities that should benefit consumers as well as the dairy industry.**

Bacteria that are competent for natural genetic transformation are able to take up naked DNA from the environment and incorporate it into their genomes by homologous recombination. Several streptococcal species belonging to the mitis, anginosus, and mutans phylogenetic groups have been shown to possess this property (4, 9, 17), but the phenomenon has never been demonstrated in most members of the genus *Streptococcus*. One of the best-studied naturally competent bacteria is *Streptococcus pneumoniae.* In this species and other streptococci shown to be naturally transformable, competence is not a constant property; rather, it is a transient state regulated by a quorum-sensing mechanism consisting of ComABCDE (4). *comC* encodes the precursor of a secreted peptide pheromone, the competence-stimulating peptide, which triggers development of the competent state when its external concentration in a pneumococcal culture reaches a critical threshold (7). The competence-stimulating peptide is secreted by ComAB (11) and acts through a two-component signal transduction pathway consisting of the histidine kinase ComD and the cognate response regulator ComE (4, 8, 21). The early genes are regulated by ComE, whereas the alternative sigma factor ComX is needed for expression of the late genes (5, 14, 22). Late genes share an 8-bp sequence in their promoter regions that is specifically recognized by a ComX-directed RNA polymerase holoenzyme (14). Circumstantial evidence indicates that ComX is encoded by one of the early genes and therefore depends on ComE for its expression (4). The 14 pneumococcal proteins known to be necessary for uptake of extracellular DNA and for subsequent incorporation of this DNA into the recipient's genome are all encoded by late genes (5, 22). Interestingly, recent genome sequencing has shown that the ComX regulon appears to be present in all streptococcal species (17). This finding suggests that most streptococci are naturally transformable provided that growth conditions promoting development of competence can be identified. Alternatively, the late genes of streptococcal species not known to be competent may have other functions or represent nonfunctional relicts inherited from a competent ancestor.

* Corresponding author. Mailing address: Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway. Phone: 47-64965883. Fax: 47-64965901. E-mail: sigve.havarstein@umb.no.

MATERIALS AND METHODS

Bacterial strains and growth media. *Streptococcus thermophilus* strains LMG 18311 (= ATCC BAA-250) and LMD-9 (= ATCC BAA-491) were cultivated in Todd-Hewitt broth (Difco Laboratories) supplemented with 0.8% glucose (THG) or in Hogg-Jago glucose broth (HJG) consisting of 3% tryptone, 1% yeast extract, 0.2% beef extract, 0.5% KH_2PO_4, and 0.5% glucose. HJGL is Hogg-Jago glucose broth supplemented with 0.5% lactose, whereas HJGLS is HJGL supplemented with 0.4 M D-sorbitol. Agar plates were prepared by adding 1.5% (wt/vol) agar to the media.

Construction of plasmids. The reporter plasmids pXP, pEAP, and pBP were constructed by fusing the promoters of *comX*, *comEA* (a late gene encoding part of the DNA uptake apparatus), and *stbD* (encoding a putative bacteriocin) to the firefly luciferase gene and ligating the resulting fragments into the pTRKH2 shuttle vector (20). Briefly, the luciferase gene was amplified in three separate PCRs using primer pairs LXP/LR (pXP), LCB/LR (pEAP), and LBP/LR (pBP) and a plasmid (pR424) carrying the *luc* gene as the template (3). Similarly, PCRs performed with primer pairs CXPF/CXPL, CBF/CBL, and BPF/BPL and genomic DNA from *S. thermophilus* LMG 18311 were used to amplify fragments corresponding to the *comX* (~440-bp), *comEA* (~210-bp), and *stbD* (~250-bp) promoters, respectively. Then promoter and *luc* gene fragments with complementary overlapping ends were combined and amplified by PCRs using the appropriate external primers. Primer pairs CXPF/LR, CBF/LR, and BPF/LR were used to generate the XL, EAP, and BP fragments, respectively. Finally, the three fragments were cloned into the pCR 2.1-TOPO vector (Invitrogen), excised by XhoI and PstI, and ligated into the corresponding sites of the pTRKH2 vector. The resulting reporter plasmids, pXP, pEAP, and pBP, were electroporated into *S. thermophilus* LMG 18311 as described below, giving rise to the XP, EAP, and BP strains.

To construct the pXL plasmid, a DNA fragment containing the *comX* gene joined to the promoter of *stbD* was ligated into the pEAP plasmid (see above). The fragments corresponding to the bacteriocin promoter (P_{stbD}) and the *comX* gene were amplified from *S. thermophilus* LMG 18311 DNA, using primers P1 and P2 and primers X1 and X2, respectively. The P2 and X1 primers contain NcoI sites at their 5′ ends coinciding with the start codon of the *comX* gene. Next, the P_{stbD} and *comX* fragments were cloned separately into the pDrive vector (QIAGEN). The *comX* fragment was excised from pDrive by digestion with NcoI and XbaI and ligated into the corresponding sites of the pDrive vector carrying the *stbD* promoter fragment. Then the joined P_{stbD}::*comX* fragment was excised

TABLE 1. Primers used in this study

Primer	Sequence
CBF	5′-AGTGTAACTGCAGAATACTTGCAGGTCTA TCGATCGAT-3′
CBL	5′-TTTGGCGGATCTCATAAGGACCTCCTCAT AAACCTATTC-3′
CXPF	5′-CGCTTTCTGCAGCTATCACTCTAATACAAT CCTGTGGAA-3′
CXPL	5′-TTGGCGGATCTCATTGAACCTCCAATAAT AAATATAAATTCTGT-3′
BPF	5′-GTAAATCTGCAGCTTCAAGGTCTAGTCCT CTCT-3′
BPL	5′-TTGGCGGATCTCATGGCAACTACCTCCTA AAATTTTTATC-3′
LCB	5′-TATGAGGAGGTCCTTATGAGATCCGCCAA AAACAT-3′
LR	5′-CATATGGCTCGAGTGCACTCTCAGTACAA TCTGCTC-3′
LXP	5′-TATTGGAGGTTCAATGAGATCCGCCAAAA ACATAAAGAAAGGC-3′
LBP	5′-TAGGAGGTAGTTGCCATGAGATCCGCCAA AAACATAAAGAAAGGC-3′
P1	5′-GTTTGAGTTG CCATGGCAACTACCTCC-3′
P2	5′-ATTAGGATCCTTCAAGGTCTAGTCCTCTCT TTTATGACG-3′
X1	5′-ATTATCTAGACCAAGAATTACTGGAAACA CAATAGAGG-3′
X2	5′-GGAGGTTCCATGGAACAAGAAGTTTTTGT TAAGGC-3′
EC1	5′-GAGGCATCATTGGAAGAATAGAGCAGC-3′
EC2	5′-AAGCTTAAGATCTAGAGCTCGAGGATCAA AAACTAGAGAGAAGATTGCCGTCAG-3′
EC3	5′-AGCATGCATATGCATCCGGAGTCCTAGCT TGTTTCAGTTTGTCTCAATG-3′
EC4	5′-CCATCCCTTAAACCGAATGGCACC-3′
Kana-F	5′-ATCCTCGAGCTCTAGATCTTAAGCTT-3′
Kana-R	5′-ACTCCGGATGCATATGCATGCT-3′

from the pDrive vector by digestion with PstI and SacI and ligated into pTRKH2 precleaved with the same enzymes. Finally, the P_{stbD}::*comX* fragment was excised from pTRKH2 by digestion with BamHI and EcoRV and ligated into pEAP precleaved with BamHI ans SmaI. The resulting construct, pXL, contained an expression module (P_{stbD}::*comX*) and a reporter module (P_{comEC}::*luc*) inserted in opposite orientations. All PCRs described above were carried out with the Phusion high-fidelity DNA polymerase (Finnzymes). The sequences of the primers used are shown in Table 1.

