Characterization of *Streptococcus gordonii* prophage PH15: complete genome sequence and functional analysis of phage-encoded integrase and endolysin

Jan R. van der Ploeg

Institute of Oral Biology, University of Zürich, Plattenstrasse 11, 8032 Zürich, Switzerland

Correspondence

Jan R. van der Ploeg

jan.vanderploeg@zzmk.uzh.ch

Received 1 April 2008

Revised 1 July 2008

Accepted 2 July 2008

Streptococcus gordonii OMZ1039, isolated from supragingival dental plaque, was found to harbour a prophage, PH15, whose excision could be induced by mitomycin treatment. Phage PH15 belongs to the *Siphoviridae*. The complete genome sequence of PH15 was determined. The genome was 39 136 bp in size and contained 61 ORFs. The genome of PH15 was most similar in the structural module to the temperate bacteriophages MM1 and ϕ NIH1.1 from *Streptococcus pneumoniae* and *Streptococcus pyogenes*, respectively. In strain OMZ1039, PH15 was found to reside as a prophage in the cysteinyl-tRNA gene. A plasmid, harbouring the *attP* site and the integrase gene downstream of a constitutive promoter, was capable of site-specific integration into the genomes of different oral streptococcal species. The phage endolysin was purified after expression in *Escherichia coli* and found to inhibit growth of all *S. gordonii* strains tested and several different streptococcal species, including the pathogens *Streptococcus mutans*, *S. pyogenes* and *Streptococcus agalactiae*.

INTRODUCTION

In healthy individuals, mitis streptococci make up a large fraction of the total bacterial microbiota of the oral cavity. One of the members of the mitis streptococci, *Streptococcus gordonii*, is among the first colonizers of teeth. Compared to *Streptococcus mutans*, the principal aetiological agent of dental caries, *S. gordonii* has low cariogenic potential (Tanzer *et al.*, 2008). On the contrary, it has been suggested that *S. gordonii* can modulate the virulence properties of *S. mutans* by interfering with the *comCDE* quorum-sensing system, which regulates biofilm formation, bacteriocin production and genetic competence (Wang & Kuramitsu, 2005). Because of its pioneer colonizing properties and because of its potential to elicit immune responses in infants and children, it has been proposed to use

engineered strains of *S. gordonii* for live vaccine delivery after mucosal colonization (Fischetti *et al.*, 1996; Lee, 2003). Ideally, such antigens would be exported or displayed on the cell surface of *S. gordonii* (Lee, 2003). Although *S. gordonii* belongs to the GRAS (generally regarded as safe) organisms, it has been implicated in causing infective endocarditis (Douglas *et al.*, 1993). This colonization of damaged heart valves can occur when *S. gordonii* or other viridans streptococci enter the bloodstream after dental procedures.

It is becoming increasingly apparent that temperate bacteriophages play an important role in bacterial genome diversification, in particular in the dissemination of virulence factors. The availability of bacterial genome sequences has contributed considerably to this understanding. Prophages have been detected in the majority of the sequenced genomes (Canchaya *et al.*, 2003). Some *Streptococcus pyogenes* and *Streptococcus agalactiae* strains have been found to contain multiple prophages or prophage remnants (Canchaya *et al.*, 2003). It has been suggested that epidemic outbreaks caused by *S. pyogenes* are related to the acquisition of prophages, and hence phage-encoded virulence factors (Banks *et al.*, 2002). The function of prophages in contributing to the virulence of *Streptococcus pneumoniae* is less evident, but it has recently been reported that adherence of *S. pneumoniae* to inert surfaces and pharyngeal cells is increased after lysogenization with phage MM1 (Loeffler & Fischetti, 2006).

Abbreviation: CRISPR, clustered regularly interspaced palindromic repeats.

The GenBank/EMBL/DBJ accession number for the complete genome sequence of *Streptococcus gordonii* prophage PH15 is FM163528.

Two supplementary tables, showing the oligonucleotides used in this study and the ORFs encoded by *S. gordonii* prophage PH15, and three supplementary figures, showing phylogenetic analysis of part of *sodA* from strains OMZ1038, OMZ1039 and OMZ1081, electron microscopy of phage PH15 adsorption to *S. gordonii* and *S. oralis*, and a comparison of PH15 ORFs with proteins encoded by streptococcal phages, are available with the online version of this paper.

Although the role played by oral streptococci in the ecology of the oral cavity has been studied to some extent (Kuramitsu *et al.*, 2007), little is known about the presence and impact of bacteriophages from these organisms. The genome sequence of the virulent phage M102, which infects *S. mutans*, has recently been published (van der Ploeg, 2007). The sequenced genomes of three oral streptococcal strains, including *S. gordonii* CH1, do not appear to contain prophages (Ajdić *et al.*, 2002; Tanzer *et al.*, 2008; Vickerman *et al.*, 2007; Xu *et al.*, 2007). Temperate phages from *Streptococcus mitis*, an abundant inhabitant of the oral cavity, have been described (Romero *et al.*, 2004a; Siboo *et al.*, 2003). *S. mitis* phage SM1 has been found to encode two proteins, PblA and PblB, which mediate binding to platelets and affect virulence in a rat model of infective endocarditis (Mitchell *et al.*, 2007; Siboo *et al.*, 2003).

Here, the isolation and the genome sequence of what is believed to be the first temperate bacteriophage from a fresh isolate of *S. gordonii* and the functional analysis of the phage-encoded endolysin and integrase are reported.

