ucc

Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland

IDENTIFICATION OF GENES INVOLVED IN THE DNA INJECTION PROCESS BY LACTOCOCCAL P335 PHAGES

A Thesis Presented to the National University of Ireland Cork

By

Andrea Carolina Erazo Garzon

BSc Pharmaceutical Biotechnology

For the Degree of

PhD in Microbiology

School of Microbiology

National University of Ireland Cork

Supervisors: Prof. Douwe van Sinderen and Dr. Jennifer Mahony

Head of School: Prof. Gerald Fitzgerald

February 2019

Dedications

To my beloved abuelitos,

Who made me the person that I am today.

Your unconditional love has always been a source of encouragement to make the most of the opportunities I've been given in my life.

To my beloved tía Yolanda and tío Mario,

The reason I appreciate life.

The way you embraced the eventualities of life with such grace and courage will always be a source of inspiration.

DECLARATION

I certify that the thesis submitted is my own work and has never been submitted for another degree, either at University College Cork or elsewhere.

Andrea Carolina Erazo Garzon

_____

Table of Contents

Abstract

Bacteriophages (phages) belonging to the P335 lactococcal group are one of the most frequently isolated phages in fermentation processes using *Lactococcus lactis* strains as the primary starter culture. Despite efforts to prevent phage attack, problems caused due to phage infection during large scale milk fermentations are still regularly reported.

Lactococcus lactis subsp. *cremoris* (*L. lactis*) 3107 is a dairy starter strain and host for the model lactococcal P335 phage TP901-1. Three derivatives of this strain that are resistant to TP901-1, while remaining sensitive to infection by another P335 phage called LC3 were employed to define molecular players involved in the phage infection pathway. Comparative genomic analysis of *L. lactis* 3107 and its phage-resistant derivatives identified several mutations that may cause the observed resistance. Complementation assays with one specific construct, termed pNZ*gtfA*, restored phage sensitivity and lysogenization frequency in two of the derivative strains. pNZ*gtfA* expresses the *L3107_1442* gene (renamed here as *gtfA*) which encodes a predicted membrane-associated glycosyltransferase (GTF), GtfA. Interestingly, the mutant strains also exhibited reduced sensitivity to the rare lactococcal 949 group phages, 949 and WRP3, while complementation assays fully restored infectivity by these phages, thus suggesting that *gtfA* is also involved in phage-host interactions of these phages. TP901-1 escape mutants capable of infecting these phage-resistant derivatives possess mutations in the gene encoding the structural and/or lytic domain of the Tail-associated lysin in order to adapt to cells which have acquired phage resistance.

Interestingly, the gene *gtfA* is ubiquitous among *Lactococcus* genomes, *L. lactis* subsp. *cremoris* NZ9000 harbours a 99 % identical homologue of $gtfA_{3107}$ on its genome, referred to here as $gtfA_{NZ9000}$ (corresponding to locus tag *LLNZ_08920*). Therefore, a mutation (stop codon) was introduced into the coding sequence of $gtfA_{NZ9000}$ by targeted mutagenesis in order to expose the mutant strain generated to NZ9000-infecting 936 group phages. Results showed that a non-functional GtfA protein has no effect on phages belonging to the lactococcal 936 group. $GtfA_{3107}$ possess similar characteristics to proteins belonging to the GT-C family of GTFs which are known to be involved in glycosylation of substrates at the out-facing side of the

cytoplasmic membrane. Therefore, GtfA is presumed to glycosylate a substrate, most likely a cell envelope-associated glycopolymer, such as a cell wall polysaccharide (CWPS) or teichoic acid (TA). If $gtfA_{NZ9000}$ is involved in the modification of CWP components, the inactivation of this gene may have generated a mutant strain deficient in the glycosylation of PSP or rhamnan. In an attempt to identify the GtfA substrate, the biochemical composition and structure of CWPS components (rhamnan and PSP) of *L. lactis* strains 3107 and NZ9000, and a number of derivative strains were analysed. Results showed that derivatives possessing a non-functional *gtfA* or over-expressing *gtfA* do not possess any structural modifications in either the rhamnan or PSP components of the CWPS as compared to the 3107 or NZ9000 parent strains. In addition, the biochemical composition and structure of lipoteichoic acids (LTAs) in *L. lactis* 3107 and a representative phage-resistant derivative was also analysed, results revealed that *gtfA* is not responsible for terminal galactose substitution of LTA. Currently, further biochemical analysis of TA composition in *L. lactis* 3107 is being performed in order to identify the GtfA substrate and its relevance to phage infection.

Attempts to generate LC3-resistant derivatives of *L. lactis* 3107 resulted in the isolation of lysogens of this phage that displayed resistance to multiple P335 phages. The LC3-prophage encoded superinfection exclusion (Sie) protein $Sie_{2009/LC3,}$ which is an identical (100 % identity) homologue of the previously identified Sie system Sie_{2009} was shown to be responsible for the phage resistance observed. Furthermore, superinfection immunity (Sii) against certain phages belonging to the P335 group was also exhibited in the presence of the LC3-prophage encoded repressor protein Rep_{LC3}.

Overall this study has identified bacterial and phage-encoded protein(s) required for, or interfering with, the infection process of the dairy starter strain *L. lactis* 3107 by phages belonging to the lactococcal P335 group, as well as 936 and 949 group phages. This study forms the basis for future studies aimed to limit the proliferation of phages in the industrial context.

Chapter I

Literature review

Chapter I contents

1.1 Introduction to bacteriophages

Bacteriophages, also known as phages, are bacteria-infecting viruses; biochemically phages represent complex macromolecular assemblies which typically consist of proteins and nucleic acids. Bacteriophages were first discovered in 1915 by Frederick Twort (1) and in 1917 by Felix d'Herelle (2). Phages and their bacterial hosts are presumed to been subject to co-evolution, and although the precise time dimension of this is unknown, it may have been in operation for some 3 billion years. The fact that phages do not possess a universally conserved molecule that can be used as an evolutionary clock, combined with a complete lack of a fossil record, leaves us to rely solely on comparative phage genome analysis as a tool to understand phage evolution (3). Phages are ubiquitous and are considered to represent the most numerous biological entities in the biosphere, with our planet harbouring an estimated 10^{30} to 10^{32} phage particles (4), thus outnumbering bacteria by ~ 10-fold (5). It was previously estimated that ocean water and sediment contains up to 10^{7} phage particles ml^{-1} (6). Following a comparison between the number of bacteria and the vast number of phages coexisting in these niches, it has been suggested that ~ 10^{25} phage infections occur every second worldwide (7).

Phages are obligate parasites, thus while they carry all information necessary to direct their own reproduction in a susceptible bacterial host, they require the replication machinery, energy and ribosomes of their host to propagate. Therefore, phages reside in any ecosystem where bacteria can be found (8), including soil (9, 10), water (11), and the human gastrointestinal tract (12, 13). Each phage typically infects a specific group of bacteria, which is often a (small) subset of one species, though several related species can sometimes be infected by the same phage (14). For example, the lytic *Streptococcus thermophilus* (*St. thermophilus*) phages DT1, DT4, MD1, MD2 and Q5 can only infect specific *St. thermophilus* strains (15), whereas the *Lactobacillus plantarum* (*Lb. plantarum*) phages 14-E10 and 22-A2 are capable of infecting *Lactobacillus brevis* as well as its host strain *Lb. plantarum* (16). Additionally, the virulent phage SH7 belonging to the T4 virus genus of the *Myoviridae* family has been shown to infect *Escherichia coli* (*E. coli*) 0157:H7, *Salmonella* Paratyphi, *Shigella dysenteriae* as well as its original host strain *Shigella flexneri*, all of which belong to the Enterobacteriaceae family (17). Phages maintain a dynamic balance among the wide variety of bacterial species due to their high-level of specificity, their ability to reproduce rapidly in suitable bacterial hosts, and long-term

infectivity as many phages can infect after decades of not encountering a host (18). Therefore, phages also play an important role in bacterial population dynamics, and represent an important driver of bacterial evolution (19, 20). Furthermore, phages affect bacterial pathogenesis (21, 22). In the past 90 years phages have been used as bactericidal agents to treat bacterial infections in humans, an application also known as phage therapy, and currently explored in Western medicine as an alternative to antibiotics (23, 24). Furthermore, phage-derived proteins are used as molecular machines (25), diagnostic agents (26), and therapeutic agents (26-28), and have been employed to aid in drug discovery (29). Moreover, temperate phages have been shown to be capable of transferring desirable or undesirable traits on their bacterial host strains. For example, lactococcal temperate phages may confer natural phage defence mechanisms on their host including adsorption inhibition, DNA injection blocking, abortive infection and restriction-modification systems (30-36). For more information on temperate phages and some of the traits they may confer on their lysogenized host see Section 1.5.3: Lytic or lysogenic phage cycle. Despite the many advantages of bacteriophages, these bacteria-infecting viruses are a major nuisance to various biotechnology processes that rely on bacteria to produce their high quality end products. For example, bacteria are required to manufacture fermented dairy products such as cheese and yogurt, or produce a pharmaceutical molecule (e.g. insulin) in large-scale, and in such cases phage infections may cause serious issues that can lead to production delays, low quality product or even completely arrest the fermentation process. Phage problems have previously been reported in the food, chemical, pharmaceutical, feed and pesticide industries (4). However, by far the most frequently documented phage problems are those occurring in the dairy industry, with the first incident of phages affecting a dairy starter culture being reported by Whitehead and Cox in 1935 (37). Since then phage contamination events in the dairy industry have been consistently reported to occur at any major milk fermentation facility across the globe (38-42).

1.2 Bacteriophages and the dairy industry

According to The Portal for Statistics website (https://www.statista.com; information relates to 2018), the economic value of the global dairy industry, including fermented and non-fermented products was estimated to be ~ 300 billion dollars, with cheese production accounting for 27 % of this value. The economic success of the

fermented dairy industry is dependent on the consistent and reliable activity of starter cultures composed of bacterial strains belonging to the Gram-positive, anaerobic, non-spore forming lactic acid bacteria (LAB), which owe their name to the fact that they produce lactic acid as the primary metabolic end-product of hexose fermentation (43). LAB strains are currently among the most intensively studied bacteria due to their large-scale use in dairy starter cultures, being demonstrated by the fact that ~ 10^{14} bacterial cells are required to produce a ton of cheese (8). The most commonly used LAB as starter cultures include strains of *Lactococcus lactis* (*L. lactis*), *St. thermophilus*, and various *Leuconostoc* and/or *Lactobacillus* species (8), which, like all bacteria, are susceptible to phage attack. The dairy industry has developed a series of strategies to prevent phage attack of starter cultures, such as improved factory design, enhanced sanitation, the use of specific starter cultivation media, strain rotation, and the employment of phage-resistant strains (8). Despite substantial and ongoing efforts to prevent this natural phenomenon from occurring, phage contamination in the dairy industry is still a very common phenomenon with a significant impact on the economic success of this industry (4, 44).

The manufacture of cheese involves the inoculation of (typically, pasteurized) milk with a starter culture so as to initiate and perform the fermentation/acidification process and to obtain high-quality end-products (8). *L. lactis* is the major LAB starter species employed in the cheese industry. Lactococcal phages are ubiquitous in the dairy environment, as they reside in non-sterile raw milk, and are not fully inactivated by pasteurization treatments (4). This (relatively) low phage concentration of raw milk combined with the presence of phage-sensitive bacterial cells in a starter culture will allow the rapid development of a specific phage population. This will negatively impact on final product quality and production regimes, causing significant economic losses (31). The primary source of phage contamination in the dairy fermented industry appears to be raw milk, although other sources of phage contamination may include processed or recycled ingredients, starter cultures and air/surfaces (8). Phages may be present in processed or recycled ingredients such as whey proteins (45), which are added to improve the texture/taste, to increase the nutrient value, and to increase the yield of the final product. Moreover, the starter culture may itself act as a phage reservoir as many LAB strains harbour prophages (46, 47), which may be triggered to enter the lytic cycle during the fermentation process when exposed to heat, salts, antimicrobials, starvation, or UV (48, 49). For more information on prophages see

Section 1.5: The infection process and life cycle of a bacteriophage. The presence of airborne lactococcal phages in the fermented dairy industry has previously been documented (50-53) making the air and surfaces within a dairy factory another potential source of phage contamination. Studies have also shown that phages can persist as infectious particles over a long period of time in dairy factories (54, 55). Consequently, lactococcal phages have been the focus of intensive research for many years as a thorough understanding of how these problematic viral parasites interact with their hosts is necessary in order to devise strategies that would limit their negative impact. Such strategies may involve the construction of robust starter cultures through genetic engineering of LAB (56), thus creating mutant LAB strains resistant to specific phages, or the development of strain rotations/blends.