Preparation of electrocompetent *S. thermophilus* LMG 18311 cells. An overnight culture grown at 37°C was diluted 100-fold in preheated HJG (37°C) and incubated until the optical density at 660 nm (OD_{660}) was 0.3. The culture (50 ml) was then diluted 1:1 in prewarmed HJG containing 20% glycine. After incubation at 37°C for 1 h, cells were harvested by centrifugation (4,000 × *g* for 10 min at 4°C) and washed twice with 1 volume of ice-cold electroporation buffer (5 mM $KHPO_4$, 0.4 M D-sorbitol, 10% glycerol; pH 4.5). Finally, pelleted cells were resuspended in 4 ml ice-cold electroporation buffer, divided into aliquots, and frozen in an ethanol-dry ice bath. Electrocompetent *S. thermophilus* LMG 18311 cells were stored at −80°C.

Electroporation. A Bio-Rad MicroPulser unit was used to transform *S. thermophilus* LMG 18311 cells by electroporation. Electrocompetent cells were thawed on ice, and an 80-μl cell suspension was mixed with 1 μg of recombinant pTRKH2 plasmid DNA. After incubation for 30 min on ice, the cells were transferred to an electroporation cuvette with a 0.1-cm gap between the electrodes. A single pulse of 1.6 kV lasting 2.5 ms was delivered. The electroporated cells were immediately resuspended in 1 ml of ice-cold HJGLS and incubated for 3 h at 37°C, before they were spread on HJGL plates containing 2 μg/ml erythromycin. Transformants were picked following 24 to 48 h of incubation at 37°C. Isolated clones were verified by PCR using primers M13F and M13R, which were complementary to sequences flanking the multiple cloning site of the pTRKH2 plasmid.

Luciferase reporter assay. Detection of luciferase activity was performed essentially as previously described by Chastanet et al. (3). Strains were grown in THG to an OD_{550} of 0.4, aliquoted, and maintained as glycerol stocks at −80°C. Shortly before use, glycerol stocks were thawed and diluted 10 times in THG. For each test sample, 280 μl of diluted culture was mixed with 20 μl of firefly D-luciferin (10 mM solution in THG) and transferred into a 96-well Corning NBS plate with a clear bottom. If appropriate, the peptide pheromone known as *Streptococcus thermophilus* pheromone (STP) was added to a final concentration of 250 ng/ml immediately before the experiment was started. The plate was incubated at 37°C in an Anthos Lucy 1 luminometer for 7.5 h. The OD_{492} and luminescence were measured automatically by the luminometer at 10-min intervals.

Natural transformation of *S. thermophilus* LMG 18311. *S. thermophilus* LMG 18311 cells harboring pXL were grown overnight at 37°C in Todd-Hewitt broth (Difco Laboratories) supplemented with 0.8% glucose and 2 μg/ml erythromycin. The next day the culture was diluted to obtain an OD_{550} of 0.5 in the same medium prewarmed to 37°C. Then 1 ml of the diluted culture was transferred to a 1.5-ml Eppendorf tube containing 250 ng STP, and the sample was placed in a water bath at 37°C. Two hours later transforming DNA was added. At this stage the OD_{550} of the culture did not exceed ~0.3. The preparation was then incubated for an additional 2 h at 37°C. Finally, the sample was put on ice, serially diluted, and spread on HJGL agar plates containing the appropriate antibiotic (50 μg/ml streptomycin or 100 μg/ml kanamycin). To avoid loosing the pXL plasmid, 2 μg/ml erythromycin had to be added to the HJGL agar plates. Curing of the pXL plasmid was performed by cultivating transformants in antibiotic-free medium for about 100 generations.

Disruption of the *comEC* gene. The *comEC* gene disruption cassette consists of a kanamycin resistance gene flanked by two 800- to 1,000-bp DNA fragments corresponding to the 5′ and 3′ regions of the *comEC* gene. In the first step for disruption of the *comEC* gene, the kanamycin resistance gene was amplified by PCR from the pFW13 vector (23), using primers Kana-F and Kana-R. In the second step, the 5′ and 3′ flanking fragments were generated in two separate PCRs with primer pairs EC1/EC2 and EC3/EC4 and genomic DNA from *S. thermophilus* LMG 18311 as the template. The EC2 and EC3 primers used to amplify the flanking sequences contained 22-bp extensions homologous to the 5′ and 3′ ends of the kanamycin gene, respectively. After agarose gel purification of all PCR fragments, the kanamycin resistance gene was first joined to the 5′ flanking fragment in a PCR performed with the two DNA fragments and the EC1 and Kana-R primers. In the same way, the kanamycin resistance gene was joined to the 3′ flanking fragment in a PCR performed with both fragments and the Kana-F and EC4 primers. Finally, the two combined fragments were joined in a PCR performed with the EC1 and EC4 primers. The resulting *comEC* gene disruption cassette was purified with a PCR purification kit obtained from QIAGEN and was used directly to transform competent *S. thermophilus* LMG 18311 cells carrying the pXL plasmid. In addition, the gene disruption cassette was cloned into the pCR 2.1-TOPO vector (Invitrogen) according to the manufacturer's instructions.

RESULTS

ComX is expressed in *S. thermophilus* during early logarithmic growth. The recent sequencing of three *S. thermophilus* genomes opened the door to rational metabolic engineering of this important dairy species (2, 10, 25). However, in order to take full advantage of the available sequence information, better genetic tools have to be developed (26). Natural transformation is a highly efficient tool for genetic manipulation that has been used successfully with *S. pneumoniae* for more than 60 years. Drawing on our experience with *S. pneumoniae*, we therefore set out to determine whether *S. thermophilus* is a naturally transformable species. We chose to work on *S. thermophilus* LMG 18311, which was isolated from yoghurt produced in the United Kingdom in 1974, because close inspection of its genome sequence indicated that the ComX regulon is intact in this strain. Initially, experiments were carried out to establish if transformants could be obtained by adding homologous genomic DNA containing a streptomycin resistance marker to LMG 18311 cultures grown under various conditions. All results were negative, suggesting that ComX and/or the late genes are not expressed under the conditions used. To

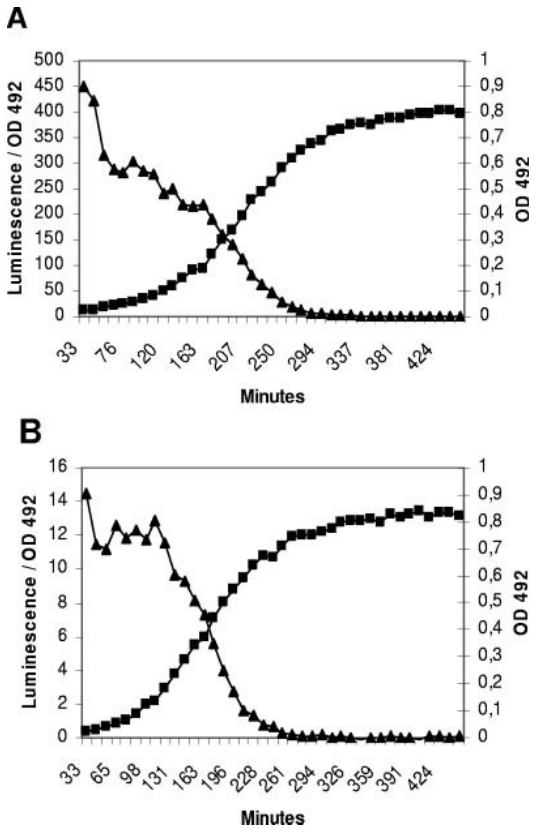

FIG. 1. Expression of the *luc* reporter gene (▲) during growth of *S. thermophilus* LMG 18311 from logarithmic to stationary phase (■). (A) Expression of luciferase driven by the *comX* promoter. (B) Expression of luciferase driven by the *comEC* late gene promoter.

be able to monitor the activity of the *comX* promoter in a growing culture of LMG 18311 cells over time and under various conditions, we used the shuttle plasmid pTRKH2 (20) to construct a reporter plasmid, pXP, harboring a transcriptional fusion between the *comX* promoter and the firefly luciferase gene. The pXP plasmid was subsequently introduced into *S. thermophilus* LMG 18311 by electroporation, giving rise to the XP strain. Luciferase activity was monitored by growing cultures of the XP strain at 37°C in a Lucy 1 luminometer (Anthos) as described in Materials and Methods. Unexpectedly, we discovered that the *comX* promoter is active during early to approximately mid-logarithmic phase in XP cells grown in THG at 37°C. As the culture approached stationary phase, the activity of the *comX* promoter declined to zero (Fig. 1A).