METHODS

Bacterial strains and growth conditions. *Streptococcus*, *Lactococcus lactis* and *Enterococcus faecalis* strains are listed in Table 1 and were routinely grown without agitation in Todd–Hewitt broth supplemented with 0.3 % (w/v) yeast extract (THY) at 37 °C under aerobic conditions or in a mixture of 10 % CO₂ and 90 % air. When required, antibiotics were used at the following final concentrations: erythromycin, 15 µg ml⁻¹; kanamycin, 750 µg ml⁻¹. Streptococci from human supragingival plaque were isolated on mitis salivarius agar (Difco). *Escherichia coli* strains JM109 and BL21(DE3) were used for routine plasmid propagation and protein production, respectively, and grown in Luria–Bertani (LB) medium. When required, antibiotics were added at the following final concentrations: erythromycin, 200 µg ml⁻¹; kanamycin, 50 µg ml⁻¹.

DNA manipulations. Oligonucleotides used in this study are listed in Supplementary Table S1. Plasmid purification from *E. coli*, restriction enzyme digestion, agarose gel electrophoresis, DNA ligation and transformation of *E. coli* were carried out using standard methods (Ausubel *et al.*, 1987). Chromosomal DNA was isolated by using the GenElute Bacterial Genomic DNA kit (Sigma). DNA sequencing employed dye terminators and was performed on an ABI 3730xl DNA analyser (Applied Biosystems).

For strain identification, part of the 16S rRNA gene was amplified by using PCR with the universal primers 27f and 1492r. The PCR product was sequenced using primer 519r. For sequencing of *sodA*, PCR was performed with primers *sodAf* and *sodAr*, and the product was sequenced with both primers.

Transformation of streptococci was done as follows. An overnight culture of the strain of interest was diluted 40-fold in fresh THY medium and incubated for 3–4 h at 37 °C. In some cases, heat-inactivated horse serum was added to a final concentration of 5 % (v/v). A suitable amount of plasmid (0.1–1.0 µg) was added to 0.5 ml of the culture, which was further incubated for 1 h at 37 °C. Appropriate dilutions were plated on THY medium containing erythromycin. Plates were incubated for 48 h in 10 % CO₂/90 % air at 37 °C.

Phage propagation, isolation of phage DNA and sequencing strategy. For isolation of phage DNA from OMZ1039, an overnight culture was diluted 100-fold in fresh THY medium. After 3 h of incubation at 37 °C, mitomycin was added to a final concentration of 0.1 µg ml⁻¹. The culture was incubated for another 20 h at 37 °C. To remove debris, the lysate (designated as phage lysate below) was centrifuged for 10 min at 7000 g and passed through a 0.4 µm pore-size filter. Phage DNA was isolated from the lysate by using the LambdaPrep kit (Promega). Digestion of phage DNA with *EcoRV* yielded 11 fragments with a total size of about 38 kb. Nine of these *EcoRV* restriction fragments (with a total size of about 28 kb) were cloned in the vector pSMART-Lc-Kan (Lucigen). The inserts of the resulting plasmids were completely sequenced by using a primer-walking strategy. In addition, a shotgun library was prepared by cloning fragments of ~1.5 kb, obtained by hydro-shearing of phage DNA, in pGEM-T-Easy (MWG-Biotech). These plasmids were end-sequenced. Sequences were assembled with the program CAP3 (Huang & Madan, 1999). Direct sequencing of PCR fragments obtained with custom primers was utilized to confirm junction sequences across the ends of cloned fragments and to fill gaps.

Sequence analysis. The assembled sequence of PH15 was analysed for the presence of ORFs with the program Genemark.hmm (Lukashin & Borodovsky, 1998), with the codon usage of *S. pneumoniae* strain R6 and a minimum size of 42 bases. Further sequence analysis used the programs from the GCG package (Accelrys). BLAST sequence similarity searches were done at <http://www.ncbi.nlm.nih.gov/blast/>. For multiple sequence alignments, CLUSTAL W (<http://www.ebi.ac.uk/clustalw/>) was used. Analysis of CRISPR (clustered regularly interspaced palindromic repeats) sequences in the genome of *S. gordonii* was done at <http://crispr.u-psud.fr/>. Phylogenetic analysis of *sodA* sequences was performed at <http://www.phylogeny.fr/>, using the one-click mode (Dereeper *et al.*, 2008).

Construction of streptococcal integration vector. The integrase gene and *attP* from PH15 were amplified by PCR using primers PH15-int1 and PH15-int2. The product was digested with *SacI* and *EcoRI* and cloned into plasmid pNE1 (Bartilson *et al.*, 2001) to give pPH15int. A fragment harbouring a constitutive promoter (*cp32*) (Jensen & Hammer, 1998) was generated by annealing of two complementary primers (*cp32f* and *cp32r*) and cloned upstream of the *int* gene in the *Bam*HI/*Sma*I sites of pPH15int to give pcPH15int. Subsequently, the *Bam*HI/*Eco*RI fragments from pPH15int and pcPH15int were cloned into similarly digested plasmid pVA891 (Macrina *et al.*, 1983) to make plasmids pOMZ292 and pOMZ291, respectively. A plasmid with a truncated integrase gene was constructed by digestion of pOMZ291 with *Bgl*II, filling in the ends with Klenow polymerase and religation, resulting in plasmid pOMZ308.