1.3 Taxonomy of lactococcal phages

The negative impact of lactococcal phages on dairy fermentations as well as their biodiversity in this ecological niche has led to the isolation and characterization of many lactococcal phages (57). Currently, the International Committee on Taxonomy of Viruses (ICTV) which classifies phages according to their phage morphology and genome organization has included lactococcal phages as members of the *Caudovirales* order of tailed phages, which represent over 95 % of all known phages (58). Phages of the order *Caudovirales* possess a nucleic acid-containing capsid, which may be isometric or elongated and a tail structure to recognize their bacterial host and to deliver their genome with high specificity (59). Phages belonging to the order *Caudovirales* are classified into three families (60): (i) the *Siphoviridae* family which represents bacterial viruses with long, non-contractile tails (e.g., *Bacillus subtilis* (*B. subtilis*)-infecting SPP1 phage (61)), and which comprises ~ 60 % of the characterized phages, (ii) the *Myoviridae* family which represents viruses with long, complex contractile tails (e.g., *E. coli*-infecting T4 phage (62)) and which constitutes around 25 % of the characterized phages, and (iii) the *Podoviridae* family which encompasses viruses with short non-contractile tails (e.g., *Salmonella typhimurium*-infecting P22 phage (63, 64)), comprising ~ 15 % of the characterized phages. These families are further divided into genera and subgenera by criteria such as host range, genome size and genome configuration (linear, circular, and supercoiled) (60). All currently known *L. lactis* phages possess a double-stranded DNA (dsDNA) genome

and a non-contractile tail, making them members of the *Siphoviridae* family, with a small number belonging to the *Podoviridae* family (57).

More than two decades ago, lactococcal phages were classified into 12 groups, based mainly on phage morphology and DNA homology criteria: 936, c2, P335, 949, P087, 1358, 1706, 1483, T187 and BK5T phage groups, which belong to the *Siphoviridae* family, and phage groups P034 and KSY1, which belong to the *Podoviridae* family (65-68). This classification method was employed to compare lactococcal phage isolates and revealed that most phage isolates from dairy settings belong to one of three main groups, i.e. the 936, c2, and P335 groups. Of note, based on genome analysis it has since been proposed to merge the BK5-T group with the P335 group (69), while in addition DNA-DNA hybridization studies have suggested to merge the 1483 and T187 group with the P335 group (69-72). Furthermore, lactococcal phages Q54 and 1706 were shown to be unrelated to known lactococcal phages leading to the emergence of new groups (57). Thus, the classification system for lactococcal phages, as based on their morphological and genomic analyses, currently distinguishes ten groups, as demonstrated in Table 1. To date, eight groups including the 936, c2, P335, 949, P087, 1358, 1706 and Q54 phages belong to the *Siphoviridae* family due to their long, non-contractile tails, whereas the P034 and KSY1 phage groups belong to the *Podoviridae* family due to their short tails (57). Examples of some of the published lactococcal phages with some of them including their genomic sequence are shown in Table 2. Members of the P335, 936 and c2 groups are still the most frequently encountered lactococcal phages in dairy fermentations globally (69). Consequently, lactococcal phage studies have in recent years mainly concentrated on members of these three groups. Currently, two multiplex PCRs are available to rapidly assign new phage isolates to one of these three main lactococcal phage groups, P335, 936 and c2 (73, 74).

Table 1. Biodiversity of bacteriophages infecting *L. lactis* strains; modified from (57).

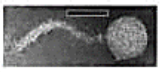
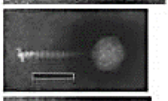
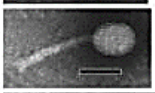
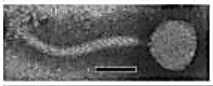
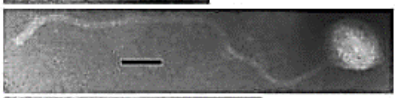
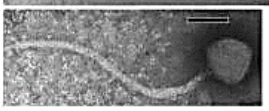
Family	Group	Phage	Tail length (nm)	Electron Micrograph
Siphoviridae				
	936	bIL170	126	
	P335	ul36	104	
	1358	1358	93	
	c2	c2	95	
	Q54	Q54	109	
	P087	P087	163	
	949	949	490	
	1706	1706	276	
Podoviridae				
	P034	P369	19	
	KSY1	KSY1	32	

Table 2. Examples of lactococcal phages plus associated genome and host characteristics.

Family	Group	Phage	P335 Sub-group	CWPS of host strain(s)	Genbank accession number	Reference
Siphoviridae	**P335**	BK5-T	I	ND	NC_002796	(75)
		4268	I	ND	NC_004746	(76)
		C41431	I	A & C	KX160219	(77)
		38502	I	A	KX160204	(77)
		Tuc2009	II	A	NC_002703	(78)
		P335	II	B	DQ838728	(79)
		TP901-1	II	C	NC_002747	(80)
		ul36	II	A & C	NC_004066	(79)
		53801	II	C	KX160207	(77)
		53802	II	C	KX160208	(77)
		49801	II	C	KX160205	(77)
		98103	II	C	KX160214	(77)
		98104	II	C	KX160215	(77)
		98102	II	C	KX160213	(77)
		98101	II	C	KX160212	(77)
		98204	II	A	KX160218	(77)
		98202	II	A	KX160216	(77)
		98203	II	A	KX160217	(77)
		LC3	III	C	NC_005822	(81)
		r1t	III	C	U38906	(82)
		Dub35A	III	C	KX160220	(77)
		Q33	IV	A, B & C	JX564242	(83)
		BM13	IV	ND	JX567312	(83)
		58502	IV	A	KX160209	(77)

		50902	IV	A	KX160206	(77)
		62502	IV	A	KX160210	(77)
		62503	IV	A	KX160211	(77)
		63301	ND	ND	KX456211	(84)
		86501	III	ND	KX456212	(84)
		50101	III	ND	KX456207	(84)
	936	bIL170	N/A	N/A	NC_001909	(85)
		sk1	N/A	N/A	NC_001835	(86)
		jj50	N/A	N/A	NC_008371	(87)
		712	N/A	N/A	NC_008370	(87)
		P008	N/A	N/A	NC_008363	(87)
	c2	c2	N/A	N/A	NC_001706	(88)
		bIL67	N/A	N/A	NC_001629	(89)
	949	949	N/A	N/A	NC_015263	(90)
		bIL168	N/A	N/A	NC_011046	(91)
	P087	P087	N/A	N/A	NC_012663	(92)
	1358	1358	N/A	N/A	NC_027120	(93)
	1706	1706	N/A	N/A	NC_010576	(94)
	Q54	Q54	N/A	N/A	NC_008364	(95)
Podoviridae	**P034**	ascc28	N/A	N/A	NC_010363	(96)
	KSY1	KSY1	N/A	N/A	NC_009817	(97)

Not determined (ND); Not applicable (N/A); cell wall polysaccharide (CWPS)

1.4 General structure and function of a tailed bacteriophage

In order to understand how tailed bacteriophages infect their hosts, we first need to understand their structure and the function of each component. Electron microscopy (EM) studies of single particles of model phages such as SPP1 (61), HK97 (98), TP901-1 (99), and Tuc2009 (100) have shown that these nanomachines are composed of a phage capsid, connector, tail and baseplate (BP) or tail tip (Fig. 1.1), each of which will be discussed below in more detail.

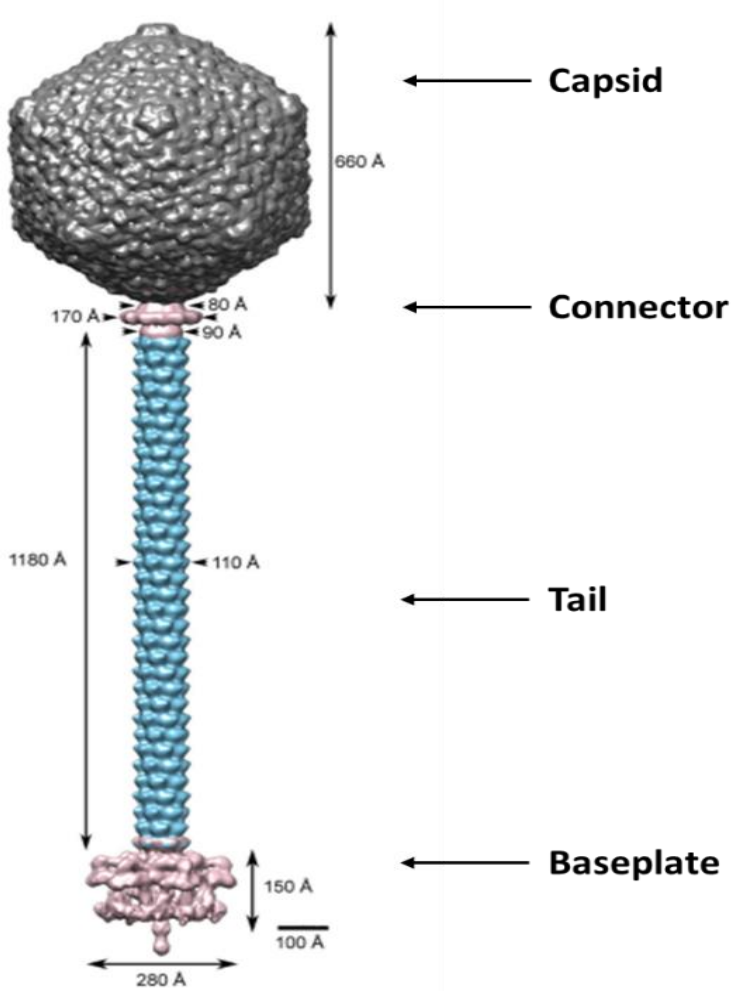


Figure 1.1. Complete structure of the lactococcal P335 bacteriophage TP901-1; modified from (99).

1.4.1 The capsid (head) and head-to-tail connector

The capsid (also known as the head) of a bacteriophage is a robust proteinaceous shell which allows phages to carry and protect their genetic material (101). The majority of known bacteriophages possess an icosahedral capsid which harbours dsDNA in a highly condensed state; the combined nucleic acid and capsid are known as the nucleocapsid (18). In many dsDNA phages, the assembly of the phage structure begins with the formation of the procapsid/prohead (which is a preformed protein shell). The procapsid structure is formed from multiple copies of the major capsid protein (MCP), which together with scaffolding protein forms a stabilising core until the DNA is packaged. Once it is filled with DNA, the procapsid undergoes maturation to form the mature phage capsid (102, 103).

In general, following infection of the host, which involves release of the phage genome from the capsid into the host cytoplasm, viral genome replication leads to the accumulation of DNA concatemers, which are substrates for DNA encapsidation (i.e. the DNA packaging process). Most dsDNA phages pack their DNA into the cavity of the procapsid through the connector, with the aid of a packaging complex (composed of a small and large terminase subunit and the portal protein) (104, 105). The small subunit of the terminase possesses DNA binding activity, while the large subunit is required for procapsid binding, DNA translocation and DNA cleavage (106). The portal protein plays an important role in DNA packaging as this protein forms a circular oligomer which surrounds the channel through which DNA enters and exits the head. This portal structure ensures an efficient DNA packaging process and determines the length of DNA that will be packaged into the procapsid (107, 108). The interaction of DNA-terminases with the portal protein allows translocation of DNA into the procapsid (109, 110). The successful completion of DNA packaging depends on the recognition of a specific sequence called *cos* or *pac* (111). DNA packaging begins when the small terminase recognises and binds to the DNA (*cos* or *pac* site) and subsequently the large terminase cleaves the DNA and interacts with the phage portal protein where DNA translocation into the procapsid is driven by ATP hydrolysis (112). DNA is translocated unidirectionally and begins with the generated end following cleavage. When the head has been filled to capacity, a termination cleavage occurs at a specific (Lambda, T3, T7) or non-specific (P1, P22, T1, SPP1) site producing the mature form of viral DNA present in virion particles (112). The DNA concatemer is cleaved before and after packaging by the nuclease activity of the large

terminase, while the energy for the DNA packaging process is provided by ATPase activity (105).