Low-level expression of late genes during early logarithmic phase. Even though ComX is expressed in early logarithmic phase, we were not able to obtain transformants when cultures at this stage of growth were exposed to purified genomic DNA from a streptomycin-resistant mutant of strain LMG 18311. One possible explanation for this negative result is that there were undetected loss-of-function mutations in the transformation machinery. Alternatively, the level of ComX produced might have been too low to significantly activate expression of the late genes. To determine whether this could be the case, we constructed a plasmid similar to pXP, except that we exchanged the *comX* promoter with the promoter of the late gene *comEA* (stu1562). The resulting plasmid, pEAP, was electroporated into *S. thermophilus* LMG 18311, giving rise to the EAP strain. The activity of the *comEA* promoter was monitored by growing the EAP strain in the Lucy 1 luminometer exactly as described above for the XP strain. The data obtained revealed that there was low-level luciferase expression during early logarithmic phase, a pattern that roughly coincided with the activity of the *comX* promoter (Fig. 1A and B). From the reporter assay alone, it was not possible to determine whether the *comEA* promoter operated at a very low level in all bacteria in the culture or if it was highly expressed in just a tiny fraction of the cells. In any case, the results indicate that the level of ComX produced was too low to turn on the competent state in a significant fraction of the bacterial population.

Overexpression of ComX. Morrison and coworkers have shown that in *S. pneumoniae* ectopic expression of ComX from a raffinose-inducible promoter in a *comX* mutant does not result in development of competence. However, when the *comX* gene was cloned into a multicopy plasmid and expressed from a nisin-inducible promoter, the number of transformants obtained was almost 10% of the wild-type number (15). In later studies these workers were able to show that a product of the early genes, termed ComW, is needed in addition to ComX for efficient induction of competence. Evidence indicates that ComW contributes to the stabilization of ComX against proteolysis and that it is also required for optimal functionality of the sigma factor (16, 24). We were not able to identify any homologue of ComW in *S. thermophilus*, but it is reasonable to assume that ComX is unstable also in this species. Judging from our results and the results of the studies carried out by Morrison and coworkers, it should be possible to induce competence in *S. thermophilus* if sufficiently high levels of ComX can be obtained. Using a strong constitutive promoter would probably not work, as constant high levels of ComX would interfere with the normal transcription pattern of the cell. Instead, a system that could provide high-level transient expression of ComX is needed. We recently developed an inducible expression system for *S. thermophilus*, and we decided to use this method for overexpression of ComX (1a). This method exploits a quorum-sensing system consisting of a peptide pheromone (STP) encoded by *stbC* (stu1688), its secretion apparatus (StbAB), its histidine kinase receptor (StbH) (stu1687), and the cognate response regulator StbR (stu1686). The STP secretion apparatus appears to be nonfunctional in *S. thermophilus* LMG 18311, due to a frameshift mutation in the *stbB* gene. Nevertheless, we found that synthetic STP (NH_2-SGWMDYIN GFLKGFGGQRTLPTKDYNIPQV-COOH), added to a culture of the LMG 18311 strain, activates transcription of the bacteriocin-like gene, *stbD* (stu1685), situated immediately upstream of *stbR* in the opposite transcriptional orientation. Consequently, by constructing transcriptional fusions between the *stbD* promoter and target genes, inducible expression of recombinant proteins can be obtained in *S. thermophilus.* To determine the efficiency of the expression system under various growth conditions, a *luc* reporter gene was placed behind the *stbD* promoter. The resulting reporter construct (pBP) was introduced into *S. thermophilus* LMG 18311 by electroporation, giving rise to the BP strain. Using this strain, we tested the STP-inducible expression system in different media and found that the level of lumines-

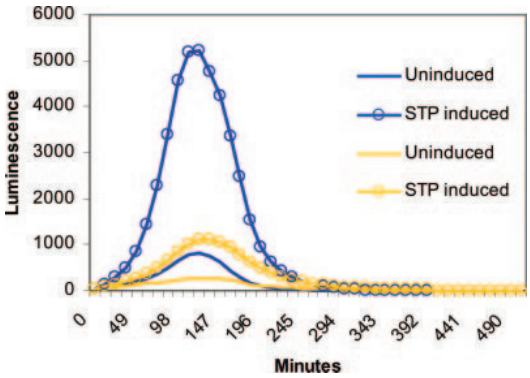

FIG. 2. Effect of the growth medium on STP-induced overexpression of the luciferase reporter protein. Expression of the *luc* reporter gene was driven by the promoter of the putative bacteriocin gene, *stbD*. The *S. thermophilus* BP strain was grown in THG (blue lines) or HJGL (yellow lines) from logarithmic to stationary phase. The STP peptide pheromone was added at time zero.

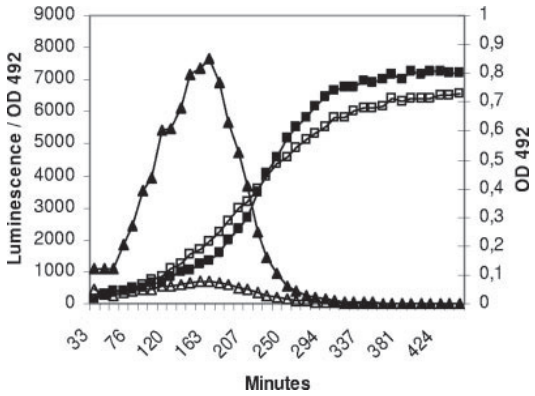

FIG. 3. Transcriptional activation of late competence genes by STP-induced overexpression of ComX in *S. thermophilus* LMG 18311 carrying the pXL plasmid. Plasmid pXL contains an expression cassette consisting of the STP-responsive *stbD* promoter transcriptionally fused to the *comX* gene. In addition, it contains a reporter cassette consisting of the *comEC* late gene promoter transcriptionally fused to the *luc* reporter gene. As ComX activates transcription from late gene promoters, this construct makes it possible to monitor expression of late competence genes in STP-induced (▲) and uninduced (△) *S. thermophilus* cultures. ■, growth curve for STP-induced culture; □, growth curve for uninduced culture. The STP peptide pheromone was added at time zero.

cence obtained was by far the highest in THG (Fig. 2). This medium was therefore used for subsequent overexpression of ComX.

Induction of the competent state in *S. thermophilus*. To determine whether the new expression system could drive ComX production to the level required for activating transcription of the late genes, a new plasmid based on pEAP was constructed. This plasmid (pXL) was constructed by ligating a DNA fragment consisting of a transcriptional fusion between the *stbD* promoter and the *comX* gene into unique restriction sites of the pEAP plasmid. To avoid transcriptional readthrough, the expression and reporter modules of the pXL plasmid were inserted in opposite directions. The resulting construct was introduced into the LMG 18311 strain by electroporation. The effect of STP-induced overproduction of ComX on late gene expression was subsequently monitored by measuring light emission from XL cells with the Lucy 1 luminometer. The results showed that a culture of XL cells treated with 250 ng/ml of STP displayed an approximately 600-fold increase in luminescence compared to a corresponding culture of bacteria harboring the pEAP plasmid (Fig. 1B and 3). No effect on luciferase expression was seen when cultures of the EAP and XP strains were treated with the STP peptide pheromone, demonstrating that the dramatic increase in light production observed with the XL strain must have been due to STP-induced overexpression of ComX. Our results also revealed that ComX is expressed in uninduced XL cells due to a leaky *stbD* promoter. However, the peak luminescence of cultures treated with STP was about sevenfold higher than the luminescence of uninduced cultures (Fig. 3). By chance we discovered that in the absence of selection pressure the pXL plasmid is rapidly lost from its host. Presumably, the presence of ComX, expressed from the leaky *stbD* promoter, disturbs the normal functions of the bacterial cell.