Cloning, expression and activity of the lytic enzyme of PH15. The *lysA* gene from PH15 was amplified by PCR using primers PH15lysf and PH15lysr and DNA from strain OMZ1039 as template. The product was digested with *Nde*I and *Xho*I and cloned into similarly digested pET28b (Novagen) to give plasmid pOMZ276. For expression of *lysA*, a culture of *E. coli* BL21(DE3) (Novagen) harbouring plasmid pOMZ276 was grown at 37 °C to OD₆₀₀ 0.8, upon which IPTG was added to 0.1 mM. Incubation of the culture at 37 °C was continued for 5 h. Cells were harvested, resuspended in 1× Bugbuster (Novagen) and incubated with shaking for 1 h at room temperature. The extract was centrifuged for 5 min at 16000 g and the supernatant purified by Ni-NTA agarose affinity chromatography as recommended by the manufacturer (Qiagen), using 250 mM imidazole to elute the protein. The elution buffer was exchanged with 20 mM sodium phosphate buffer, pH 7.1, by using an Amicon Ultra-4 centrifugal filter device

Table 1. Micro-organisms used in this study and their lysis by purified endolysin from PH15

Organism	Strain number	Alternative name, characteristics	Source or reference	Lysis
<i>S. gordonii</i>	OMZ1039	Harbours prophage PH15	This study	+
<i>S. gordonii</i>	OMZ504	ATCC 12396 (SK6)	Kilian <i>et al.</i> (1989)	+
<i>S. gordonii</i>	OMZ505	ATCC 10558T (SK3)	Kilian <i>et al.</i> (1989)	+
<i>S. gordonii</i>	OMZ606	NCTC 3165 (SK51)	Kilian <i>et al.</i> (1989)	+
<i>S. gordonii</i>	OMZ938	Challis DL1*	H. Jenkinson†	+
<i>S. oralis</i>	OMZ8	ATCC 9811		–
<i>S. oralis</i>	OMZ507	SK2 (NCTC 7864)	Kilian <i>et al.</i> (1989)	–
<i>S. oralis</i>	OMZ607	SK248	M. Kilian‡	–
<i>S. oralis</i>	OMZ1038	Harbours prophage PH10	This study	–
<i>S. oralis</i>	OMZ1081	Harbours prophage PH5	This study	ND§
<i>S. mitis</i>	OMZ508	SK24 (NCTC 8029)	Kilian <i>et al.</i> (1989)	–
<i>S. mitis</i>	OMZ991	DSMZ12643 (NCTC 12261)	DSMZ	–
<i>S. mitis</i>	OMZ1043	BMS	H. Busscher	–
<i>S. mitis</i>	OMZ1049	BA	H. Busscher	–
<i>S. pneumoniae</i>	OMZ810	Wu2	B. Diamond¶	–
<i>S. sanguinis</i>	OMZ499	ATCC 10556	Kilian <i>et al.</i> (1989)	+
<i>S. sanguinis</i>	OMZ500	SK36	Kilian <i>et al.</i> (1989)	+
<i>S. mutans</i>	OMZ381	P42	G. Tiraby#	+
<i>S. mutans</i>	OMZ918	UA159 (ATCC 700610)	ATCC	ND§
<i>S. mutans</i>	OMZ1001	Derivative of UA159	van der Ploeg (2005)	+
<i>S. mutans</i>	OMZ62	AHT		+
<i>S. mutans</i>	OMZ125	LM7		+
<i>S. mutans</i>	OMZ175			+
<i>Streptococcus cricetus</i>	OMZ165	HS6		+
<i>Streptococcus sobrinus</i>	OMZ176			–
<i>Streptococcus anginosus</i>	OMZ510	NCTC 10713	Kilian <i>et al.</i> (1989)	+
<i>S. agalactiae</i>	OMZ1092		R. Zbinden**	+
<i>S. pyogenes</i>	OMZ1091		R. Zbinden**	+
<i>Streptococcus constellatus</i>	OMZ511	ATCC 27823	Kilian <i>et al.</i> (1989)	–
<i>Streptococcus intermedius</i>	OMZ512	ATCC 27335	Kilian <i>et al.</i> (1989)	+
<i>Streptococcus salivarius</i>	OMZ513	NCTC 8618	Kilian <i>et al.</i> (1989)	+
<i>S. thermophilus</i>	OMZ1082	DSM 20617	M. Loessner††	+
<i>Ent. faecalis</i>	OMZ940	GC110	D. Clewell‡‡	+
<i>L. lactis</i>	OMZ1041	DSM 20729	DSMZ	–
<i>L. lactis</i>	OMZ1045	NZ9800	M. Kleerebezem§§	–

**S. gordonii* Challis DL1 originates from the same ancestor as strain CH1, whose genome has been sequenced, but the strains have apparently diverged during maintenance in different laboratories (Howard Jenkinson, personal communication).

†University of Bristol.

‡Århus University.

§ND, Not done.

||University of Groningen.

¶Albert Einstein College of Medicine of Yeshiva University, New York.

#Université Paul Sabatier, Toulouse.

**University of Zurich.

††Swiss Federal Institute of Technology Zurich.

‡‡University of Michigan.

§§NIZO Food Research.

(cut-off 10 000 Da). Cell wall lytic activity of the purified protein with different bacterial strains was evaluated as follows. One hundred microlitres of an overnight culture was mixed with 4 ml THY soft agar and poured on a THY plate. After the soft agar had hardened, 5 µl enzyme was spotted on the plate. The plate was incubated for 16 h at 37 °C in 10 % CO₂ and 90 % air and examined for clear zones indicative of growth inhibition.

Electron microscopy. Grids were prepared by applying drops of phage lysates (prepared as described above) to Formvar-coated, carbon-reinforced copper grids for about 45 s, then excess solution was removed with filter paper. Subsequently, a drop of 1 % phosphotungstic acid (PTA), pH 4.2, was applied for ~15 s and similarly removed with filter paper. The grids were allowed to dry and were examined under a Philips EM 400T electron microscope at

60 kV. Images were taken at magnifications of $\times 33\,000$, $\times 42\,000$ and $\times 54\,500$.

To visualize phages adhering to bacteria, 5 ml of exponentially grown cultures of *S. gordonii* OMZ1039 and *Streptococcus oralis* OMZ607 were incubated with 5 ml of phage lysate PH15 for 30 min at 37 °C. Glutaraldehyde was added to a final concentration of 2.5 % for 30 min at room temperature. Fixed cells were harvested by centrifugation, washed three times with distilled water and resuspended to a final volume of 1 ml. Electron microscopy was as described above.