The *pac* site refers to a sequence which signals initiation of the phage DNA packaging process. The *pac* type phages, e.g. P2 (111), employ a headful approach mechanism whereby in the first round of replication the DNA is initially cut at the *pac* site and then translocated into the procapsid until it is filled, representing an amount of DNA that is greater than the length of a single complete genome. Each subsequent round of DNA replication results in the DNA being cut at non-specific sites downstream of the *pac* site releasing the packed DNA, resulting in packaging of circularly permuted DNA molecules with more than one unit length of genome. The end generated by the headful cleavage of the previous chromosome is the start of packaging events of the DNA, thus leading to the packaging of terminally redundant molecules (112). This so-called head-full packaging mechanism employed by *pac*-site phages appears to be a very tightly regulated process in order to prevent the formation of phages with unfilled or incompletely filled capsids (64, 98, 113, 114). The *cos* site refers to a specific cohesive end site on the phage DNA, *cos* type phages cut the DNA at the *cos* site generating single genomic units with self-complementary 3' overhangs (105). For example, the *cos* site of phage Lambda is a 22 bp sequence that is cut asymmetrically to allow the concatemeric DNA to be packaged in a single genome-specific, linearized manner into a prophage head leaving an overhang *cos* site at either end of the genome. Thus, when the DNA is injected into the cytoplasm the two cohesive ends anneal generating a covalently closed circular genome (111). Sequenced lactococcal P335 group phages have been reported to utilise the *cos* or *pac* method of packaging (115), while lactococcal 936 and c2 group phages utilise the *cos* method of packaging (87, 95).

The head-to-tail connector is a ring-shaped structure which connects the phage capsid with its tail (Fig. 1.1), the connector is composed of the portal protein, head completion proteins and a stopper protein. The connector is involved in phage DNA packaging (116-118) and DNA ejection (119). The connector of phage Lambda has been well studied and is located inside the capsid of the phage at the vertex where the tail attaches (119). Morphogenesis of the icosahedral Lambda head has been suggested to begin with the assembly of the connector which is composed of molecules of the minor head protein gpB which are involved in DNA entry (120). Fusion and cleavage reactions take place to form the procapsid. Terminase then binds to the connector and

concatameric DNA, and packages linear monomers of the dsDNA genome into the procapsid. The packaged DNA within the head is stabilized by incorporating molecules of major head protein which prevents the loss of DNA after it has been packaged (120). The gpW protein then binds to the connector region, which allows the incorporation of gpFII (121). Lambda gpFII protein is located at the bottom of the connector, representing 8 % of the mass of the connector, and forming the site on the head to which the tail binds (120).

1.4.2 The tail and baseplate / tail tip

The assembly of the tail and associated base plate (BP, sometimes also called the tail tip) of phages belonging to the lactococcal P335 and 936 groups has been the focus of several studies (79, 122). The tail of a siphophage is located between the connector and the BP/tail tip, and represents a structure with a continuous central channel which is used as a genome ejection tunnel to allow efficient viral genome delivery into the target host cell cytoplasm (61, 123-127). The BP or tail tip, also referred to as the host adsorption device, is located at the distal part of the phage tail and is responsible for host recognition, attachment, and initiation of infection. The structure of this adsorption device varies among lactococcal phages depending on the phage type (79, 122). The tail tube of the lactococcal 936-group phage p2 is formed by stacking of the major tail protein (MTP) and a tail terminator hexamer at the point linking to the connector, the tail length is determined by the tape measure protein (TMP) (103, 128). In the case of phages belonging to the *Myoviridae* family, additional ring-like structures are formed around the tail tube to form a contractile sheath, this differ from *Siphoviridae* phages which possess non-contractile tails (129). The BP structure of phage p2 is assembled from the distal tail protein (Dit), tail-associated lysin (Tal) and eighteen receptor-binding proteins (RBPs). The Dit is attached to the last (i.e. most distal relative to the head) MTP ring (130). The RBPs of the 936-group phage p2 appear to exist in an inactive or active form. In the presence of the cation Ca^{2+}, the RBPs appear to undergo a change in orientation to rotate the C-terminus 200° downwards towards the host cell receptor. This conformational change appears to activate the BP of p2 during attachment to the host as the binding sites are orientated towards the host, leading to the opening of a channel at the bottom of the BP for DNA passage ultimately resulting in DNA ejection (130).

The BP and tail assembly of the P335-group phage TP901-1 involves the gene products of *dit* (*orf46*), *tal* (*orf47*), *bppU* (*orf48*), and *bppL* (*orf49*) (131). The multicomponent BP structure of TP901-1 (99) (Fig. 1.2) consists of a Dit hexamer, which anchors an upper and lower BP disk composed of multiple copies of the upper BP protein (BppU) and lower BP protein (BppL), respectively, and a central tail fiber (Tal) protruding below the BP. The TP901-1 BP harbours an assembly of RBPs which recognize a specific saccharidic receptor located on the surface of a host cell, therefore the TP901-1 BP is fundamental for host infection as the BppL represents the RBP (131). The BP of TP901-1 harbours 54 RBP monomers (specified by ORF49) which are orientated towards the distal extremity and are not influenced by the presence of Ca^{2+} ions, indicating that TP901-1 is ready for host adhesion and does not need to undergo a conformational change to reach an activated state (as in the case of p2) (123). The RBP of TP901-1 is 163 amino acids long and forms a homotrimer which is composed of three domains, the N-terminal, β-prism neck, and C-terminal domain. The helical N-terminal domain comprises 63 residues, whereas the C-terminal domain is represented by residues 64-163 and is a six-stranded β-barrel which is also known as the head recognition domain. The β-prism neck domain is also known as the β-helix and is represented by residues 40-63. Fluorescence quenching experiments revealed that the RBP of TP901-1 exhibits high affinity for glycerol, muramyl-dipeptide, and other saccharides in solution (132).

In contrast, the tail tip of the model *B. subtilis*-infecting phage SPP1 (61, 124, 127) harbours a single component BP, also referred to as a "stubby end" which does not contain a sugar binding RBP. It has previously been suggested that such a discrete tail tip complex is found in phages adsorbing to proteinaceous receptors, such as SPP1 (133) and the lactococcal phage c2 (134), due to the high-affinity and specificity of such protein-protein interactions (135). Further studies have demonstrated that the tail tip of some of the c2-group phages may be capable of attaching to a secondary carbohydrate receptor (136).

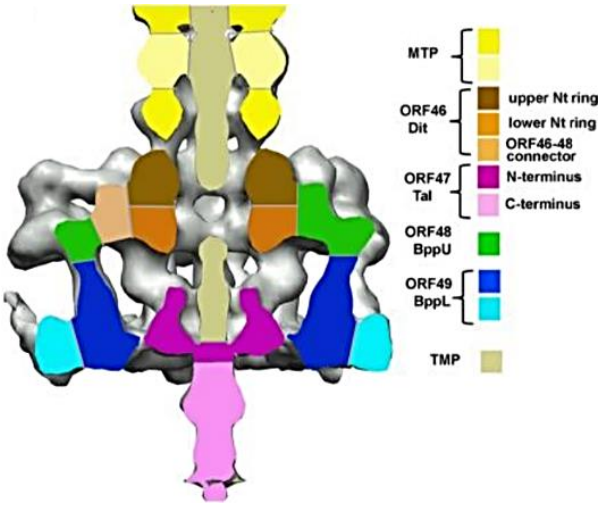


Figure 1.2. Representation of TP901-1 baseplate structure and its components represented by different colours. Figure modified from (137). N-terminal (Nt).

1.5 The infection process and life cycle of a bacteriophage

The infection process of tailed bacteriophages involves a number of tightly controlled steps which need to be understood in order to gain a proper appreciation of how phages infect their specific bacterial hosts.

1.5.1 Adsorption

The first step involved in the interaction of a tailed bacteriophage with its specific bacterial host is the recognition of and attachment/adsorption to a specific component of the host cell envelope. This binding can occur as a reversible and/or irreversible interaction, and is thought to be mediated by the tail tip or BP structure which acts as an adsorption device that binds to specific host receptor molecules that are present at the host cell surface (61, 123, 138-140). The attachment of some phages, such as coliphages (141), the model *Bacillus* phage SPP1 (133), and a number of lactococcal phages (140), seem to involve two separate adsorption stages and two distinct receptors to achieve successful attachment to its bacterial host. In Gram-negative bacteria the most common components of the cell envelope that act as host receptors for phages are porins, transporters and lipopolysaccharides (LPS). In contrast, in the structurally distinct cell envelope of Gram-positive bacteria receptor molecules for phages include cell wall polysaccharides (CWPS) (77, 140, 142-144), cell surface-exposed proteins and/or teichoic acids (TAs) (133, 145, 146), which are found in the outer regions of the thick peptidoglycan cell wall.

Particular external factors and the physiological state of the host appear to determine the adsorption efficiency of a phage to its host. For example, many phages require specific cofactors such as calcium or magnesium for attachment (147, 148), while the coliphage Lambda receptor is only expressed in the presence of the sugar maltose (149). Moreover, studies have shown that bacteria can develop resistance to specific phages through mutational loss or alteration of receptors (134, 150, 151). Phages in turn are capable of adapting to altered receptors or may even adapt to attach to a different receptor (152, 153).

1.5.2 DNA injection

Following attachment of the phage to its host, the next step is injection of the viral genome through the cell envelope and into the cytoplasm of the host. This event can be achieved in different ways, for example by tail contraction or the release of proteins which cause the bacterial cell membranes (in the case of a Gram-negative host) to become distorted/fused (141). Once the phage releases its viral genome from its capsid, it remains attached to the cell wall of its host as a "ghost" phage. Distinct mechanisms of DNA injection have been described for different phages. For example, phage T4 which has a contractile tail and a cell-puncturing device in the BP (154) has the fastest rate for DNA transport as its 169 Kb genome is ejected in 30 seconds (155, 156). In contrast, even though phage T7 has a smaller genome (40 Kb) than T4, its DNA ejection lasts 9 - 12 minutes. The T7 viral proteins gp16 and gp15 control the amount of DNA that enters the cell by first delivering 850 bp of the genome into the cytosol. This DNA includes three promoters which are recognized by RNA polymerase which then 'pulls' the remainder of the genome into the cytoplasm due to transcription-mediated template transfer. Thus, transcription provides the driving force for the entry of the T7 genome (157, 158). Another example of a DNA injection mechanism is the genome transfer of T5 (121 Kb) which is performed in two steps. First, 8 % of the DNA is transferred to the cytoplasm (159) where, during a 4 minute pause, two proteins (A1 and A2) encoded by this fragment are synthesized. Following the synthesis of these proteins the remaining DNA is transferred (160). The exact mechanism by which these proteins catalyse transfer of the remaining DNA is unknown, and it has been suggested that the DNA binding protein A2 may pull the remaining DNA into the cytoplasm (161). The push-pull mechanism of the DNA injection process of the Gram-positive-infecting *B. subtilis* phage 29 has been well

studied. Phage 29 has a non-contractile tail and a 19 Kb linear genome which is injected into the host cell using a tightly regulated two-step energy-dependent process. Once the phage recognizes its entry site in the membrane of the host, it first pushes ~ 65 % of the phage genome with right-left polarity into the host cell at the expense of the packaging pressure built inside the viral capsid. Following this, synthesis of at least four viral proteins is expressed due to the early promoter present in the entered fragment. These proteins form a multiprotein complex which pulls the remaining DNA into the cell, thereby completing DNA injection. Unlike the other proteins forming the multiprotein complex, protein p17 appears to be the most essential protein for this second DNA entry step (162).