Having constructed an inducible ComX expression system that efficiently activated transcription from late gene promoters, we were anxious to find out whether the transformation machinery of *S. thermophilus* LMG 18311 was still functional. To our delight, 3×10^3 streptomycin-resistant CFU per ml of culture (standard error, $\pm 0.9 \times 10^3$ CFU per ml; $n = 4$) were obtained when genomic DNA (1 μg/ml) that was isolated from a streptomycin-resistant LMG 18311 mutant (Str^r through spontaneous conversion) was added to cultures of the XL strain pretreated with the STP pheromone (see Materials and Methods for experimental details). The total number of CFU in the culture was estimated in parallel and was determined to be 5×10^8 CFU per ml (standard error, $\pm 1 \times 10^8$ CFU per ml; $n = 4$). To ensure that the observed acquisition of streptomycin resistance had taken place by natural transformation, we checked whether the process depended on a functional *comEC* gene. ComEC, also called CelB, has been shown to be essential for natural transformation in *Bacillus subtilis*, *S. pneumoniae*, and other competent bacteria and is believed to encode a transmembrane channel required for DNA internalization (1, 6). The gene encoding ComEC is located on the same transcription unit as *comEA*. To disrupt the *comEC* gene, we used PCR to generate a DNA fragment consisting of a kanamycin marker fused to ~1,000-bp flanking regions amplified from the 5′ and 3′ halves of the *comEC* gene of *S. thermophilus*. This fragment was added to an STP-induced culture of the XL strain at a concentration of 1 μg/ml. After incubation for 2 h at 37°C, the bacteria were spread on agar plates containing 100 μg/ml of kanamycin and then incubated at 37°C for 18 to 24 h. We obtained 7×10^3 CFU per ml (standard error, $\pm 1 \times 10^3$ CFU per ml; $n = 4$) on the agar plates containing kanamycin, and $\sim 5 \times 10^8$ CFU per ml on the control plates lacking the antibiotic. Correct integration of the gene disruption cassette into the *comEC* gene by double-crossover homologous recombination was verified by PCR in 10 randomly picked kanamycin-resistant colonies. Next, we tested the transformability of the XL Δ*comEC* strain using genomic DNA from the streptomycin-resistant LMG 18311 mutant as a selectable marker. No transformants were obtained, demonstrating that the XL strain

became nontransformable when the *comEC* gene was disrupted.

When transformation is performed with a linear DNA fragment, such as the *comEC* gene disruption cassette described above, the ends of the fragments may be attacked and shortened by nucleases, resulting in reduced transformation efficiency (13). In an attempt to further increase the transformation efficiency, we protected the ends of the *comEC* gene disruption cassette by cloning it into the pCR2.1-TOPO plasmid (Invitrogen). By using this strategy we obtained 3×10^6 kanamycin-resistant CFU per ml (standard error, $\pm 0.4 \times 10^6$ CFU per ml; $n = 4$) when 3 μg/ml of plasmid DNA was added to STP-induced cultures of the XL strain. The total number of CFU per ml of competent culture was the same as before ($\sim 5 \times 10^8$ CFU per ml). As described above, correct integration of the *comEC* gene disruption cassette was verified by PCR in 10 randomly picked colonies. These results showed that approximately 1% of the streptococcal chains present in the competent culture receiving 3 μg/ml of recombinant plasmid DNA gave rise to a colony when they were cultivated on agar plates containing kanamycin.

Unexpectedly, preliminary results indicated that *S. thermophilus* strain LMD-9 transformed very poorly using the method described above for strain LMG 18311. Therefore, more strains must be tested before it becomes clear whether the method described here works well with most *S. thermophilus* strains or whether we were just lucky to pick LMG 18311.

DISCUSSION

By using the highly efficient transformation procedure described above it should be possible to introduce mutations into the genome of *S. thermophilus* LMG 18311 without the use of a selectable antibiotic resistance marker. DNA fragments containing the desired insertion/deletion or point mutation(s) can be made by PCR or other molecular methods and cloned into a suitable plasmid. Following uptake of this construct by competent *S. thermophilus* LMG 18311 cells, targeted integration of the mutated region into the bacterial genome is mediated by ~1,000-bp flanking regions through double-crossover homologous recombination. Due to the high transformation efficiency it should be relatively easy to identify transformants containing the sought-after genotype against the background of wild-type streptococci. Standard colony hybridization with a labeled oligonucleotide probe designed to specifically recognize mutants could be used for this purpose. Before plating, a mechanical blender (e.g., Ultra-Turrax model T25 [Ika Labotechnik, Staufen, Germany]) must be used to disrupt the long chains of *S. thermophilus* cells, as described previously (18). After identification of the desired mutant, it is easily cured of the unstable pXL helper plasmid by growth in the absence of erythromycin. *S. thermophilus* mutants made by using this technique fulfill the safety criteria described by Johansen (12) and should therefore attain "generally recognized as safe" status provided that DNA from organisms that are not generally recognized as safe is not introduced into the genetically engineered strain.

In the present work we showed that overexpression of ComX induces the competent state in *S. thermophilus* LMG 18311. An important question that remains to be answered, however, is how natural transformation is turned on spontaneously in this strain. Although it is possible that the mechanism controlling the development of competence has degenerated during adaptation to the dairy niche, it is more likely that spontaneous development of competence in *S. thermophilus* LMG 18311 requires special, as-yet-undiscovered growth conditions. The regulatory pathway controlling expression of the *comX* gene is not known, but our results show that this gene is actively transcribed during early logarithmic phase when the XP strain is grown in THG at 37°C. In spite of this, the levels of transcription of the late genes under these conditions stayed very low, indicating that the level of *comX* transcription was too low or that ComX was prevented from accumulating to levels required for late gene expression by a regulatory mechanism operating at the posttranscriptional level. It has been reported that the ClpP protease negatively controls ComX in *S. pneumoniae* (24) and *Streptococcus pyogenes* (19), and it is therefore likely that the same control mechanism operates in *S. thermophilus*. The fact that overexpression of ComX efficiently induces expression of the late genes suggests that a system that negatively controls the accumulation of ComX becomes saturated under these circumstances. In sum, preliminary data indicate that spontaneous induction of competence in *S. thermophilus* requires the joint action of at least two converging regulatory pathways. However, in contrast to other streptococci that have been shown to be competent for natural transformation, a quorum-sensing system of the ComCDE type does not seem to be involved.

Considering the high degree of degeneracy detected in the genome of *S. thermophilus*, it is remarkable that the genes involved in natural transformation have remained intact. Bolotin et al. (2) found that 10% of the genes in *S. thermophilus* strains LMG 18311 and CNRZ 1066 are nonfunctional pseudogenes, and they concluded that these strains have adapted to the dairy niche mainly through loss-of-function events. The intactness of the late competence genes in strain LMG 18311 indicates that even in a constant milk environment natural competence provides a selective advantage. Indeed, evidence of lateral gene transfer from other dairy bacteria to *S. thermophilus* LMG 18311 has been reported. A 17-kb mosaic region found in the *pepD* gene contains fragments with high homology to corresponding sequences in *Lactobacillus bulgaricus* and *Lactococcus lactis*, species that come into close contact with *S. thermophilus* during fermentation of yoghurt and cheeses, respectively (2). In light of the results presented here, it seems plausible that at least some of these gene transfer events have taken place by natural genetic transformation.

ACKNOWLEDGMENT

This work was supported by grants from the Research Council of Norway.

REFERENCES

1. **Berge, M., M. Moscoso, M. Prudhomme, B. Martin, and J. P. Claverys.** 2002. Uptake of transforming DNA in Gram-positive bacteria: a view from *Streptococcus pneumoniae*. Mol. Microbiol. **45:**411–421.