RESULTS AND DISCUSSION

Presence of temperate phages in fresh streptococcal supragingival plaque isolates

It was investigated whether cultures of fresh isolates from human supragingival plaque harboured prophages. Out of 15 isolates tested, 10 lysed following addition of mitomycin. The lysates from these 10 isolates were examined for the presence of bacteriophages by electron microscopy. In five lysates, bacteriophages were found to be present, suggesting that these isolates harboured a prophage. Analysis of the sequence of the 16S rRNA gene showed that four isolates belonged to the mitis group of streptococci; the fifth isolate was identified as *Staphylococcus epidermidis*. Two of the sequences were identical and might represent the same isolate. Therefore, investigations were continued with the three streptococcal isolates, which were designated OMZ1038, OMZ1039 and OMZ1081. Since identification of mitis streptococci on the basis of 16S rRNA sequences at the species level is not reliable (Kawamura *et al.*, 1999), part of the *sodA* gene was sequenced (Poyart *et al.*, 1998). Strains OMZ1038 and OMZ1081 were both assigned as *S. oralis*, whereas OMZ1039 was identified as *S. gordonii* (Supplementary Fig. S1). In this study, the temperate phage (designated PH15) from OMZ1039 was characterized.

Phage PH15 did not form plaques with *S. gordonii* strains OMZ505 and OMZ938, which suggests either that the host range is narrow or that the phage is defective. Electron microscopic analysis of PH15 showed that it belongs to the *Siphoviridae*, having a long, presumably non-contractile tail of 231 ± 39 nm and an icosahedral head with a diameter of 59.7 ± 5.4 nm (Fig. 1). Incubation of PH15 with bacteria indicated that the phage adhered to OMZ1039 (Supplementary Fig. S2a), but not to *S. oralis* strain OMZ607 (Supplementary Fig. S2b), which suggests that phage adherence is specific.

Genome sequence of PH15

The complete genome sequence of PH15 was determined to be 39 136 bp in size, with a GC percentage of 40.51 %, which is identical to the GC percentage of the genome sequence of *S. gordonii* Challis CH1 (accession no. CP000725).

Fig. 1. Electron microscopic analysis of phage PH15, showing the long, presumably non-contractile tail and icosahedral head. Two different micrographs are shown to give a better impression of the head and tail of the phage.

A total of 61 ORFs were encoded on the phage genome. They are listed in Supplementary Table S2, and a sequence comparison of the ORFs from PH15 with proteins from other streptococcal phages is summarized in Supplementary Fig. S3. Three of the ORFs had a GTG start codon, five a TTG start codon and the rest an ATG start codon. The transcription of most ORFs was directed to the right, but six ORFs, all located near *attL* (Fig. 2), were oriented in the opposite direction. There were four sequences with 100 % identity to the *E. coli* consensus promoter (ttgaca-n17-tataat). These were located just upstream of ORF5, ORF9, ORF56 and ORF59.

According to their putative functions and assuming a genetic organization similar to that in other bacteriophages from low-GC-content Gram-positive organisms (Canchaya *et al.*, 2003), the ORFs were assigned to different functional groups, i.e. lysogeny, replication and regulation, packaging, morphogenesis and host lysis (Fig. 2).

The lysogeny module is located on the most leftward part of the genome. The module consists of ORFs, the order of which follows the conserved pattern found in *Siphoviridae* from low-GC-content Gram-positive bacteria (Lucchini *et al.*, 1999). ORF1 showed similarity to numerous putative integrases from low-GC-content Gram-positive streptococci, lactococci, lactobacilli and bacilli. The highest similarity (78 % identity) was found with the putative integrase from bacteriophage LambdaSa2, which has been identified as a prophage in *S. agalactiae* strain 2603 V/R by genome sequencing (Tettelin *et al.*, 2002). One of the few similar proteins with demonstrated activity is the integrase

Fig. 2. Organization of the *S. gordonii* phage PH15 genome and comparison with the genome from *S. pneumoniae* phage MM1. ORFs with similarity to proteins encoded by bacteriophage MM1 (NC_003050; Obregón *et al.*, 2003) are indicated by blocks. The percentage amino acid identity is represented by the colour of each block (light yellow, 20–39 % identity; dark yellow, 40–59 % identity; brown, 60–79 % identity). The colour code for the different modules was adopted from Obregón *et al.* (2003). Lysogeny and lysogenic conversion modules are shown enlarged. Putative promoters are indicated by hooked arrows, putative transcriptional terminators by Ω . The upper and lower scale bars indicate the size in bp.

from *L. lactis* phage TPW22 (Petersen *et al.*, 1999). ORF1 appears to be preceded by its own promoter and separated from ORF2 by an inverted repeat (free energy of 14.6 kcal mol⁻¹; 61.1 kJ mol⁻¹). This suggests that ORF1 is transcribed independently of the genes upstream.

ORF2 and ORF3 did not show sequence similarity to genes of known function. ORF4 was similar to cI-like repressors from lactococcal phages LC3 (Blatny *et al.*, 2003) and Tuc2009 (Kenny *et al.*, 2006). As in most streptococcal and lactococcal phage genomes, the genetic switch for lysogeny/lysis is most likely located upstream of the gene encoding the cI-like repressor. This region contained two putative operators for the cI-like repressor which overlap with the two putative promoters for cI and ORF5. ORF5 was preceded by an *E. coli* σ^{70} promoter sequence with an extended –10 region. The imperfect inverted repeat ccacgaaacgtgt, which overlaps with this promoter, might serve as a binding site for the cI-like repressor. The gene encoding the latter protein was also preceded by a putative operator sequence (ccacgaaaggaaa). The sequences of the operators resemble those of the operator from lactococcal phage Tuc2009 (Kenny *et al.*, 2006). An additional copy of the operator (ccacgaaacgtgg) with perfect dyad symmetry was present in the 3' region of ORF8, encoding a putative Cro-like repressor. Such an additional copy of the operator has also been found in the Cro-like repressor of lactococcal phages LC3 and Tuc2009 (Blatny *et al.*, 2003; Kenny *et al.*, 2006). In contrast to most other streptococcal and lactococcal phages (Lucchini *et al.*, 1999), the Cro-like repressor was separated from the cI

repressor by three ORFs. Of these, ORF5 was very small and without similarity to other genes, whereas ORF6 and ORF7 showed similarity to ORFs from *S. mitis* phage SM1 (Siboo *et al.*, 2003).