1.5.3 Lytic or lysogenic phage cycle

Following entry of the nucleic acid into the host cell, the viral genome can propagate itself following either a lytic or lysogenic phage cycle (163). Bacteriophages that only replicate following the **lytic cycle** are known as **lytic** or **virulent phages**; once the phage multiplication cycle is completed the virulent phage causes lysis of its host in order to release newly produced phage particles (or virions). During the **lytic cycle** of propagation, the phage 'hijacks' (partly or completely) the RNA transcription, DNA replication and/or protein synthesizing machinery of the bacterial cell to make phage components. Once new phage genomes have been made individual genome units are packaged into the phage capsid, followed by the attachment of the phage tail to this DNA-filled capsid (163). Following assembly, the completed progeny viral particles release two proteins, the holin and lysin, encoded by the lysis cassette. These enzymes weaken the bacterial cell envelope leading to lysis of the host cell, i.e. rupture of the cell membrane and cell wall, and consequently the mature, infective phages make their way into the surrounding environment. Cell lysis is achieved as large pores/holes are formed in the cell membrane following the accumulation of holin proteins within the cell membrane and spontaneous aggregation at a genetically determined time-point. These pores provide the lysin access to its substrate, the cell wall peptidoglycan (164, 165). The lysin binds to a specific carbohydrate component of the cell wall using its C-terminal domain which ensures efficient substrate binding, while its N-terminal domain harbours the catalytic activity (166). The holin and lysin proteins are involved in the timed release of newly synthesized phage particles. This process varies among different phages and can be analysed with phage latent period

assays (which determine the time after infection before new phages are released). The latent period for dairy phages appears to differ depending on the phage. For example, the latent period for lactococcal phages has been reported to range from 40 - 90 min, for phages infecting *St. thermophilus* the latent period was reported to vary from 20 - 40 min, while for phages infecting *Lactobacillus* a latent period of 30 - 100 min has been reported (93, 167).

Previously it has been shown that lytic infection of *L. lactis* subsp. *cremoris* (*L. lactis*) UC509.9 with two distinct phages, the P335-group phage Tuc2009 and c2-group phage c2, demonstrates a rather modest impact on host transcription, yet with a phage-specific host response to infection. This study revealed that most of the host gene transcription changes occur during the late infection stage, whereas during the immediate and early infection stages only relatively minor changes in host transcription occur. Host transcriptome alterations during infection appear to be specific to Tuc2009 or c2, with relatively few differentially transcribed genes shared between infection with these phages (168). Additionally, another study revealed the defence strategy of *L. lactis* subsp. *lactis* (*L. lactis*) IL1403 in response to infection by lytic phage c2. The host response to c2 phage-attack involves induction of membrane stress proteins, cell wall D-alanylation, maintenance of the proton motive force (PMF) and energy conservation. The phage presence appears to be sensed as an extra-cytoplasmic stress as the host response is strongly targeted to the cell wall, thus affecting membrane integrity (169).

Bacteriophages that are able to replicate using both the lytic and lysogenic cycle (163) are known as temperate or lysogenic phages. The major difference between these two phage replication cycles is that in the lytic phage cycle, the phage DNA does not undergo integration into the host genome, instead the viral genome exists as a separate molecule within the cytoplasm of the bacterial cell and replicates independently from the host bacterial DNA. This differs from the lysogenic cycle which involves the integration of the phage genome into the bacterial chromosome, following entry of the viral nucleic acid, resulting in a prophage. During the lysogenic cycle the bacterial host cell remains alive and undergoes its natural life cycle, during which the phage genome replicates as part of the bacterial host chromosome. However, at any stage the integrated phage can exit the bacterial chromosome entering the lytic cycle, thus replicating/proliferating and subsequently killing the infected bacterial cell (163, 170). The lysogenic state of a phage is highly evolved as a

coevolution event between the phage and its host, frequently conferring advantages to the host (18). The presence of a temperate phage in a bacterial genome can lead to significant changes in the properties of the host such as resistance to infection by other phages and antibiotics, as well as the acquisition of restriction-modification systems (167, 171-174).

The lysogeny module of a temperate phage encodes genes involved in the process of genome insertion into the bacterial chromosome and maintenance of the temperate state. The lysogeny module encodes an integrase which facilitates the site-specific recombination between the genomes of the phage and host at specific homologous attachment sequences known as *attP* (phage) and *attB* (bacterium) (175). For phages infecting LAB, the site-specific integration system was first identified in phage LC3, a member of the P335 lactococcal group (176), homologous systems have since been identified in other P335 phages (82, 177). To maintain the integrated prophage the lytic genes of the temperate phage must be repressed in order to prevent the prophage from entering the lytic cycle. This process is controlled by one of the prophage-encoded proteins produced during the lysogenic state, a repressor protein which blocks transcription of phage genes required for lytic infection. This repressor system is responsible for the lysogenic state of many temperate phage genomes (178-180) including the majority of lysogenic LAB phages (181). For example, upon infection of the Lambda phage into *E. coli*, the phage encodes the transcriptional activator CII which is accumulated within the infected cell until a sufficient amount is present to activate the transcription of the repressor protein, CI, and the phage integrase. Following this, CI blocks the expression of genes required for lytic development by binding to an operator sequence within the intergenic region consequently blocking the transcription of the regulator protein, Cro, which prevents transcription of CI. The transcription of both CI and the integrase favours lysogeny allowing the phage genome to become incorporated within the host genome (179). The constitutive expression of these regulatory repressor proteins confers the lysogenised host with immunity to (secondary) infection by the same or related phages, this type of immunity is known as superinfection/repressor immunity (182, 183).

Previous studies have shown that prophages from Gram-negative bacteria encode a variety of distinct adsorption inhibition systems which alter the cell surface or other cell envelope components thus preventing attachment of homologous phages to the cell (184-186). For example, the D3 prophage from *Pseudomonas aeruginosa*

harbours a three-gene operon encoding proteins which modify the O-antigen of the cell surface LPS of its host. This results in an alteration of its serotype, which prevents infection by phages which use the O-antigen as a receptor, including phage D3 itself (184). The *E. coli* prophages 80 and N15 express the *cor* gene producing the Cor protein which physically interacts with FhuA, this interaction prevents the infection of phages that require FhuA as a receptor (185, 186). Prophages from Gram-negative and Gram-positive bacteria can also encode superinfection exclusion (Sie) proteins which interfere with the phage DNA injection process following phage adsorption to the cell (187). For example, the *E. coli* prophage HK97 harbours a gene (15), which encodes an inner membrane protein capable of blocking (secondary) HK97 DNA entry into the cytoplasm. Phage-resistance to other phages using the same host receptor as HK97 is not conferred suggesting that this Sie system does not inhibit phage adsorption but interferes with DNA injection (188). Similarly, the SieA system expressed by prophage P22, from *Salmonella typhimurium*, prevents the DNA entry of phages L, MG178, MG40 and P22 itself (189), while an Sie system expressed by coliphage P1 provides phage-immunity against coliphages c, c4 and P1 itself (190). Sie systems encoded by prophages of (Gram-positive) LAB species have also been reported including the well characterized Sie_{2009} encoded by Tuc2009, the resident prophage of the dairy starter strain *L. lactis* subsp. *cremoris* UC509. The sie_{2009} gene encodes a membrane-associated protein that allows phage adsorption to occur while preventing DNA injection of specific phages belonging to the lactococcal 936 group (32). Additionally, Sie systems encoded by prophages of the lactococcal strains *L. lactis* subsp. *cremoris* (*L. lactis*) MG1363 and *L. lactis* IL1403 have demonstrated a similar spectrum of activity against members of the 936 group of lactococcal phages (191). The *St. thermophilus* prophage TP-J34 expresses a small lipoprotein, Ltp, that provides phage resistance against the 936-group lytic lactococcal phage P008, and TP-J34 itself (192). This lipoprotein appears to block phage DNA injection by interfering with the process of DNA release from the head of the phage (32, 192). Another Sie system from an *St. thermophilus* prophage, Sfi21, is expressed by *orf203* resulting in phage resistance against heterologous streptococcus phages (193). Prophages constitute 10 - 20 % of a given bacterial genome consequently playing an important role in the evolution of bacterial species as prophages can contribute important biological properties to their bacterial host such as boosting virulence of their host (194). Lysogenic phages are widespread among dairy starter cultures, and although

prophages may be an advantage to bacterial hosts, they may also pose a significant threat to the fermentation process (195).

1.6 Structure of the bacterial cell envelope.

The cell envelope of bacteria is a key element to phage-host interactions as specific components of the cell envelope act as receptors for specific phages (see above). To survive unpredicted and often hostile environments, bacteria have developed a complex, multi-layered structure known as the cell envelope, whose components are synthesised at the cytoplasmic membrane. The function of the bacterial cell envelope is to determine and maintain the shape of the cell, to protect the cell from lysis by osmotic pressure, and to allow selective passage of nutrients from the outside environment and waste products from within the cell (196). Bacteria can be classified as either Gram-negative or Gram-positive depending on the structure of their cell envelope. Gram-negative bacteria (Fig. 1.3) possess a cytoplasmic membrane that is surrounded by a thin peptidoglycan layer and an outer membrane containing LPS, whereas Gram-positive bacteria (Fig. 1.4) are surrounded by a thick peptidoglycan layer and lack an outer membrane (18). The cell envelope of bacteria has been most thoroughly studied in the Gram-negative *E. coli* (197-199) and Gram-positive bacterium *B. subtilis* (200, 201) and *Staphylococcus aureus* (*S. aureus*) (202, 203), with some studies also involving Gram-positive LAB (142, 204, 205).

1.6.1 The cell envelope of Gram-negative bacteria.

The structure of the Gram-negative cell envelope was first clearly observed in 1969 by Glauert and Thornley (206), which revealed that the Gram-negative cell envelope is composed of an outer membrane, a peptidoglycan cell wall and a cytoplasmic or inner membrane. The aqueous cellular compartment in the space between the inner cytoplasmic membrane and the outer membrane was first named the periplasm in 1961 by Peter Mitchell (197), which was isolated in 1967 with its distinct set of proteins being characterized (198). Furthermore, in 1968 and 1972 the composition of the inner and outer membranes were characterized (199, 207).

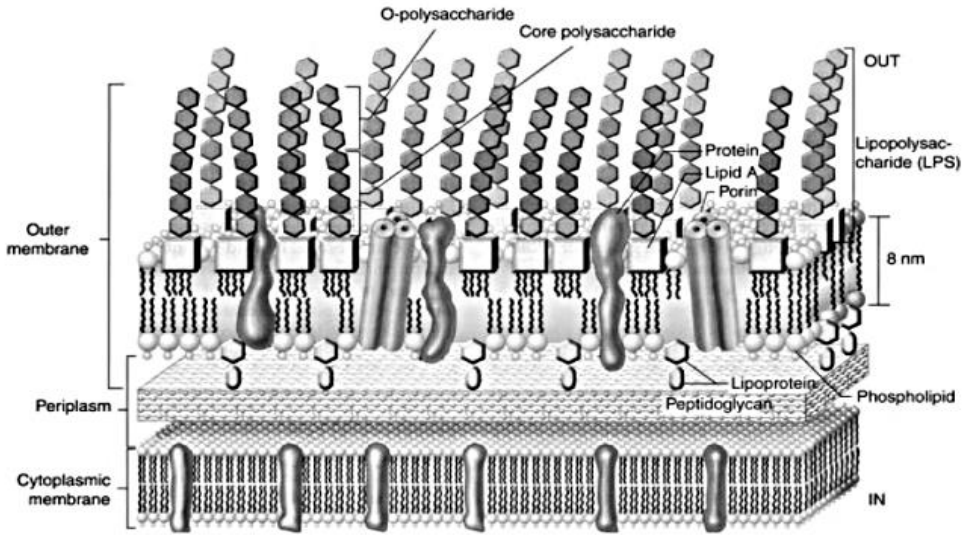


Figure 1.3: The structure of the cell wall of Gram-negative bacteria (18).

1.6.1.1 The outer membrane.

The outer membrane of the cell envelope of Gram-negative bacteria is the first layer encountered starting from the outside and proceeding inward. This is a key bacterial structure in phage-host interactions as the outer membrane harbours porins, affinity transporters and LPS which have been shown to act as Gram-negative host receptors for phage infection (141, 150, 208-210). The outer membrane is essential for survival of the cell and represents a protective barrier making Gram-negative bacteria generally more resistant to antibiotics and harsh environments when compared to Gram-positive bacteria (196). Many antibiotics target intracellular processes of bacteria such as replication, nucleic acid and protein biosynthesis, protein folding, and cell division, and therefore must penetrate the cell envelope in order to be effective (211). The outer membrane contains phospholipids, which are essentially confined to the inner leaflet, and glycolipids, principally LPS (212), which are confined to the outer leaflet. Most proteins of the outer membrane can be divided into lipoproteins or β-barrel proteins. Lipoproteins contain lipid moieties which are attached to an amino-terminal cysteine residue (213), and which allow such lipoproteins to become embedded in the inner leaflet of the outer membrane. *E. coli* contains one hundred distinct outer membrane lipoproteins, although the function of most of these are unknown (214). Most of the integral, transmembrane proteins of the outer membrane adopt a β-barrel conformation. These outer membrane proteins include the OmpF and

OmpC porins (215), which allow passive diffusion of amino acids and small molecules such as mono- and di-saccharides across the outer membrane. Together, these porins are present at ~ 250,000 copies per cell. Other outer membrane porins include LamB (216), which is responsible for the diffusion of maltose and maltodextrins, and PhoE (215), which allows the diffusion of anions such as phosphate across the outer membrane. Another abundant outer membrane porin is the outer membrane protein A (OmpA) which is monomeric and can exist in two different conformations (217), meaning that it exists either as an active or inactive porin. Outer membrane porins larger than β-barrels are also present but at much lower levels and are responsible for high affinity transport of large ligands such as Fe-chelates or vitamins such as vitamin B-12 (218). The outer membrane of *E. coli* contains a small number of enzymes including phospholipase (PldA) (219), a protease (OmpT) (220) and an enzyme that modifies LPS (PagP) (221). LPS consist of a hydrophobic domain, a nonrepeating oligosaccharide core and an extended polysaccharide (PS) chain known as the O-antigen (222). LPS in the outer membrane protects the cell from hydrophobic molecules and porins control the diffusion of hydrophilic molecules larger than ~ 700 Daltons, making the outer membrane of Gram-negative bacteria a very effective and selective permeability barrier (218).