1a. **Blomqvist, T., H. Steinmoen, L. S. Håvarstein.** Pheromone-induced expression of recombinant proteins in *Streptococcus thermophilus*. Arch. Microbiol., in press.

2. **Bolotin, A., B. Quinquis, P. Renault, A. Sorokin, S. D. Ehrlich, S. Kulakauskas, A. Lapidus, E. Goltsman, M. Mazur, G. D. Pusch, M. Fonstein, R. Overbeek, N. Kyprides, B. Purnelle, D. Prozzi, K. Ngui, D. Masuy, F. Hancy, S.**

Burteau, M. Boutry, J. Delcour, A. Goffeau, and P. Hols. 2004. Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. Nat. Biotechnol. **22:**1554–1558.

3. **Chastanet, A., M. Prudhomme, J. P. Claverys, and T. Msadek.** 2001. Regulation of *Streptococcus pneumoniae clp* genes and their role in competence development and stress survival. J. Bacteriol. **183:**7295–7307.
4. **Claverys, J. P., and L. S. Håvarstein.** 2002. Extracellular-peptide control of competence for genetic transformation in *Streptococcus pneumoniae*. Front. Biosci. **7:**d1798–1814.
5. **Dagkessamanskaia, A., M. Moscoso, V. Henard, S. Guiral, K. Overweg, M. Reuter, B. Martin, J. Wells, and J. P. Claverys.** 2004. Interconnection of competence, stress and CiaR regulons in *Streptococcus pneumoniae*: competence triggers stationary phase autolysis of *ciaR* mutant cells. Mol. Microbiol. **51:**1071–1086.
6. **Draskovic, I., and D. Dubnau.** 2005. Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulphide bonds. Mol. Microbiol. **55:**881–896.
7. **Håvarstein, L. S., G. Coomaraswami, and D. A. Morrison.** 1995. An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae*. Proc. Natl. Acad. Sci. USA **92:**11140–11144.
8. **Håvarstein, L. S., P. Gaustad, I. F. Nes, and D. A. Morrison.** 1996. Identification of the streptococcal competence-pheromone receptor. Mol. Microbiol. **21:**863–869.
9. **Håvarstein, L. S., R. Hakenbeck, and P. Gaustad.** 1997. Natural competence in the genus *Streptococcus*: evidence that streptococci can change pherotype by interspecies recombinational exchanges. J. Bacteriol. **179:**6589–6594.
10. **Hols, P., F. Hancy, L. Fontaine, B. Grossiord, D. Prozzi, N. Leblond-Bourget, B. Decaris, A. Bolotin, C. Delorme, S. D. Ehrlich, E. Guédon, V. Monnet, P. Renault, and M. Kleerebezem.** 2005. New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed by comparative genomics. FEMS Microbiol. Rev. **29:**435–463.
11. **Hui, F. M., and D. A. Morrison.** 1991. Genetic transformation in *Streptococcus pneumoniae*: nucleotide sequence analysis shows *comA*, a gene required for competence induction, to be a member of the bacterial ATP-dependent transport protein family. J. Bacteriol. **173:**372–381.
12. **Johansen, E.** 1999. Genetic engineering (b). Modification of bacteria, p. 917–921. *In* R. Robinson, C. Batt, and P. Patel (ed.), Encyclopedia of food microbiology. Academic Press, London, United Kingdom.
13. **Lataste, H., J. P. Claverys, and A. M. Sicard.** 1981. Relation between the transforming activity of a marker and its proximity to the end of the DNA particle. Mol. Gen. Genet. **183:**199–201.
14. **Lee, M. S., and D. A. Morrison.** 1999. Identification of a new regulator in *Streptococcus pneumoniae* linking quorum sensing to competence for genetic transformation. J. Bacteriol. **181:**5004–5016.
15. **Luo, P., H. Li, and D. A. Morrison.** 2003. ComX is a unique link between multiple quorum sensing outputs and competence in *Streptococcus pneumoniae*. Mol. Microbiol. **50:**623–633.
16. **Luo, P., H. Li, and D. A. Morrison.** 2004. Identification of ComW as a new component in the regulation of genetic transformation in *Streptococcus pneumoniae*. Mol. Microbiol. **54:**172–183.
17. **Martin, B., Y. Quentin, G. Fichant, and J. P. Claverys.** 2006. Independent evolution of competence regulatory cascades in streptococci? Trends Microbiol. **14:**339–345.
18. **Monnet, C., S. Pernoud, A. Sepulchre, C. Fremaux, and G. Corrieu.** 2004. Selection and properties of *Streptococcus thermophilus* mutants deficient in urease. J. Dairy Sci. **87:**1634–1640.
19. **Opdyke, J. A., J. R. Scott, and C. P. Moran, Jr.** 2003. Expression of the secondary sigma factor σ^X in *Streptococcus pyogenes* is restricted at two levels. J. Bacteriol. **185:**4291–4297.
20. **O'Sullivan, D. J., and T. R. Klaenhammer.** 1993. High-copy-number and low-copy-number *Lactococcus* shuttle cloning vectors with features for clone screening. Gene **137:**227–231.
21. **Pestova, E. V., L. S. Håvarstein, and D. A. Morrison.** 1996. Regulation of competence for genetic transformation in *Streptococcus pneumoniae* by an auto-induced peptide pheromone and a two-component regulatory system. Mol. Microbiol. **21:**853–862.
22. **Peterson, S. N., C. K. Sung, R. Cline, B. V. Desai, E. C. Snesrud, P. Luo, J. Walling, H. Li, M. Mintz, G. Tsegaye, P. C. Burr, Y. Do, S. Ahn, J. Gilbert, R. D. Fleischmann, and D. A. Morrison.** 2004. Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. Mol. Microbiol. **51:**1051–1070.
23. **Podbielski, A., B. Spellerberg, M. Woischnik, B. Pohl, and R. Lütticken.** 1996. Novel series of plasmid vectors for gene inactivation and expression analysis in group A streptococci (GAS). Gene **177:**137–147.
24. **Sung, C. K., and D. A. Morrison.** 2005. Two distinct functions of ComW in stabilization and activation of the alternative sigma factor ComX in *Streptococcus pneumoniae*. J. Bacteriol. **187:**3052–3061.
25. **Tettelin, H.** 2004. Streptococcal genomes provide food for thought. Nat. Biotechnol. **22:**1523–1524.
26. **Vadeboncoeur, C., and S. Moineau.** 2004. The relevance of genetic analysis to dairy bacteria: building upon our heritage. Microb. Cell Fact. **3:**15–18.

Paper III

Letters in Applied Microbiology

Letters in Applied Microbiology ISSN 0266-8254

ORIGINAL ARTICLE

A food-grade site-directed mutagenesis system for *Streptococcus thermophilus* LMG 18311

T. Blomqvist, H. Steinmoen and L.S. Håvarstein

Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway

Keywords
biotechnology, dairy, fermentation and processes, lactic acid bacteria, Streptococci.

Correspondence
Leiv Sigve Håvarstein, Department of Chemistry, Biotechnology, and Food Science, Norwegian University of Life Sciences, PO Box 5003, N-1432 Ås, Norway. E-mail: sigve.havarstein@umb.no

2009/1792: received 12 October 2009, revised 15 December 2009 and accepted 17 December 2009

doi:10.1111/j.1472-765X.2009.02794.x

Abstract

Aims: To develop a general method for site-directed mutagenesis in the dairy starter strain *Streptococcus thermophilus* LMG 18311 which does not depend on antibiotic-resistance genes or other selection markers for the identification of transformants.

Methods and Results: In a previous study, we demonstrated that *Strep. thermophilus* LMG 18311 can be made competent for natural genetic transformation by overexpression of the alternative sigma factor ComX. In the present study, we wanted to investigate whether the natural transformation mechanism of *Strep. thermophilus* LMG 18311 is efficient enough to make it feasible to perform site-directed mutagenesis in this strain without the use of a selection marker. Competent bacteria were mixed with a DNA fragment engineered to contain a nonsense and a frameshift mutation in the middle of the target gene (*lacZ*) and subsequently seeded on agar plates. By performing colony-lift hybridization using a digoxigenin-labelled oligonucleotide probe, we succeeded in identifying transformants containing the sought after mutation.