In the module that putatively encodes proteins involved in regulation and transcription, only a few of the ORFs could be assigned a function based on sequence similarity with known proteins or on the presence of conserved motifs. The packaging module consists of ORF33 and ORF34, which most likely encode the small and large subunits, respectively, of a terminase. The structural module of PH15 encodes the structural proteins of the phage in the order capsid and tail, which corresponds to the usual order in streptococcal and lactococcal phages. Overall, the ORFs from the structural module showed a similar synteny and sequence similarity to the structural proteins from *S. pyogenes* ϕ NIH1.1 (Ikebe *et al.*, 2002) and *S. pneumoniae* MM1 (Obregón *et al.*, 2003).

The lysis modules of *Siphoviridae* generally consist of two genes, which encode a holin and an endolysin. This appears to be also the case for PH15. ORF54 is predicted to contain two transmembrane helices and was most similar to the holins from *S. mitis* phage Sk137 (Llull *et al.*, 2006) and *S. pneumoniae* phage EJ-1 (Díaz *et al.*, 1996). ORF55 has an amidase-5 domain, suggesting that it cleaves the peptidoglycan at the amide bond between *N*-acetylmuramic acid and L-alanine of the stem peptide. ORF55 was similar to a number of putative endolysins from *Streptococcus thermophilus* bacteriophages (Sheehan *et al.*, 1999). In about half

of these *S. thermophilus* bacteriophages, an intron has been found to be present in the endolysin gene (Foley *et al.*, 2000), but there was no indication that this is the case for the endolysin gene of PH15 (see below). Following ORF55, a putative termination signal was present with a free energy of $-13.5 \text{ kcal mol}^{-1}$ (56.5 kJ mol^{-1}).

The region downstream of the lysis module and flanking the *attR* site frequently encodes lysogenic conversion genes, whose expression changes the phenotype of the host strain (Canchaya *et al.*, 2003). In PH15, this region contains six ORFs (ORF56–ORF61), whose amino acid sequences did not show sequence similarity to those of proteins of known function. However, some of the ORFs (ORF56, ORF57 and ORF60) showed similarity to ORFs encoded in the lysogenic conversion region of *Streptococcus dysgalactiae* phage 3396 (Davies *et al.*, 2007). ORF56 and ORF59 were preceded by a consensus σ^{70} *E. coli* promoter sequence, suggesting that these genes are expressed in the lysogenic state.

Presence of CRISPR in *S. gordonii* Challis

It has recently been shown that one of the defence mechanisms against bacteriophage infection involves the CRISPR system (Barrangou *et al.*, 2007). CRISPR arrays consist of identical repeats that are separated by spacer sequences. Bacteria can acquire resistance to bacteriophage infection by the incorporation of short phage-derived sequences between the repeats (Barrangou *et al.*, 2007). The precise mechanism of phage resistance mediated by CRISPR is unknown, but is thought to be the counterpart of RNA interference in eukaryotes. The genome from *S. gordonii* Challis CH1 (accession no. CP000725; Vickerman *et al.*, 2007) contains a CRISPR array with repeats of 36 bp and a total of 26 spacers of 30 bp in size. The sequences of the repeats were identical to the repeats of CRISPR1 from different *S. thermophilus* strains (Barrangou *et al.*, 2007). Analysis of the spacer sequences indicated that none of these spacers matched with the genome sequence of phage PH15. Interestingly, however, the first spacer matched with the genome sequence of the inducible *S. pneumoniae* prophage EJ-1 (Romero *et al.*, 2004b), apart from one mismatch. This suggests that *S. gordonii* Challis encountered a phage very similar to EJ-1.

Characterization of the phage attachment site in *S. gordonii* OMZ1039

To determine the phage and bacterial attachment sites *attP* and *attB*, genomic DNA from OMZ1039 was directly sequenced. The choice of primers was based on the assumption that the attachment sequences are in proximity to the putative phage integrase at the 5' end and of the putative lysin at the 3' end. Thus, one primer (15a5-1) was located in the integrase gene (ORF1), the other (15g7-7) was located in ORF61. The genomic sequencing resulted in distinct sequence traces, up to a point where overlapping

peaks were observed, probably due to a mixture of phage and bacterial sequences. By subtracting the known phage sequence from the mixed sequence, the bacterial sequence was deduced. Subsequently, the bacterial sequence that bordered the phage sequence was confirmed by PCR, combining primers located in the phage sequence and the bacterial sequence. The PCR products had the expected sizes and sequencing confirmed the correctness of the attachment sites.

The *attP* site could be deduced from these sequences and was 18 bp with the sequence 5'-TAAGGCGGTAGAC GGATT-3'. The *attB* site was found in the genome of *S. gordonii* Challis NCTC 7868. Further analysis showed that the first 16 bp of the *attP* site were identical to the last 16 bp of the cysteinyl-tRNA, indicating that in OMZ1039 phage PH15 has integrated just downstream of this gene (Fig. 3a). By PCR amplification of chromosomal DNA with primers 15att2 and 15att3, it was determined whether other *S. gordonii* strains possess the same *attB* sequence. Template DNA from *S. gordonii* strains OMZ504, OMZ505 and OMZ606 gave a PCR product of the expected size, whereas template DNA from OMZ938 did not yield a product. Sequencing of the products from OMZ504 and OMZ505 confirmed that the *attB* site was identical to that in OMZ1039.