1.6.1.2 The peptidoglycan cell wall.

The peptidoglycan layer consists of chains, each of which consist of a polymer of the disaccharide N-acetyl-glucosamine (GlcNAc) and N-acetyl-muramic (MurNac) acid, and which are crossed-linked by short peptides to make a two-dimensional sheet (223). The glycan chains of the peptidoglycan layer appear to be located perpendicular to the long axis of rod shaped bacterial cells (224), therefore determining the cell shape due to its rigidity. The peptidoglycan layer in Gram-negative bacteria is attached to the outer membrane by lipoproteins known as Lpp, murein/Braun's lipoproteins (225). This protein is the most abundant protein in *E. coli* with more than 500,000 molecules per cell. Moreover, the porin OmpA is important for outer membrane stability as OmpA is responsible for non-covalently association of the outer membrane peptidoglycan to the rest of the bacterial cell envelope (226).

1.6.1.3 The periplasm

The periplasm which is only found in Gram-negative bacteria is an aqueous cellular compartment located in between the outer membrane and inner membrane, and is densely packed with proteins as demonstrated by a higher viscosity compared to that of the cytoplasm (227). The role of the periplasm is to sequester potentially harmful degradative enzymes such as RNAse which can interfere with DNA replication (228). The first enzyme protein to be identified in the periplasm was alkaline phosphatase. Alkaline phosphatase cleaves inorganic phosphate (required for cell growth) from external phosphomonoesters for transportation to the cytoplasmic membrane (229, 230). In addition, many periplasm transport proteins responsible for the transport of substrates (e.g. sugars, amino acids, peptides, metals, etc.), chemotaxis and chaperon-like molecules which play a role in envelope biogenesis are found in the periplasm (231). Other functions of the periplasm involve protein folding and oxidation, important signalling functions, cell division regulation, and secretion of lipoproteins, glycerolphospholipids and uniquely structured β-barrel proteins to the outer membrane (232).

1.6.2 The cell envelope of Gram-positive bacteria.

The first cell envelope structure of a Gram-positive bacteria, that of *Bacillus cereus*, was visualized by electron microscopy by Chapman and Hillier in 1953 (233). As mentioned above, Gram-positive bacteria do not possess an outer membrane, which plays a role in protecting Gram-negative bacterial cells from the environment, and also indirectly stabilizes the inner membrane of Gram-negative organisms, thus explaining the presence of just a thin peptidoglycan layer in the latter bacterial group. In contrast, the lack of an outer membrane in Gram-positive organisms has led to the evolution of a 30 – 100 nm thick peptidoglycan layer containing many sheets of peptidoglycan polymers that surround the cell in order to withstand harsh environmental conditions (196). It has been shown that in *B. subtilis* changes in environmental and growth conditions lead to a dramatic change in the composition of surface-expressed proteins, illustrating the role of the cell envelope in adapting to new environments (200). The four main components of a Gram-positive cell envelope outside of the cytoplasmic membrane are peptidoglycan (also known as murein), lipo- and/or wall-teichoic acids (LTA or WTA, respectively), PS and, though not present in

all bacteria, an S-layer. Previous studies have demonstrated that all of these components, are recognized by specific phages as target host receptors for phage infection (83, 87, 133, 135, 143-145, 234, 235), and therefore understanding of the structure and composition of each of these cell envelope components is necessary in order to gain a better understanding of how phages interact with their specific hosts.

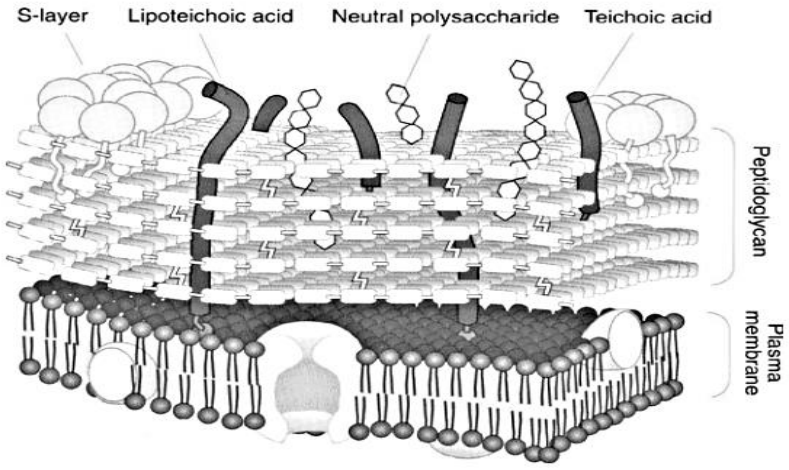


Figure 1.4: Cell envelope structure of Gram-positive lactic acid bacteria (NB. The CWPS structure is not indicated in this image) (204).

1.6.2.1. The peptidoglycan layer

The peptidoglycan layer, also known as sacculus, is composed of a PS with individual covalently cross-linked chains. The sacculus is therefore a single huge, covalently-linked macromolecule which is made up of parallel glycan threads connected by short elastic peptide crosslinks which reinforce this structure (196). The peptidoglycan layer has the ability to expand and contract when necessary with extraordinary elasticity and tensile strength. Thus, molecules can enter and exit the cell as required by the plasma membrane (236). It has been shown for the cell wall of *E. coli* and *B. subtilis* that each peptidoglycan layer allows bacterial cells to exclude molecules larger than 55 KDa (237).

In LAB, like in all eubacteria, the basic components of a glycan thread in the peptidoglycan envelope are GlcNAc and MurNAc resulting in a polymer of disaccharide N-acetyl-glucosamine-β(1-4)-N-acetyl-muramic acid. During peptidoglycan synthesis GlcNAc and/or MurNAc may undergo three different modifications, following polymerization and incorporation into the peptidoglycan

envelope. The three modifications include the following; (i) Studies involving bacilli (238, 239) and *Lactobacillus fermentum* (*Lb. fermentum*) (240) have revealed that removal of the acetyl groups from GlcNAc and MurNAc causes increased sensitivity to autolysis and external lysis procedures (241, 242), as well as increased susceptibility to phages (243). (ii) Other studies involving species of Gram-negative and Gram-positive bacteria, including *S. aureus* and several LAB species such as *Enterococcus faecalis* (244), *Lactobacillus casei* (*Lb. casei*) (245), *Lactobacillus acidophilus* (246) and *Lb. fermentum* (247) have demonstrated that MurNAc residues in the peptidoglycan envelope can be 6-O-acetylated (i.e. an acetyl is added to the C-6 hydroxyl group of MurNAc residues). O-Acetylation of the MurNAc moiety of peptidoglycan is typically associated with bacterial resistance to lysozyme and in some bacteria, such as *Bacillus anthracis*, this process is combined with N-deacetylation to confer resistance of cells to lysozyme (248). (iii) C6 of MurNac has also been reported to be substituted by cell WTA and teichuronic acids which will be further discussed below (204).

1.6.2.2. Teichoic acid (TA)

TAs, first discovered in 1958 in the LAB *Lb. plantarum* (249), are present in most cell envelopes of Gram-positive bacteria comprising half of the weight of the cell envelope (204). TAs are synthesized on the outer face of the cytoplasmic membrane as long polymers which thread through the peptidoglycan sacculus. These polymers show distinctly different molecules and are anchored in different positions in the cell envelope classifying them into cell wall teichoic acids (WTAs), lipoteichoic acids (LTAs) and lipoglycans. Depending on, among others, the bacterial species or strain, stage or rate of growth of the cell, the pH of the medium and carbon source it is exposed to, and phosphate availability, TAs will vary in structure and abundance (250-253).

1.6.2.3. Cell wall teichoic acid (WTA)

Cell WTAs include poly-glycerol/ribitol phosphate TAs which are covalently attached to the peptidoglycan cell wall. These molecules share the same overall structure composed of a chain of phosphodiester-bound glycerol or ribitol residues attached to a “linkage unit”, which allows the covalent attachment to the peptidoglycan

(204). In LAB, despite the wide variety of TAs, the structure of the linkage unit is well conserved. It is composed of the disaccharide *N*-acetylmannosaminyl*β*-(1-4)glucosamine followed by glycerol phosphate, however, differences may occur in the nature of the molecules (glycosyl or D-alanyl) which decorate the main chain of the TAs and number of glycerol phosphate residues (254). A poly (glycerol phosphate) TA has been identified as an important carbohydrate component of the cell wall of the LAB strain *L. lactis* IL1403 (255). However, not all LAB appear to contain WTAs, such as *Lb. casei* (256) and *L. lactis* subsp. *cremoris* (257), which relies on LTAs for the anionic character of the cell envelope. Teichuronic acids are phosphate-free and the main chain is composed of sugar monomers directly linked by glycosidic bonds (204). The nature of the sugar monomers varies according to the species studied. Teichuronic acids have been described for several *Bacillus* species (258, 259). Considering the phylogenetic relatedness of members of the genus *Bacillus* to LAB, it is possible that teichuronic acids are present in the cell wall of LAB as well.

1.6.2.4. Lipoteichoic acid (LTA)

Unlike WTAs, polymers known as LTAs and lipoglycan remain attached to the cytoplasmic membrane through a lipid anchor, however, a fraction of them are found free in the cell wall or even released into the medium (260). LTAs were first discovered in 1970 in the LAB *Lb. fermentum* (261). In LAB, LTAs are widely distributed and have been found in lactobacilli, lactococci, enterococci, leuconostocs, and streptococci (252). The main component of the hydrophobicity of the cell envelope is provided by LTAs which may be involved in bacterial adhesion (204). The main chains of cell WTAs and LTAs are decorated with glycosyl or D-alanyl substituents which greatly influence the functionality of these anionic polymers. Altering the substitution process of glycosyl or D-alanyl through genetic means has been shown to cause changes in the phenotype of the mutant strains such as phage resistance (145, 262-264). TAs are potent immunogens which are able to weakly stimulate cytokine synthesis, a process inhibited by removal of the D-alanine substituents. Additionally, TAs are involved in the functionality of the cell, for example when phosphate is scarce WTAs composed of polymers of glycerol- or ribitol-phosphate act as a phosphate source (265). Furthermore, cell WTAs are believed to maintain the correct ionic environment for cation-dependent membrane

systems as these anionic polymers provide a constant reservoir of bound cations (266). Moreover, LTAs (267), TAs and teichuronic acids (268, 269) have been shown to influence the activity of autolysins which are required by the lytic phage cycle to induce lysis of the bacterial host in order to release virions. Interestingly, TAs have been shown to act as receptors for phage adsorption (133, 145, 146), which will be further discussed in Section 1.8: Phage-host interactions of Gram-positive bacteria.

1.6.2.5. Cell wall polysaccharides (CWPS)

The cell envelope of many Gram-positive bacteria is known to contain or is covered by CWPS, which have been classified into: (i) capsular PS which decorate the outside of the cell envelope and are either covalently bound (246) or loosely associated (270) with the peptidoglycan, and (ii) extracellular polysaccharides (EPS) which do not remain attached to cells and are released into the cell surroundings (271). CWPS are complex structures which vary according to the nature of sugar monomers, modification (e.g. acetylation and reduction), mode of linkage, branching, and substitution. The structure has also been shown to depend on growth conditions (272). In several LAB, rhamnose has been shown to be a constituent of CWPS, in strains including lactobacilli (246, 272-274), streptococci (275), and lactococci (257). Furthermore, in *L. lactis* strains a novel CWPS, the so-called pellicle (PSP), has been reported (205). Moreover, studies have shown that specific CWPS structures act as receptors for the adhesion of specific phages (140, 142, 144, 257, 274, 276-278) which will be further discussed in Section 1.8: Phage-host interactions of Gram-positive bacteria.