Conclusions: By exploiting the natural transformability of *Strep. thermophilus* LMG 18311 and standard molecular methods, we have demonstrated that the genome of this bacterium can be altered at preselected sites without introduction of any foreign DNA.

Significance and Impact of the Study: A food-grade site-directed mutagenesis system has been developed for *Strep. thermophilus* LMG 18311 that can be used by the dairy industry to construct starter strains with novel and/or improved properties.

Introduction

Streptococcus thermophilus is a starter organism used by the dairy industry for the manufacture of yogurt and Italian and Swiss-type cheeses. It is the only species in the genus *Streptococcus* that has generally recognized as safe (GRAS) status (Delorme 2008). Because of consumer scepticism, no yogurts produced with genetically modified bacteria have so far been approved by The European Food Safety Authority (EFSA). In our opinion, it is likely that GMO strains that contain no foreign DNA, and only deviates from the parental strain at one or a few nucleotide positions, will be the first to be approved by consumers as well as EFSA. Based on this reasoning, we decided to investigate whether it is possible to carry out site-directed mutagenesis in the starter strain *Strep. thermophilus* LMG 18311 without the use of a selection marker. Our goal was to change only a few nucleotides in a preselected gene.

Most techniques used to engineer specific genetic changes into the genomes of target bacteria are very inefficient. Consequently, antibiotic-resistance genes, or other types of selection markers, are employed to isolate bacteria containing the desired mutation(s). However, when constructing genetically modified bacteria intended for use in the manufacture of food products, such markers should be avoided (Johansen 1999). We have previously shown that *Strep. thermophilus* LMG 18311 can be made

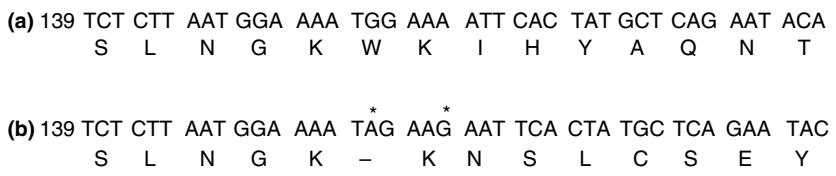

Figure 1 Sequence of mutated region of the *Streptococcus thermophilus lacZ* gene. (a) Wild-type sequence. (b) Mutated sequence. Asterisks indicate the positions of introduced nonsense and frameshift mutations.

competent for natural genetic transformation (Blomqvist *et al.* 2006). In the competent state, these bacteria take up external DNA and incorporate it into their genomes by homologous recombination. In the present study, we have investigated the efficiency of this process and found that mutants engineered to contain single base-pair mutations in their genomes can easily be identified by colony hybridization using a digoxigenin (DIG)-labelled oligonucleotide probe. Consequently, we have developed a general procedure for generating food-grade mutants in *Strep. thermophilus* LMG 18311 that can be used to introduce specific base-pair changes anywhere in the genome of this strain without leaving behind foreign DNA.

Materials and methods

Bacterial strains and growth conditions

Streptococcus thermophilus strain LMG 18311 harbouring the XL plasmid (pXL) (Blomqvist *et al.* 2006) was grown in Hogg-Jago glucose broth (HJGL) (3% tryptone; 1% yeast extract; 0·2% beef extract; 0·5% KH_2PO_4; 0·5% glucose; 0·5% lactose) or Todd-Hewitt broth (Difco Laboratories) supplemented with 0·8% glucose (THG). The XL plasmid was derived from the shuttle vector pTRKH2 (O'Sullivan and Klaenhammer 1993) as previously described (Blomqvist *et al.* 2006). Agar plates were prepared by adding 1·5% (w/v) agar to the media. For blue/white colony screening, 32 μl of 50 mg ml^{-1} X-gal dissolved in dimethylformamide was spread on each plate, prior to spreading the bacterial cells. All incubations were carried out at 37°C.

Construction of *lacZ* mutant

An approx. 1700-bp PCR fragment containing a nonsense mutation and a 1-bp frameshift insertion in the middle (see Fig. 1) was constructed using overlap extension PCR. In the first round, two separate PCRs were performed using the primer pairs LacZ-F1/LacZ-R1 and LacZ-F2/LacZ-R2, respectively, using an annealing temperature of 53°C and 30-s elongation time. Genomic DNA from *Strep. thermophilus* LMG 18311 was used as template. The two products from the first round were used in a second PCR together with the primers LacZ-F1 and Lac-R2 (annealing temperature = 55°C, 75-s elongation time), resulting in a spliced fragment containing the mutated region in the middle. This fragment was cloned into the pCR 2.1-TOPO vector (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions, giving rise to a plasmid termed pΔBG. The insert of this plasmid was sequenced using the primer Lac-sekv-F1 to confirm the introduced mutations. Prior to the transformation of the *Strep. thermophilus* LMG 18311 XL strain, the pΔBG was cleaved by *Nco*I to generate a linear fragment.

The PCRs described earlier were carried out using the Phusion high-fidelity DNA polymerase (Finnzymes, Espoo, Finland), and DNA sequencing was performed using BigDye 3.1 (Applied Biosystems, Foster City, CA). The sequences of the primers used are given in Table 1.

Transformation of *Streptococcus thermophilus* LMG 18311 XL

Streptococcus thermophilus LMG18311 cells, harbouring a plasmid (pXL) for the overexpression of *comX* (Blomqvist *et al.* 2006), were grown overnight at 37°C in THG supplemented with 2 μg ml^{-1} erythromycin. The next day, the culture was diluted to OD_{550} = 0·05 in the same medium prewarmed to 37°C. Then, 250 ng of synthetic *Streptococcus thermophilus* pheromone (STP) (NH2-SGWMDYINGFLKGFGGQRTLPTKDYNIPQV-COOH) was added to 1 ml of diluted culture, and the sample was placed in a water bath at 37°C. Two hours later, when OD_{550} of the culture was between 0·2 and 0·3, approx. 1 μg ml^{-1} of transforming DNA (linearized pΔBG) was added. The sample was then further incubated at 37°C

Table 1 Primers and probe used in this study

Primer/probe	Sequence
LacZ-F1	5′-GGACACCGTGATGAATCACTTA-3′
LacZ-R1	5′-TTCTGAGCATAGTGAATTCTTCTATTTTC-3′
LacZ-F2	5′-GGAAAATAGAAGAATTCACTATGCTCAGAA-3′
LacZ-R2	5′-ACTGGAACGACTTCAACAACTT-3′
LacZ-sekv-F1	5′-CACAGTACTTGTCACTGTAACTA-3′
LacZ-sekv-F2	5′-CCTCAAGTTCCTCAAGAATCA-3′
LacZ-sekv-F3	5′-CTATGGTAAGAGAATTGTCTT-3′
LacZ-sekv-F4	5′-GGACTTCATTGACCAAGCTATT-3′
LacZ-sekv-F6	5′-GGATCGTGCCAATAACATGTT-3′
LacZ-Probe-1	5′-CTTAATGGAAAATAGAAGAATTCACTATGC-DIG-3′

for 2 h. To separate long chains of cells into single cells or short chains, the culture was diluted to 4 ml using THG medium and treated with a mechanical blender (Ultra-Turrax model T25; Ika Labotechnik, Staufen, Germany) at maximum speed (24 000 rev min^{-1}) for 1 min. This process was repeated three times with cooling in between to avoid overheating. Following this treatment, the cells were allowed to recover at 37°C for 30 min before they were put on ice. Finally, the sample was serially diluted and spread on HJGL agar plates prepared with X-gal for blue/white screening, or without X-gal for colony hybridization. The plates were incubated over night at 37°C. The following morning, eight positive (white) transformants were selected from plates containing X-gal and grown in HJGL. The treatment in the mechanical blender was repeated, and the culture was spread on fresh agar plates containing X-gal. The presence of the correct mutations in the *lacZ* gene of the white transformants was verified by sequencing of the PCR product generated with primers LacZ-F1 and LacZ-R2. The oligonucleotide termed LacZ-sekv-F1 was used as a sequencing primer (Table 1). Curing of the pXL plasmid was performed by cultivating transformants in antibiotic-free medium for about 100 generations.