The *att* sequence from PH15 was similar at the 5' end to the *attP* sequence from *L. lactis* phage TPW22. This phage has been shown to integrate in *L. lactis* at the identical position of the cysteinyl-tRNA (Petersen *et al.*, 1999; Williams, 2002). Due to sequence differences in the cysteinyl-tRNA gene, the attachment sites are also different. The integrase from TPW22 has 49% identity to the integrase of PH15. It thus appears that the integrase has coevolved with the cysteinyl-tRNA gene. As described above, the putative integrase from *S. agalactiae* phage LambdaSa2 shows very high sequence similarity to the PH15 integrase. Analysis of the sequence downstream of the LambdaSa2 integrase gene indicated the presence of an identical attachment site (not shown).

BLAST searches showed that, apart from *S. gordonii*, the PH15 *attB* sequence was fully conserved in *S. agalactiae*. Allowing for one mismatch at the last base pair, the *attB* sequence was also found in all other streptococci whose genome has been sequenced up to now: *Streptococcus suis*, *S. mutans*, *S. pneumoniae*, *Streptococcus sanguinis*, *S. pyogenes* and *S. thermophilus*.

Construction of an integration vector for (oral) streptococci

To confirm the functionality of the putative integrase, the *int* gene (ORF1) and *attP* were cloned into vector pVA891 (Macrina *et al.*, 1983), which cannot replicate in streptococci. In plasmid pOMZ292, the *int* gene is preceded by the putative promoter present in PH15. Plasmid pOMZ291 contains an additional synthetic promoter (cp32) upstream

Fig. 3. (a) Integration site of PH15. The gene organization around the attachment site of the *S. gordonii* CH1 genome (accession no. CP000725; Vickerman *et al.*, 2007) is shown. (b) Map of the streptococcal integration vector pOMZ291.

of this putative promoter (Fig. 3b). It was examined whether the plasmids could be integrated into the genome of *S. gordonii* OMZ938 after introduction by transformation of naturally competent cells. Plasmid pOMZ291 gave rise to transformants (4×10^3 c.f.u. μg^{-1}), whereas transformation with plasmids pVA891 or pOMZ292 did not yield colonies. This confirms the functionality of both integrase and attP. It also suggests that successful integration requires the presence of the synthetic promoter. The putative promoter located upstream of the integrase may not be functional in plasmid pOMZ292, since a phage-encoded transcription factor might be lacking. In addition, a plasmid with a C-terminal-truncated integrase gene (pOMZ308) could not be introduced into *S. gordonii* OMZ938, which shows that the integrase is absolutely required for integration. PCR and sequence analysis confirmed that pOMZ291 had inserted at attB (not shown). The stability of the integrated plasmid was evaluated by cultivation in THY medium without antibiotic selection and counting the number of antibiotic-sensitive colonies after plating. After about 50 generations in non-selective medium, all colonies were antibiotic-resistant. After about 100 generations, six out of 50 colonies had lost antibiotic resistance.

The conservation of the attB sequence in streptococci should in principle allow for integration of pOMZ291 in the

genomes of a wide variety of oral streptococcal species. Indeed, transformation of *S. mutans* strains OMZ381 and OMZ918, *S. sanguinis* strain OMZ500 and *S. mitis* strain OMZ991 with plasmid pOMZ291, but not with pVA891, yielded erythromycin-resistant colonies. Combining PCR and sequence analysis of the product, it was confirmed that integration occurred at the expected position in the genome of *S. mutans* OMZ918. Since the attB sequence is present in all sequenced streptococcal genomes, it is expected that plasmid pOMZ291 will integrate in most streptococcal genomes, provided that the organism can be transformed. It may thus be a valuable tool for delivery of foreign genes in single copy. The system may also be used to introduce single-copy transcriptional fusions at an ectopic site.

Production of active endolysin from the *lysA* gene of PH15

Phage endolysins are potentially useful as antibacterial agents to treat infections because of their specificity and high activity. To explore the specificity of the PH15 endolysin, the *lysA* gene (ORF55) was cloned into the *E. coli* expression vector pET28b, expressed as a fusion to a C-terminal His-tag. Upon induction with IPTG, a 31 kDa protein was produced, which is somewhat smaller than the calculated size of 32.4 kDa. Addition of the protein to a

culture of *S. gordonii* OMZ1039 led to a more than threefold decrease in optical density within 5 min (Fig. 4). The protein also showed weak activity towards *Ent. faecalis* OMZ940, but not towards *L. lactis* OMZ1041 (Fig. 4). The protein retained activity for 6 months when stored at 4 °C or at -20 °C in 44 % glycerol. The activity of the purified protein towards different bacterial species was evaluated by spot assays (Table 1). The protein inhibited growth of all tested strains of *S. gordonii*, *S. mutans* and *S. sanguinis*. The protein was also active with the pathogenic streptococci *S. pyogenes* and *S. agalactiae*. However, *S. mitis* and *S. oralis* strains and one *S. pneumoniae* strain were not inhibited. This could be caused by a difference in murein structure, since *S. mitis* and *S. oralis* contain a direct link between the stem peptides (Lys-direct type peptidoglycan), whereas *S. gordonii*, *S. mutans* and *S. sanguinis* contain a Lys-Ala₁₋₃ or Lys-Ala₂₋₃ type peptidoglycan (Schleifer & Kilpper-Bälz, 1987). The requirement for a Lys-Ala₁₋₃ type peptidoglycan could indicate that the enzyme acts as an endopeptidase, cleaving a bond within the peptide cross-link. Growth of *S. thermophilus* OMZ1082 was also inhibited by LysA, which is not surprising, since the protein was most similar in sequence to endolysins from *S. thermophilus* bacteriophages. The enzymes from *S. thermophilus* have not been characterized in detail, but the enzyme from bacteriophage O1205 has also been reported to be active towards *S. mutans* (Sheehan *et al.*, 1999).