1.6.2.6. S-layer

A variety of proteins are also present at the surface of the peptidoglycan layer, some of which are analogous to proteins located in the periplasm of Gram-negative bacteria (279). These proteins may contain membrane-spanning helices; be attached to a lipid anchored in the membrane; bind to TAs; or be covalently attached to or associated tightly with the peptidoglycan layer (280). The S-layer is a surface layer composed of protein subunits which are non-covalently, yet strongly anchored to the underlying peptidoglycan. The function of the S-layer is unknown but studies have shown that it might be involved in virulence in the case of pathogenic species (281)

and/or in adhesion (282). With regards to LAB, the S-layer has only been identified in the cell wall of certain members of the genus *Lactobacillus* (283-285).

1.7 Phage-host interactions of Gram-negative bacteria.

As mentioned above, the first step involved in the interactions of a phage with its specific host is the recognition of and adsorption to a specific cell surface-exposed, host-encoded receptor. Adsorption of the phage to its host may require binding of the phage to one or more host encoded receptor(s) in a reversible or irreversible manner. Currently, the adsorption process of phage to host-encoded receptors in certain Gram-negative bacteria has been studied in detail. As mentioned above, Gram-negative host receptors are present in the outer membrane which contains surface-exposed proteinaceous or saccharidic structures, such as pili, flagella or LPS. The most common Gram-negative outer membrane components that act as host receptors for phages are porins, affinity transporters and LPS (141). Following reversible binding to a primary receptor, irreversible binding to a secondary receptor occurs, which triggers DNA release, resulting in phage DNA transfer through the bacterial cell envelope into the host cell cytoplasm where phage DNA replication can begin. Though the process of viral DNA injection is reasonably well understood for a small number of phages infecting Gram-negative bacteria (286).

The DNA injection process of Gram-negative bacteria has been shown to vary among morphologically similar phages. For example, phage Lambda requires host proteins to orchestrate the formation of the channel required for DNA translocation (287), whereas phage T5 relies on its own phage-encoded proteins to form the channel for DNA transport (288). Previous Gram-negative bacterial studies have shown that specific bacterial gene products are involved in phage genome injection as specific phages require either the activity of a host membrane-associated protein (151) or the protein itself to trigger viral DNA ejection (208). For example, *in vitro* assays with purified LamB from *E. coli* have shown that this outer membrane protein, responsible for producing transmembrane channels, inactivates phage Lambda DNA injection, thereby providing the first evidence of the involvement of this protein in DNA transfer. It has not only been shown that the outer membrane protein LamB acts as a phage receptor for the *Siphoviridae* phage Lambda, but it has also been revealed that the bacterial inner membrane PtsM complex, responsible for the transfer and phosphorylation of sugars, is involved in the phage infection process (208). Similarly,

the protein FhuA, an outer membrane-bound ferrichrome transporter of *E. coli*, acts as a receptor for the *Siphoviridae* coliphage T5 (289). Furthermore, other *E. coli* phage receptors including LPS or porin proteins (e.g. OmpC) are recognized by coliphages belonging to the *Myoviridae* T4 superfamily, where this particular phage-host interaction occurs as a reversible event (290). Additionally, phages C1 and C6 have been shown to first adsorb to the *E. coli* proteins DcrA (inner membrane protein) and DcrB (periplasmic protein), respectively, where such an initial adsorption appears to trigger formation of diffusion channels across the outer membrane of *E. coli*. Following this, a secondary adsorption to the outer membrane protein receptors BtuB and FhuA occurs, which in turn triggers phage DNA injection (150). Similarly, coliphage T1 irreversibly binds to the outer membrane protein FhuA, but it requires the activity of the inner membrane protein TonB for successful DNA injection (209, 210).

Moreover, phage HK97 encodes a gp15 moron element, within its morphogenesis region, which encodes a novel Sie protein (as previously mentioned). The HK97 prophage actively expresses gp15 which is responsible for the resistance of the bacterial host cell to phages HK97 and HK75. Furthermore, the construction of a HK97 hybrid phage showed that the TMP is involved in the susceptibility of HK97 to inhibition by gp15, demonstrating that gp15 acts to prevent the entry of HK97 DNA into the cytoplasm of the infected cell (188). A model for the *in vivo* genome injection process of the temperate phage HK97 has been proposed based on a study involving single deletions in *E. coli* genes predicted to encode membrane-associated proteins. Results revealed that the activities of the protein PtsG (a membrane-associated glucose transporter protein) and FkpA (a periplasmic chaperone) act as host receptors for phage HK97. A potassium efflux assay, which is used to detect phage-membrane penetration, confirmed that deletion of these specific genes prevented the passage of the phage genome through the cell envelope (160, 291). This study did not only identify FkpA and PtsG as host receptors required for phage injection, it also genetically mapped the determinants for these requirements to the HK97 phage TMP. Additionally, the specific region of the HK97 TMP that is sensitive to DNA injection inhibition by the gp15 superinfection exclusion protein was identified. Construction of hybrid HK97 phages carrying TMP regions from other phages revealed that the TMP region required for PtsG interaction is also involved in gp15-mediated superinfection exclusion (151).

Overall, these studies demonstrate that even within a common Gram-negative bacterial host different host-encoded receptors can be involved in DNA injection by specific bacteriophages. Furthermore, the fact that different phages require distinct inner membrane proteins as host receptors for phage infection shows that such receptors are involved in the genome entry process of phage. These proteins may either act as channels to allow the viral DNA to pass through the cell envelope, or may anchor a region of the TMP of the phage to the inner membrane of the host cell allowing efficient passage of the viral genome through the cell wall into the cytoplasm of the infected cell.

1.8 Phage-host interactions of Gram-positive bacteria.

Unlike Gram-negative bacteria, the extent of knowledge regarding phage interactions with their Gram-positive bacterial hosts is much more limited. In order for a phage to successfully infect its Gram-positive bacterial host it must first attach/adsorb to exposed receptors and penetrate the cell wall to allow the viral DNA to enter the cytoplasm of the infected bacteria cell. The process by which this occurs is not well understood. Siphophage SPP1 and the interaction with its *B. subtilis* host has been studied in considerable detail, allowing this phage to become a model for the process of DNA infection for phages infecting Gram-positive bacteria. It has been reported that the protein encoded by the *B. subtilis* gene *yueB* is involved in irreversible binding of SPP1, following an initial reversible binding to a carbohydrate moiety on the cell surface. This was demonstrated by introducing deletions within the *yueB* gene which led to the formation of a *B. subtilis* mutant strain which was resistant to SPP1 (292). Following this particular study the extracellular region (ectodomain) of the YueB protein (a 36.5 nm long fiber) was purified using metal affinity and size exclusion chromatography (293). Incubation of SPP1 virions with purified YueB ectodomain revealed that viral DNA ejection was triggered *in vitro* following irreversible binding of SPP1 to the ectodomain of YueB. SPP1 DNA ejection was confirmed by the reduced number of plaques and the decreased amount of phage DNA remaining in the capsids of SPP1 following the interaction of SPP1 virions with purified YueB ectodomain. This was the first receptor to be purified from a Gram-positive bacterium that was actively capable of triggering viral DNA ejection *in vitro*. The SPP1 bacteriophage possesses a long non-contractile tail composed of a tube and a tail tip acting as the adsorption device. The head-to-tail connector of SSP1 is

composed of gp6, gp15 and gp16 dodecameric rings, with the gp16 ring being responsible for closing the portal channel structure to prevent DNA exit from the nucleocapsid (294, 295). It has been suggested that one end of the dsDNA which is coiled within the capsid is bound to gp16, when ejection of the SPP1 genome is triggered, and that this is the first DNA segment to exit the SPP1 phage capsid (296). Additionally, a further study determined the structure of the SPP1 tail before and after DNA ejection, revealing extensive structural rearrangements during this process (61).

Only a small number of phages infecting Gram-positive bacteria have been shown to adsorb to a membrane-bound proteinaceous receptor. For example, protein purification and chromatography techniques have revealed that the membrane-associated protein PIP (phage infection protein), which is encoded by the Gram-positive *L. lactis* subsp. *lactis* C2 (*L. lactis* C2) is required for phage c2 infection (134). Moreover, irreversible binding of the phage to this membrane protein occurs following an initial and reversible adsorption to a cell wall carbohydrate (140).

Studies have shown that specific phages which infect Gram-positive bacteria recognize TA structures as receptors for phage infection (133, 145, 146, 262, 297). Furthermore, the modification of the WTA or LTA structures has been shown to result in phage-resistance (145, 298, 299). For example, the presence of a galactose-containing moiety in lipoteichoic acid prevents phage adsorption to *L. lactis* subsp. *cremoris* SK110 (145). Phage-resistant *S. aureus* mutants revealed that glycosylated cell WTAs act as a receptor for siphovirus 47, Sa2mw, 13, and 77, while non-glycosylated TA acts as a receptor for myovirus K and 812 (146). The major target for the model phage SPP1 is glucosylated TAs to which SPP1 first reversibly adsorbs to before irreversibly binding to YueB (133). Other phages infecting *B. subtilis* also adsorb to glucosylated polyglycerol TA, and thus phage-resistance has been reported in *B. subtilis* mutant strains deficient in glucosylation of their TAs (298). Furthermore, the lack of GlcNac in TA of *Listeria* mutant strains also results in phage-resistance (299). Additionally, phage P35h4 was previously shown to directly adsorb to galactosylated WTA in *Listeria monocytogenes* (300).

1.9 Lactococcal phage-host interactions

Lactococcal phage-host interactions are (like any other phage-host system) determined by two factors, the phage-encoded RBPs and the host-encoded target receptor molecule(s), which may be a saccharidic, proteinaceous and/or TA

component of the cell envelope (141). It had previously been revealed that rhamnose present in the CWPS in *L. lactis* subsp. *cremoris* KH acts as a receptor for the 936 bacteriophage kh (257). Furthermore, it was shown that seven c2-group lactococcal phages infecting *L. lactis* C2 adsorb to a rhamnose component of the CWPS (140). Fluorescence quenching experiments demonstrated that the RBP of the P335-group phage TP901-1 displays high affinity for saccharides in solution (301). Further phage-host recognition and attachment studies using phages belonging to different lactoccocal phage groups also suggest that their primary receptor is saccharidic in nature. These phages include P335 phages Tuc2009 (278) and TP901-1 (132), 936 phages p2 (302) and bIL170 (303), and the 949, P087 (144), and 1358 phages (142, 304).

Over a decade ago it was shown that infection of the 936 phages bIL170 and 645 requires the *L. Lactis* IL1403 *rgpE* and *ycaG* genes, and the *L. lactis* subsp. *cremoris* Wg2 *ycbB* and *ycbC* genes, respectively. The *rgpE* and *ycbB* genes encode putative glycosyltransferases, while *ycbB* and *ycbC* encode putative membrane-spanning proteins with unknown functions for phage adsorption. Interestingly, these genes were predicted to be involved in CWPS biosynthesis, suggesting that both phages bind to different CWPS structures (305). The structure and biological functions of the CWPS have previously been investigated in certain lactococcal strains (255, 306-311). The lactococcal CWPS consists of two covalently-linked constituents: the surface-exposed polysaccharide pellicle (PSP) and the recently described rhamnan PS, which is embedded within the peptidoglycan layer. Analysis of the structural components of the cell envelope of *L. Lactis* MG1363 reported the presence of a novel layer forming an outer pellicle surrounding these Gram-positive cells (205, 307) and acting as a host receptor of (certain) 936-group phages (205). Such a PSP structure was also reported in *L. lactis* subsp. *cremoris* strains 3107 (309) and SMQ-388 (142). A bioinformatic study based on the comparative genome analysis of six lactococcal genomes revealed that the CWPS-specifying DNA region contains a highly conserved region, as well as a region of diversity. This led to the classification of *L. lactis* CWPS-associated genotypes into three major groups (types A, B and C) based on the differences in the gene cluster that is presumed to be involved in the CWPS pellicle biosynthesis. Moreover, a direct correlation was reported between the CWPS pellicle genotype of a given *L. lactis* host strain and the host range of tested 936-group phages, with 936 phages infecting lactococcal strains belonging to the B- and C-type CWPS.