Colony hybridization with DIG-labelled probe

Colonies were transferred to nylon membranes (Hybond-N+; Amersham Bioscience, Little Chalfont, UK) according to the manufacturer's instructions. Subsequent treatment of the membranes (lysis, denaturation, neutralization and washing) was performed on solution-saturated 3MM paper filters (Whatman, Maidstone, UK). Lysis of bacterial cells was accomplished by incubating the membranes for 45 min at 37°C in a 10 mmol l^{-1} Tris–HCl buffer (pH 7·5) containing 250 mmol l^{-1} sucrose and 5 mg ml^{-1} of lysozyme. The next steps were performed at room temperature. DNA was denatured for 15 min in 1·5 mol l^{-1} NaCl/1 mol l^{-1} NaOH and thereafter neutralized for 4 min in a 0·5 mol l^{-1} Tris–HCl buffer (pH 8) containing 1·5 mol l^{-1} NaCl. Finally, the membranes were washed for 4 min in 2× SSC (0·3 mol l^{-1} NaCl, 0·03 mol l^{-1} sodium citrate) followed by baking at 80°C for 2 h.

Prior to the hybridization, the membranes were incubated with gentle agitation at 45°C in Easy Hyb for 30 min (see DIG Easy Hyb protocol from Roche for details). Then, the prehybridization solution was removed, and a smaller volume of fresh Easy Hyb solution containing 10 pmol ml^{-1} LacZ-Probe-1 (Table 1) was added. Membranes were hybridized overnight at 45°C. Posthybridization washes were performed 2 × 5 min in 2× SSC using 0·1% SDS at room temperature, and 2 × 15 min in 0·5× SSC using 0·1% SDS, at 50°C under constant agitation. Membranes were developed for 4–5 h using the DIG Nucleic Acid Detection kit (Roche Diagnostics GmbH, Mannheim, Germany) as described by the supplier. To confirm that positive colonies contained bacteria with the desired mutations, they were subjected to blue/white screening and sequencing of the *lacZ* gene. The sequencing reactions were performed using the primers LacZ-R2, LacZ-sekv-F1, -2, -3, -4 and -6 on a PCR fragment generated from genomic DNA with the primers LacZ-F1 and LacZ-R2.

Results

A DNA fragment containing an engineered stop codon (codon 52) and 1-bp frameshift insertion (codon 53) in the *Strep. thermophilus lacZ* gene (Fig. 1), plus approx. 1000-bp upstream and downstream flanking regions, was made by overlap extension PCR and cloned into the pCR2.1-TOPO plasmid as described in the Materials and Methods section. Subsequently, the resulting construct was linearized by cleaving a unique *Nco*I site in the TOPO plasmid. The purpose of the flanking nonhomologous plasmid sequences was to increase the transformation efficiency by protecting the homologous donor DNA from degradation by nucleases during transformation. Following the uptake of the construct by competent *Strep. thermophilus* LMG 18311 cells, targeted integration of the mutated region into the bacterial genome by double-crossover homologous recombination will take place in a minor fraction of the population. To screen for such $LacZ^-$ transformants, bacteria were seeded on agar plates containing X-gal substrate. The simplicity of this procedure allowed us to screen a large number of bacteria to determine the ratio of white to blue colonies. A low ratio, indicating poor transformation efficiency, would imply that the sought after mutant would be hard to find with a more general method such as colony-lift hybridization. Interestingly, our results showed that several white colonies ($LacZ^-$) were present on each agar plate (see Fig. 2). We estimated that 0·7% of the colonies were white, demonstrating a surprisingly high transformation efficiency. We then repeated the experiment but used colony-lift hybridization with a 30-bp DIG-labelled probe instead of blue/white screening to identify $LacZ^-$ mutants. Twelve colonies believed to be positive were selected and streaked out on new agar plates containing X-gal. Seven of these gave rise to only white colonies, two gave blue colonies, whereas three gave a mixture of blue and white colonies (Fig 3c,d,e).

The fact that *Strep. thermophilus* grow in long chains makes it impossible to know whether a colony is formed from one cell or a chain of cells. This represents a complicating factor when mutagenesis of the type described in this work is carried out. As no selectable marker is used

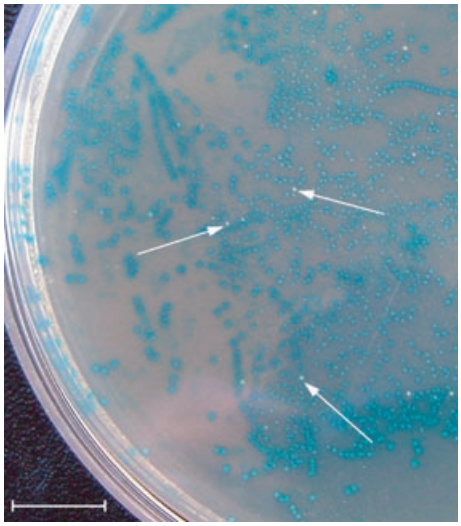

Figure 2 Estimation of transformation efficiency. After exposing a competent *Streptococcus thermophilus* LMG18311 culture to DNA containing a *lacZ* gene carrying a nonsense and a frameshift mutation in the middle, cells were spread on agar plates containing X-gal. Desired *lacZ*$^-$ mutants appear as white colonies as indicated by arrows. The scale bar represents 1 cm.

that allows growth only of the transformed cells, colonies that contains cells carrying the correct mutation(s) will also in general contain a large number of wild-type cells. We, therefore, used a mechanical blender to break the chains before the bacteria were seeded on agar plates. Microscopic inspection of cultures treated with the blender revealed that it contained single cells or chains consisting of two to *c.* five cells, while the untreated culture contained chains ranging from two to *c.* 40 cells in length (Fig. 3a,b). As mentioned earlier, some of the positive colonies identified by colony-lift hybridization turned out to consist of a mixture of LacZ$^-$ and LacZ$^+$ bacteria (see Fig. 3e). To get rid of the contaminating wild-type bacteria, treatment with the blender followed by a new round of colony-lift hybridization might be necessary. This procedure will have to be repeated until only positive colonies appear on the hybridization filters after colour development.

Discussion

The genetic engineering technique described in this work makes it is possible to introduce changes into the genome of the dairy starter strain *Strep. thermophilus* LMG 18311 without having to rely on a selection marker. The technique allows targeted substitution, insertion or deletion of a single or a few basepairs, changes that over time occur naturally but randomly in the genomes of all bacteria. In this connection, it is worth noticing that the genomes of the two commonly used *Strep. thermophilus* starter