In summary, phage PH15, which is believed to be the first *S. gordonii* bacteriophage to be identified, appears structurally related to *S. pyogenes* phage ϕ NIH1.1 and *S.*

pneumoniae phage MM1. The phage was found to harbour a putative lysogenic conversion region, which may encode virulence factors. The phage endolysin was active towards a broad range of oral streptococci, including pathogenic streptococci. Finally, the conserved nature of the phage attachment site may be exploited for the construction of streptococcal integration vectors.

ACKNOWLEDGEMENTS

I thank Verena Osterwalder and Steve Reese for excellent technical assistance, and Rudolf Gmür for critical reading of the manuscript. This work was supported by grant 215 from the research foundation of the Swiss Dental Association SSO.

REFERENCES

- Ajdić, D., McShan, W. M., McLaughlin, R. E., Savić, G., Chang, J., Carson, M. B., Primeaux, C., Tian, R., Kenton, S. & other authors (2002). Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. *Proc Natl Acad Sci U S A* **99**, 14434–14439.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. E., Seidman, J. G., Smith, J. A. & Struhl, K. (1987). *Current Protocols in Molecular Biology*. New York: Wiley.
- Banks, D. J., Beres, S. B. & Musser, J. M. (2002). The fundamental contribution of phages to GAS evolution, genome diversification and strain emergence. *Trends Microbiol* **10**, 515–521.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A. & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712.
- Bartilson, M., Marra, A., Christine, J., Asundi, J. S., Schneider, W. P. & Hromockyj, A. E. (2001). Differential fluorescence induction reveals *Streptococcus pneumoniae* loci regulated by competence stimulatory peptide. *Mol Microbiol* **39**, 126–135.
- Blatny, J. M., Ventura, M., Rosenhaven, E. M., Risøen, P. A., Lunde, M., Brüssow, H. & Nes, I. F. (2003). Transcriptional analysis of the genetic elements involved in the lysogeny/lysis switch in the temperate lactococcal bacteriophage ϕ LC3, and identification of the Cro-like protein ORF76. *Mol Gen Genom* **269**, 487–498.
- Canchaya, C., Proux, C., Fournous, G., Bruttin, A. & Brüssow, H. (2003). Prophage genomics. *Microbiol Mol Biol Rev* **67**, 238–276.
- Davies, M. R., McMillan, D. J., Van Domselaar, G. H., Jones, M. K. & Sriprakash, K. S. (2007). Phage 3396 from a *Streptococcus dysgalactiae* subsp. *equisimilis* pathovar may have its origins in *Streptococcus pyogenes*. *J Bacteriol* **189**, 2646–2652.
- Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.-F., Guindon, S., Lefort, V. & other authors (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**, W465–W469.
- Díaz, E., Munthali, M., Lünsdorf, H., Höltje, J.-V. & Timmis, K. N. (1996). The two-step lysis system of pneumococcal bacteriophage EJ-1 is functional in Gram-negative bacteria: triggering of the major pneumococcal autolysin in *Escherichia coli*. *Mol Microbiol* **19**, 667–681.
- Douglas, C. W., Heath, J., Hampton, K. K. & Preston, F. E. (1993). Identity of viridans streptococci isolated from cases of infective endocarditis. *J Med Microbiol* **39**, 179–182.

Fig. 4. Lysis of target bacteria by purified endolysin from PH15. Overnight cultures of *S. gordonii* OMZ1039, *Ent. faecalis* OMZ940 and *L. lactis* OMZ1041 were diluted 50-fold into fresh medium and grown for 6 h at 37 °C. Five microlitres of purified enzyme (11 µg) was added to 1 ml of cell suspension, and the OD₆₀₀, measured on a Hitachi U-2000 spectrophotometer, was monitored.