The comparative genome analysis of the eleven lactococcal 936-group phages used in this particular study also revealed a direct correlation between the 936-group phage-encoded RBP genotypes and the host-determined CWPS genotype (143). The variable genetic locus that is responsible for the biosynthesis of the PSP has been identified and partially characterized in a number of strains (77). The subtle differences in the biochemical make-up of PSP structures exhibited by different *L. lactis* strains appears to act as receptors to determine host specificity of many lactococcal phages (309). The variable PSP has been identified as the principal phage receptor for the industrially significant lactococcal phage groups P335 and 936, as well as the 1358, 949 and P087 (142, 144, 278, 309).

The genetic locus encompassing the CWPS biosynthesis operon of eight strains of *L. lactis* belonging to the C-type CWPS genotype have been analysed (309) (Fig. 1.5). The results revealed the presence of a variable region among the C-type CWPS biosynthesis cluster which contain genes encoding glycosyltransferases that display low or no sequence homology between the strains leading to a sub-classification of this C group into five subgroups, subtype C_1 – C_5 (309). As recently discussed CWPS is composed of the 'pellicle' and the rhamnan component. In strains exhibiting an A- and B-type CWPS, this structure is isolated as a single polymer (255, 311), while in strains exhibiting a C-type CWPS, the pellicle and rhamnan are isolated separately due to the chemically labile phosphodiester linkage (309). Based on the variable regions (PSP) of the C-type CWPS biosynthesis cluster (Fig. 1.5) the *L. lactis* strains 3107 and MG1362 have been classified as subtype C_2 and C_1, respectively (309). Genetic engineering was employed to achieve swapping of the sequence encoding the PSP structures, and results demonstrated that the subtype C_2 PSP is the host cell surface receptor of the LC3 and TP901-1 P335-group phages and that various 936-group phages require the PSP encoded by subtype C_1 as their host receptor to accomplish adsorption. Moreover, the CWPS of *L. lactis* SMQ-388 strain was also identified as a C-type strain (142). Additionally, the CWPS of *L. lactis* IL1403 was identified as a type B (255) and the CWPS of *L. lactis* UC509.9 was identified as a type A (311). The unique CWPS structures of these strains validate the genetic diversity of the encoding operon (Fig. 1.5).

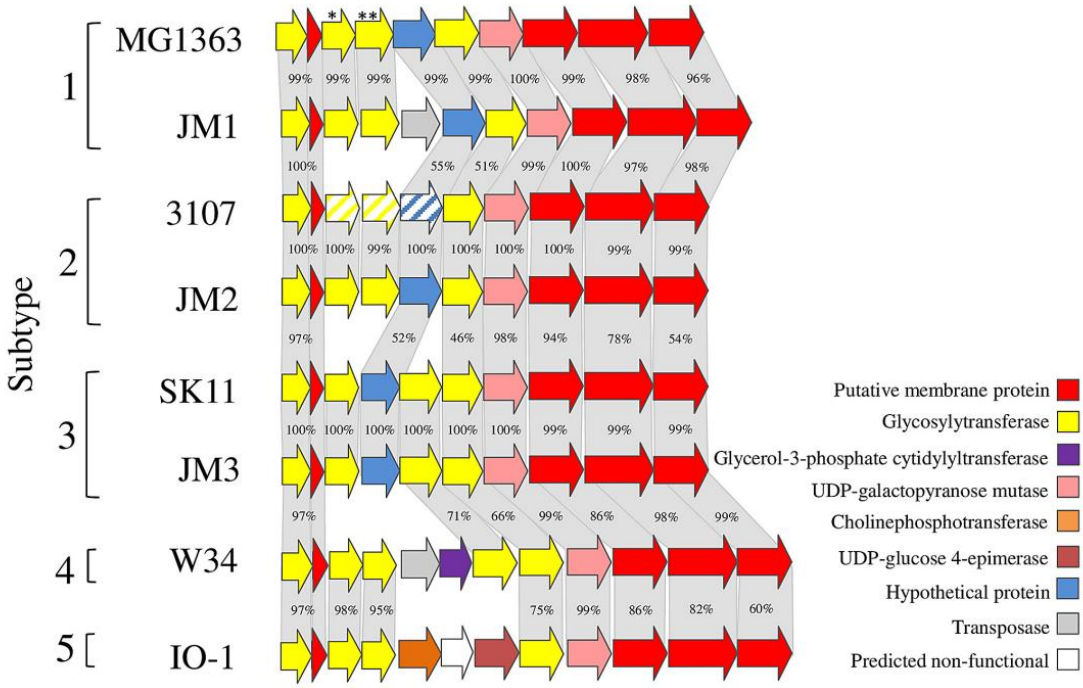


Figure 1.5. Comparison of the variable regions in the C-type CWPS biosynthesis cluster of eight lactococcal strains (309).

Lactococcal phages belonging to the P335 and 936 groups have been extensively studied due to their significant negative impact on the fermented dairy industry (83, 143). For example, by ethyl methanesulfonate exposure five bacteriophage insensitive mutants (BIMs) of *L. lactis* 3107 were generated that display different infection sensitivities to the P335-group phages LC3 and TP901-1 (312). All five mutants were resistant against TP901-1 infection, while two of these BIMs were also resistant to LC3. Furthermore, the adsorption efficiencies of TP901-1 and LC3 for both LC3-resistant mutants were severely reduced, which suggest that their CWPS biosynthesis was affected. Interestingly, near normal adsorption properties were exhibited by the three TP901-1-resistant BIMs (which had remained sensitive to LC3). This particular study suggests that both P335-group phages TP901-1 and LC3 use the same primary receptor. This conclusion was reinforced by the fact that both P335 phages require the CWPS produced by *L. lactis* 3107 (Fig. 1.5) as a host receptor (309) and the fact that the RBPs of both phages are almost identical at their respective C-terminal regions (80, 81). This particular study also suggests that TP901-1 and LC3 may adsorb to different molecules to trigger DNA injection (309). Further research involving bioinformatic analysis of these phage-resistant mutants along with

complementation assays may thus reveal novel gene(s) involved in phage-host interactions, hence expanding our knowledge regarding Gram-positive genes involved in lactococcal P335 phage infection.

1.10 Classification of phage members of the P335 group

Significant scientific attention has been directed towards the lactococcal P335 phage group as phages belonging to this particular group have frequently been isolated in the dairy industry and are the most genetically diverse. During the past decade P335 phages Tuc2009 and TP901-1 have been thoroughly studied at the molecular and structural level, and have consequently become model systems for P335 phage group infecting *L. lactis* (99, 100, 123, 131, 137, 139). The number of complete P335 lactococcal phage genome sequences (excluding predicted prophages belonging to sequenced *L. lactis* genomes) on the NCBI public database (www.ncbi.nlm.nih.gov) is rapidly increasing, some of which are exhibited in Table 1. A common characteristic among lactococcal phages appears to be their genome plasticity which is a disadvantage to the dairy industry as it allows phages to rapidly adapt to the dynamic dairy environment (135, 313). Southern hybridization and morphological analyses followed by comparative sequence analysis confirmed that phages Tuc2009, TP901-1, BK5-T, r1t, LC3, ul36 and 4268 all belong to the P335 phage group (57). Previous genome analyses revealed that among the P335 phages the genome sequences are rather divergent relying on the identification of sequence homology or analogous functional regions to classify members of the P335 group (57, 69, 79-81, 234). Identification of lactococcal P335 phage members has relied on the presence of a predicted dUTPase-encoding gene believed to be the only conserved genes among P335 phage genomes (69). However, the genome of the P335 phages 4268 and Dub35A do not exhibit a homologue of the dUTPase-encoding gene illustrating the polythetic nature of the P335 phage group (76). Currently, among all the phage genomes classified within the lactococcal P335 phage group there is not a specific genetic feature shared by all members (57). Previous functional analysis of TP901-1 confirmed the suggestion that the functional synteny of coliphage lambda is a model comparator to lactococcal phages (86). This suggests that despite sequence divergence among the morphotype phages belonging to the *Siphoviridae* family, these phages can possess similar genomic regions encoding their morphogenetic functions.

Interestingly, the recently isolated P335 phages Q33 and BM13 were shown to be genetically distinct from previously sequenced P335 phage isolates (83). As a result morphological analysis combined with comparative genomic analysis of the genomic region encoding the "adhesion device" of P335 phages was perfomed. This led to the classification of these genetically diverse phages into four sub-groups (I – IV) based on their distinct morphological features (Figs. 1.6 and 1.7) and genetic lineages of their tail tip region which is presumed to be involved in interactions with the lactococcal host. Furthermore, a more recent study based on the genetic diversity of lactococcal prophages identified a novel morphotype of the P335 lactococcal phage group, thus adding an additional sub-group (V) to the classification of this P335 group (314). Examples of published P335 phages, the sub-group they belong to, and the CWPS type they infect are demonstrated in Table 1. **Subgroup I** represents phages possessing a long fibre protruding through their tail tip. Previously, BK5-T and 4268 were the only members of this sub-group (83). However, a recent study identified the novel 38502 and C41431 phages as additional members of sub-group I due to their BK5-T-like RBP structures determined by EM and sequence similarities in the region encoding the adhesion device (77). **Sub-group II** characterizes phages with a double-disc BP structure, representatives of which include the model P335 phages Tuc2009, TP901-1, P335 and ul36. Other novel phages, shown in Table 2, exhibited a BP structure similar to that of Tuc2009 and TP901-1 (77). **Sub-groups III and IV** represent phages that exhibit a small "stubby" distal tail structure but which are genetically distinct. Sub-group III phage members include LC3 and r1t, including the recently assigned new phage member Dub35A which is highly identical to LC3 in the genomic region encoding the adhesion device. The P335 phage 50902 was classified as a sub-group IV member due to its shorter tail and small "stubby" tail tip which is typical of members of sub-group III or IV, further comparative genomic analysis based on overall proteomic content revealed sequence homology to the sub-group IV phage members (77)**. Sub-group V** represents P335 phages which possess TMPs of different lengths, well conserved Dit proteins and large RBPs (314). Large Dit proteins also known as evolved Dits have been shown to harbour carbohydrate binding domains in phages belonging to sub-group I and V. The RBPs of the phages belonging to sub-group V are comparable in size to that of phage BK5-t which belongs to sub-group I. Sub-group V phage members appear to encode a protein with fused Tal and RBP (Tal-RBP), this is consistent with BK5-t as distinct genes encoding the RBP and Tal

component are not observed in this sub-group I member. The combined Tal-RBP function has been shown to be limited to sub-group I and V phages as the genomes of sub-group II - IV phages possess distinct Tal and RBP encoding genes. Interestingly, it has been revealed that like phage BK5-t, phage members of the sub-group V possess a long tail fibre protruding through the tail tip region. Furthermore, it was shown that even though RBP sequences of the phages belonging to sub-group I and V are unrelated, perhaps similar roles and structural features are conserved between these P335 phages (314). The classification of the phage members of the lactococcal P335 group (sub-group I - V) has provided a tremendous advancement in the current understanding of lactococcal phage-host interactions and the genetic complexity of this industrially significant lactococcal phage group, reinforcing the importance of understanding phages and their hosts at a genetic and structural level.