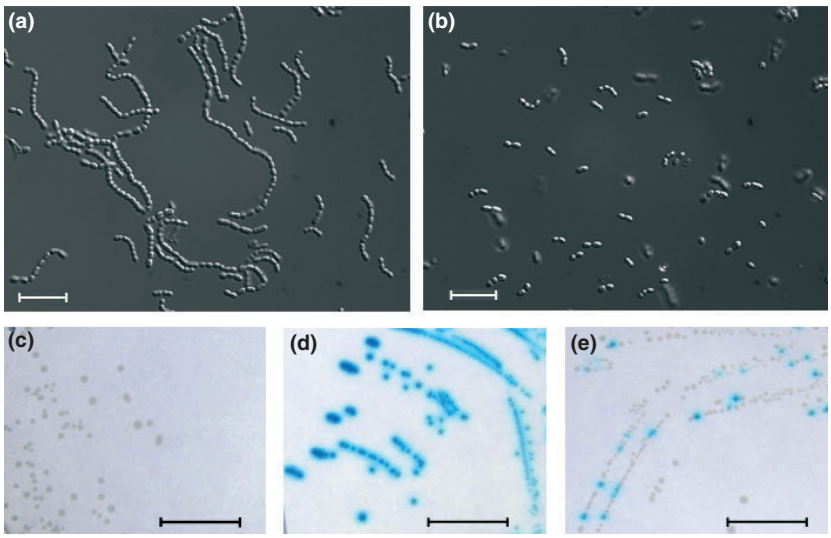

Figure 3 Mechanical breaking of bacterial chains facilitates the identification and isolation of mutants. Upper panels: microscopic images of *Streptococcus thermophilus* LMG 18311 before (a) and after (b) treatment with the Ultra-Turrax mechanical blender. Lower panel: to determine how well the mechanical separation technique worked in practice, colony-lift hybridization with a DIG-labelled oligonucleotide probe was used to identify colonies containing a mutated nonfunctional *lacZ* gene (see text for details). Colonies identified as positive with the DIG probe were rescreened on agar plates containing X-gal. Three types of results were obtained: only mutants (white colonies, panel c), mostly wild-type cells (blue colonies, panel d) and a mixture of mutant and wild type cells (panel e). This result demonstrates that the mechanical separation technique works, but that rescreening is necessary to ensure that positive colonies are not contaminated with wild-type cells. The scale bars in panels (a) and (b) represent 10 μm, while the scale bars in panels (c), (d) and (e) represent 1 cm.

strains, CNRZ1061 and LMG 18311, contain around 3000 single nucleotide differences (Bolotin *et al.* 2004; Hols *et al.* 2005).

Inactivation of gene products by introduction of stop codons or frameshift mutations can be used to construct starter strains with novel or improved properties. A concrete example of a loss-of-function mutant that could be useful for the dairy industry is a urease-negative (Ure^-) strain. A major role for *Strep. thermophilus* in milk fermentations is to provide rapid acidification. However, in urease-positive strains, hydrolysis of urea, which can be present in milk at varying concentrations depending on the type of feed provided, releases ammonia that slows down the rate of acidification. Possible consequences of this are increased fermentation time, altered texture characteristics of the product and/or growth of contaminating micro-organisms (Martin *et al.* 1997; Mora *et al.* 2004).

About 10% of the open reading frames present in the genome of *Strep. thermophilus* LMG 18311 have been reported to be pseudogenes (Bolotin *et al.* 2004). The LMG18311 strain has, for instance, a nonfunctional bacteriocin system because of a disrupted BlpB accessory protein, which together with the ABC-transporter BlpA transport bacteriocins possessing double-glycine type leaders to the exterior of the bacterial cell (Hols *et al.* 2005; Fontaine *et al.* 2007). A nonsense mutation and a 1-bp insertion have rendered the *blpB* gene nonfunctional. Using the technique described in the current study, it is possible to repair pseudogenes in LMG18311 that have become nonfunctional because of nonsense or frameshift mutations. In the case of yogurt starter strain LMG18311, a restored bacteriocin system might constitute an additional hurdle to control the growth of nonstarter strains or foodborne pathogenic species such as *Listeria monocytogenes* (Fontaine and Hols 2008).

The molecular tools described in this study can also be used to manipulate the expression level of genes and gene products. This can, for instance, be performed by introducing specific point mutations that alters promoter strength or the efficiency of ribosome-binding sites. One area of particular interest for the dairy industry is to develop novel *Strep. thermophilus* starter strains that produce more of aroma compounds such as acetaldehyde and diacetyl. In theory, this can be performed by manipulating the expression level of certain key enzymes in the carbon metabolism (Hols *et al.* 2005). It has been reported, for instance, that increased production of α-acetolactate synthase, which is involved in conversion of pyruvate into diacetyl, stimulates diacetyl production. Similarly, inactivation of α-acetolactate decarboxylase, an enzyme that coverts α-acetolactate into acetoin, boosts the level of diacetyl significantly (Hols *et al.* 2005; Monnet and Corrieu 2007). Presumably, construction of a strain with high levels of α-acetolactate synthase combined with low levels of α-acetolactate decarboxylase would give an even better result. Another area of major importance to the manufacturers of yoghurt is the production of extracellular polysaccharides (eps), which contribute to mouthfeel, taste perception and texture of the product. Rådström and co-workers have demonstrated that higher eps production can be obtained by increasing the levels of α-phosphoglucomutase, UDP-glucose pyrophosphorylase and the Leloir enzymes, and that the resulting increase in eps production influenced the rheological properties of fermented milk (Levander *et al.* 2002; Svensson *et al.* 2005).

The examples given earlier illustrate some possible applications for the genetic engineering tools described in the present study. Because of the efficiency of the transformation process, small genetic changes can be introduced into the genome of *Strep. thermophilus* LMG 18311 without the use of a selection marker. Furthermore, the desired phenotype can, in many cases, be obtained by substitution of only a single or a few base pairs. It is difficult to envision how the use of this technique can give rise to food products that represent a hazard to the consumer or the environment.

Acknowledgement

This work was supported by the Research Council of Norway.

References

Blomqvist, T., Steinmoen, H. and Håvarstein, L.S. (2006) Natural genetic transformation: a novel tool for efficient genetic engineering of the dairy bacterium *Streptococcus thermophilus*. *Appl Environ Microbiol* **72**, 6751–6756.

Bolotin, A., Quinquis, B., Renault, P., Sorokin, A., Ehrlich, S.D., Kulakauskas, S., Lapidus, A., Goltsman, E. *et al.* (2004) Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat Biotechnol* **22**, 1554–1558.

Delorme, C. (2008) Safety assessment of dairy microorganisms: *Streptococcus thermophilus*. *Int J Food Microbiol* **126**, 274–277.

Fontaine, L. and Hols, P. (2008) The inhibitory spectrum of thermophilin 9 from *Streptococcus thermophilus* LMD-9 depends on the production of multiple peptides and the activity of $BlpG_{St}$, a thiol-disulfide oxidase. *Appl Environ Microbiol* **74**, 1102–1110.

Fontaine, L., Boutry, C., Guédon, E., Guillot, A., Ibrahim, M., Grossiord, B. and Hols, P. (2007) Quorum-sensing regulation of the production of Blp bacteriocins in *Streptococcus thermophilus*. *J Bacteriol* **189**, 7195–7205.

Hols, P., Hancy, F., Fontaine, L., Grossiord, B., Prozzi, D., Leblond-Bourget, N., Decaris, B., Bolotin, A. *et al.* (2005) New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed by comparative genomics. *FEMS Microbiol Rev* **29**, 435–463.

Johansen, E. (1999) Modification of bacteria. In *Encyclopedia of Food Microbiology* ed. Robinson, R., Batt, C. and Patel, P. pp. 917–921. London: Academic Press.

Levander, F., Svensson, M. and Rådström, P. (2002) Enhanced exopolysaccharide production by metabolic engineering of *Streptococcus thermophilus*. *Appl Environ Microbiol* **68**, 784–790.

Martin, B., Coulon, J.B., Chamba, J.F. and Bugaud, C. (1997) Effect of milk urea content on characteristics of matured Reblochon cheeses. *Lait* **77**, 505–514.

Monnet, C. and Corrieu, G. (2007) Selection and properties of α-acetolactate decarboxylase-deficient spontaneous mutants of *Streptococcus thermophilus*. *Food Microbiol* **24**, 601–606.

Mora, D., Maguin, E., Masiero, M., Parini, C., Ricci, G., Manachini, P.L. and Daffonchio, D. (2004) Characterization of urease genes cluster of *Streptococcus thermophilus*. *J Appl Microbiol* **96**, 209–219.

O'Sullivan, D.J. and Klaenhammer, T.R. (1993) High-copy-number and low-copy-number *Lactococcus* shuttle cloning vectors with features for clone screening. *Gene* **137**, 227–231.

Svensson, M., Waak, E., Svensson, U. and Rådström, P. (2005) Metabolically improved exopolysaccharide production by *Streptococcus thermophilus* and its influence on the rheological properties of fermented milk. *Appl Environ Microbiol* **71**, 6398–6400.