- Fischetti, V. A., Medaglini, D. & Pozzi, G. (1996). Gram-positive commensal bacteria for mucosal vaccine delivery. *Curr Opin Biotechnol* **7**, 659–666.
- Foley, S., Bruttin, A. & Brüssow, H. (2000). Widespread distribution of a group I intron and its three deletion derivatives in the lysin gene of *Streptococcus thermophilus* bacteriophages. *J Virol* **74**, 611–618.
- Huang, X. & Madan, A. (1999). CAP3: a DNA sequence assembly program. *Genome Res* **9**, 868–877.
- Ikebe, T., Wada, A., Inagaki, Y., Sugama, K., Suzuki, R., Tanaka, D., Tamaru, A., Fujinaga, Y., Abe, Y. & other authors (2002). Dissemination of the phage-associated novel superantigen gene *speL* in recent invasive and noninvasive *Streptococcus pyogenes* M3/T3 isolates in Japan. *Infect Immun* **70**, 3227–3233.
- Jensen, P. R. & Hammer, K. (1998). The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl Environ Microbiol* **64**, 82–87.
- Kawamura, Y., Whiley, R. A., Shu, S.-E., Ezaki, T. & Hardie, J. M. (1999). Genetic approaches to the identification of the mitis group within the genus *Streptococcus*. *Microbiology* **145**, 2605–2613.
- Kenny, J. G., Leach, S., de la Hoz, A. B., Venema, G., Kok, J., Fitzgerald, G. F., Nauta, A., Alonso, J. C. & van Sinderen, D. (2006). Characterization of the lytic-lysogenic switch of the lactococcal bacteriophage Tuc2009. *Virology* **347**, 434–446.
- Kilian, M., Mikkelsen, L. & Henrichsen, J. (1989). Taxonomic study of viridans streptococci: description of *Streptococcus gordonii* sp. nov. and emended descriptions of *Streptococcus sanguis* (White and Niven 1946), *Streptococcus oralis* (Bridge and Sneath 1982), and *Streptococcus mitis* (Andrewes and Horder 1906). *Int J Syst Bacteriol* **39**, 471–484.
- Kuramitsu, H. K., He, X., Lux, R., Anderson, M. H. & Shi, W. (2007). Interspecies interactions within oral microbial communities. *Microbiol Mol Biol Rev* **71**, 653–670.
- Lee, S. F. (2003). Oral colonization and immune responses to *Streptococcus gordonii*: potential use as a vector to induce antibodies against respiratory pathogens. *Curr Opin Infect Dis* **16**, 231–235.
- Llull, D., López, R. & García, E. (2006). Skl, a novel choline-binding N-acetylmuramoyl-L-alanine amidase of *Streptococcus mitis* SK137 containing a CHAP domain. *FEBS Lett* **580**, 1959–1964.
- Loeffler, J. M. & Fischetti, V. A. (2006). Lysogeny of *Streptococcus pneumoniae* with MM1 phage: improved adherence and other phenotypic changes. *Infect Immun* **74**, 4486–4495.
- Lucchini, S., Desiere, F. & Brüssow, H. (1999). Similarly organized lysogeny modules in temperate *Siphoviridae* from low GC content Gram-positive bacteria. *Virology* **263**, 427–435.
- Lukashin, A. V. & Borodovsky, M. (1998). GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res* **26**, 1107–1115.
- Macrina, F. L., Evans, R. P., Tobian, J. A., Hartley, D. L., Clewell, D. B. & Jones, K. R. (1983). Novel shuttle plasmid vehicles for *Escherichia-Streptococcus* transgeneric cloning. *Gene* **25**, 145–150.
- Mitchell, J., Siboo, I. R., Takamatsu, D., Chambers, H. F. & Sullam, P. M. (2007). Mechanism of cell surface expression of the *Streptococcus mitis* platelet binding proteins PblA and PblB. *Mol Microbiol* **64**, 844–857.
- Obregón, V., García, J. L., García, E., López, R. & García, P. (2003). Genome organization and molecular analysis of the temperate bacteriophage MM1 of *Streptococcus pneumoniae*. *J Bacteriol* **185**, 2362–2368.
- Petersen, A., Josephsen, J. & Johnsen, M. G. (1999). TPW22, a lactococcal temperate phage with a site-specific integrase closely related to *Streptococcus thermophilus* phage integrases. *J Bacteriol* **181**, 7034–7042.
- Poyart, C., Quesne, G., Coulon, S., Berche, P. & Trieu-Cuot, P. (1998). Identification of streptococci to species level by sequencing the gene encoding the manganese-dependent superoxide dismutase. *J Clin Microbiol* **36**, 41–47.
- Romero, P., López, R. & García, E. (2004a). Characterization of LytA-like N-acetylmuramoyl-L-alanine amidases from two new *Streptococcus mitis* bacteriophages provides insights into the properties of the major pneumococcal autolysin. *J Bacteriol* **186**, 8229–8239.
- Romero, P., López, R. & García, E. (2004b). Genomic organization and molecular analysis of the inducible prophage EJ-1, a mosaic myovirus from an atypical pneumococcus. *Virology* **322**, 239–252.
- Schleifer, K. H. & Kilpper-Bälz, R. (1987). Molecular and chemotaxonomic approaches to the classification of streptococci, enterococci and lactococci: a review. *Syst Appl Microbiol* **10**, 1–19.
- Sheehan, M. M., Stanley, E., Fitzgerald, G. F. & van Sinderen, D. (1999). Identification and characterization of a lysis module present in a large proportion of bacteriophages infecting *Streptococcus thermophilus*. *Appl Environ Microbiol* **65**, 569–577.
- Siboo, I. R., Bensing, B. A. & Sullam, P. M. (2003). Genomic organization and molecular characterization of SM1, a temperate bacteriophage of *Streptococcus mitis*. *J Bacteriol* **185**, 6968–6975.
- Tanzer, J. M., Thompson, A. M., Grant, L. P., Vickerman, M. M. & Scannapieco, F. A. (2008). *Streptococcus gordonii*'s sequenced strain CH1 glucosyltransferase determines persistent but not initial colonization of teeth of rats. *Arch Oral Biol* **53**, 133–140.
- Tettelin, H., Massignani, V., Cieslewicz, M. J., Eisen, J. A., Peterson, S., Wessels, M. R., Paulsen, I. T., Nelson, K. E., Margarit, I. & other authors (2002). Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*. *Proc Natl Acad Sci U S A* **99**, 12391–12396.
- 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.
- van der Ploeg, J. R. (2007). Genome sequence of *Streptococcus mutans* bacteriophage M102. *FEMS Microbiol Lett* **275**, 130–138.
- Vickerman, M. M., Iobst, S., Jesionowski, A. M. & Gill, S. R. (2007). Genome-wide transcriptional changes in *Streptococcus gordonii* in response to competence signaling peptide. *J Bacteriol* **189**, 7799–7807.
- Wang, B.-Y. & Kuramitsu, H. K. (2005). Interactions between oral bacteria: inhibition of *Streptococcus mutans* bacteriocin production by *Streptococcus gordonii*. *Appl Environ Microbiol* **71**, 354–362.
- Williams, K. P. (2002). Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. *Nucleic Acids Res* **30**, 866–875.
- Xu, P., Alves, J. M., Kitten, T., Brown, A., Chen, Z., Ozaki, L. S., Manque, P., Ge, X., Serrano, M. G. & other authors (2007). Genome of the opportunistic pathogen *Streptococcus sanguinis*. *J Bacteriol* **189**, 3166–3175.

Edited by: M. Kilian