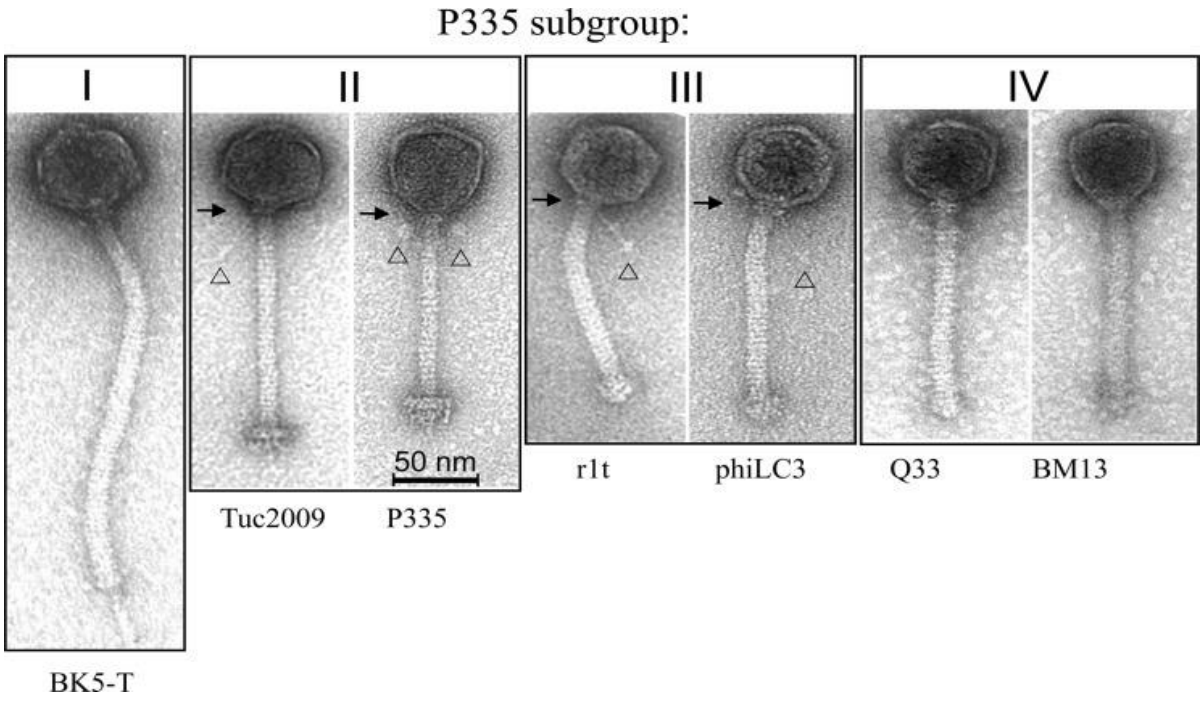


Figure 1.6. Transmission electron micrographs of lactococcal phages representing P335 subgroups I to IV (83).

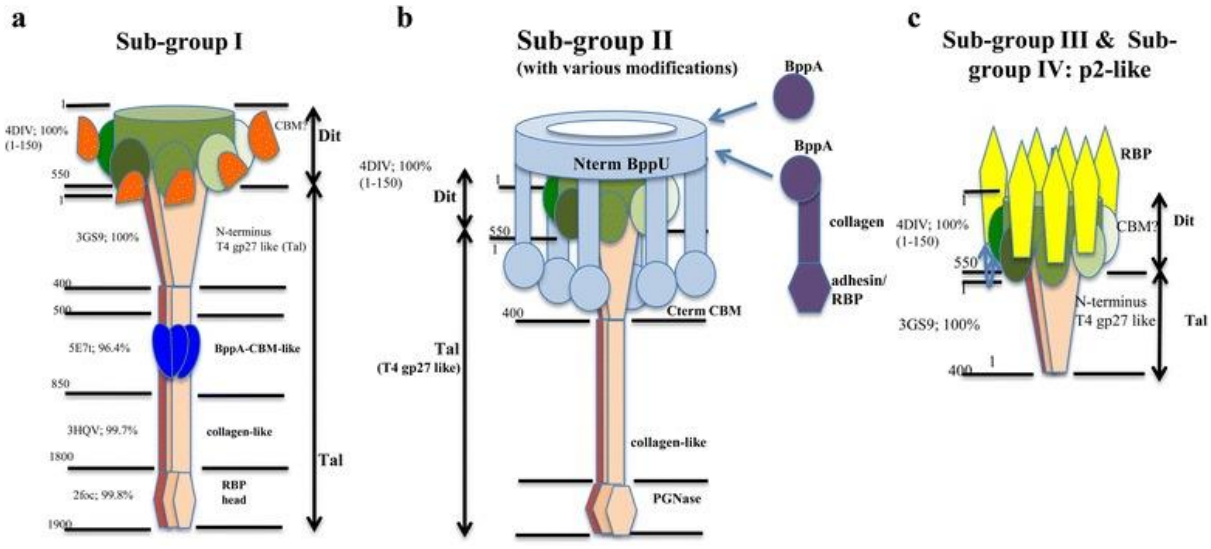


Figure 1.7: Schematic depiction of the architecture of the distal tail regions of (**a**) sub-group I, (**b**) sub-group II and (**c**) sub-group III/IV phages highlighting the domains identified by HHPred analysis (77).

1.11 Conclusions and future perspectives

It is clear that in order to fully understand phage-host interactions we first need to study and understand the structure and genetics of both, the phage and host. Phage contamination in the dairy fermented industry is a naturally occurring phenomenon, therefore it will always be a persistent problem. However, gaining insights into lactococcal phage-host interactions is a rational approach towards improving current methods used to control phage contamination and even developing new approaches to reduce the frequency of contamination. New genome sequences of lactococcal phages and bacteria are continually emerging expanding our knowledge regarding their genetic information which will gradually fully explain the interactions between phages and their bacterial hosts. Furthermore, determining the specific host receptors of Gram-positive bacteria for specific phage infection and the bacterial genes responsible for the expression of these receptors is a key element in the development of transcriptomic tools. Additionally, the identification of more proteins and carbohydrates that lactococcal phage recognise as primary and secondary receptors is necessary to isolate phage-resistant strains as the target genes can be defined as genetic markers. These will allow the development of new strategies to control phage contamination of starter cultures, hence, reducing the economic losses due to phage contamination in the dairy industry.

1.12 References

1. Twort FW. 1915. An investigation on the nature of ultra-microscopic viruses. The lancet 186:1241-1243.
2. d'Herelle F. 1917. Sur un microbe invisible antagoniste des *Bacilles Dysentériques*. Comptes rendus de l'Académie des Sciences 165:373-375.
3. Brussow H, Canchaya C, Hardt WD. 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiology and molecular biology reviews 68:560-602.
4. Emond E, Moineau S. 2007. Bacteriophages and food fermentations. Bacteriophage: genetics and molecular biology:93-124.
5. Breitbart M, Rohwer F. 2005. Here a virus, there a virus, everywhere the same virus? Trends in microbiology 13:278-284.
6. Wommack KE, Colwell RR. 2000. Virioplankton: viruses in aquatic ecosystems. Microbiology and molecular biology reviews 64:69-114.
7. Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, Lewis JA, Jacobs-Sera D, Falbo J, Gross J, Pannunzio NR, Brucker W, Kumar V, Kandasamy J, Keenan L, Bardarov S, Kriakov J, Lawrence JG, Jacobs WR, Jr., Hendrix RW, Hatfull GF. 2003. Origins of highly mosaic mycobacteriophage genomes. Cell 113:171-82.
8. Garneau JE, Moineau S. 2011. Bacteriophages of lactic acid bacteria and their impact on milk fermentations. Microbial cell factories 10:S20.
9. Marsh P, Wellington E. 1994. Phage-host interactions in soil. FEMS microbiology ecology 15:99-107.
10. Kimura M, JIA ZJ, Nakayama N, Asakawa S. 2008. Ecology of viruses in soils: past, present and future perspectives. Soil science and plant nutrition 54:1-32.
11. Suttle CA. 2007. Marine viruses—major players in the global ecosystem. Nature reviews microbiology 5:801-812.
12. Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P, Rohwer F. 2003. Metagenomic analyses of an uncultured viral community from human feces. Journal of bacteriology 185:6220-6223.
13. Mills S, Shanahan F, Stanton C, Hill C, Coffey A, Ross RP. 2013. Movers and shakers: influence of bacteriophages in shaping the mammalian gut microbiota. Gut microbes 4:4-16.

14. Kutter E. 2009. Phage host range and efficiency of plating, p 141-149, Bacteriophages. Springer.
15. Duplessis M, Moineau S. 2001. Identification of a genetic determinant responsible for host specificity in *Streptococcus thermophilus* bacteriophages. Molecular microbiology 41:325-336.
16. Lu Z, Breidt F, Plengvidhya V, Fleming H. 2003. Bacteriophage ecology in commercial sauerkraut fermentations. Applied and environmental microbiology 69:3192-3202.
17. Hamdi S, Rousseau GM, Labrie SJ, Tremblay DM, Kourda RS, Slama KB, Moineau S. 2017. Characterization of two polyvalent phages infecting *Enterobacteriaceae*. Scientific reports 7:40349.
18. Guttman B, Raya R, Kutter E. 2005. Basic phage biology. Bacteriophages: biology and applications 4.
19. Burroughs N, Marsh P, Wellington E. 2000. Mathematical analysis of growth and interaction dynamics of *Streptomycetes* and a bacteriophage in soil. Applied and environmental microbiology 66:3868-3877.
20. Hendrix RW. 2003. Bacteriophage genomics. Current opinion in microbiology 6:506-511.
21. Boyd EF, Davis BM, Hochhut B. 2001. Bacteriophage–bacteriophage interactions in the evolution of pathogenic bacteria. Trends in microbiology 9:137-144.
22. Wagner PL, Waldor MK. 2002. Bacteriophage control of bacterial virulence. Infection and immunity 70:3985-3993.
23. Abedon ST, Kuhl SJ, Blasdel BG, Kutter EM. 2011. Phage treatment of human infections. Bacteriophage 1:66-85.
24. Thacker PD. 2003. Set a microbe to kill a microbe. JAMA 290:3183-3185.
25. Smith DE, Tans SJ, Smith SB, Grimes S. 2001. The bacteriophage phi29 portal motor can package DNA against a large internal force. Nature 413:748.
26. Schuch R, Nelson D, Fischetti VA. 2002. A bacteriolytic agent that detects and kills *Bacillus anthracis*. Nature 418:884.
27. Loeffler JM, Nelson D, Fischetti VA. 2001. Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. Science 294:2170-2172.
28. Nelson D, Loomis L, Fischetti VA. 2001. Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a

bacteriophage lytic enzyme. Proceedings of the national academy of sciences of the United States of America 98:4107-4112.

29. Liu J, Dehbi M, Moeck G, Arhin F, Bauda P, Bergeron D, Callejo M, Ferretti V, Ha N, Kwan T. 2004. Antimicrobial drug discovery through bacteriophage genomics. Nature biotechnology 22:185.
30. Nyengaard N, Vogensen FK, Josephsen J. 1995. Restriction-modification systems in *Lactococcus lactis*. Gene 157:13-18.
31. Coffey A, Ross RP. 2002. Bacteriophage-resistance systems in dairy starter strains: molecular analysis to application, p 303-321, Lactic acid bacteria: genetics, metabolism and applications. Springer.
32. McGrath S, Fitzgerald GF, Sinderen Dv. 2002. Identification and characterization of phage-resistance genes in temperate lactococcal bacteriophages. Molecular microbiology 43:509-520.
33. Tükel Çl, Şanlibaba P, Özden B, Akcelik M, Akçelik M. 2006. Identification of adsorption inhibition, restriction/modification and abortive infection type phage resistance systems in *Lactococcus lactis* strains. Acta biologica hungarica 57:377-385.
34. Murray NE. 2000. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiology and molecular biology reviews 64:412-434.
35. Pingoud A, Fuxreiter M, Pingoud V, Wende W. 2005. Type II restriction endonucleases: structure and mechanism. Cellular and molecular life sciences 62:685.
36. Chopin M-C, Chopin A, Bidnenko E. 2005. Phage abortive infection in lactococci: variations on a theme. Current opinion in microbiology 8:473-479.
37. Whitehead HR, Cox GA. 1935. The occurrence of bacteriophage in cultures of lactic streptococci: A preliminary note. Government Printer.
38. Ayad E, Nashat S, El-Sadek N, Metwaly H, El-Soda M. 2004. Selection of wild lactic acid bacteria isolated from traditional Egyptian dairy products according to production and technological criteria. Food microbiology 21:715-725.
39. Brussow H, Fremont M, Bruttin A, Sidoti J, Constable A, Fryder V. 1994. Detection and classification of *Streptococcus thermophilus* bacteriophages

isolated from industrial milk fermentation. Applied and environmental microbiology 60:4537-4543.

40. Suárez V, Quiberoni A, Binetti A, Reinheimer J. 2002. Thermophilic lactic acid bacteria phages isolated from Argentinian dairy industries. Journal of food protection 65:1597-1604.
41. Moineau S, Borkaev M, Holler BJ, Walker SA, Kondo JK, Vedamuthu ER, Vandenbergh PA. 1996. Isolation and characterization of lactococcal bacteriophages from cultured buttermilk plants in the United States. Journal of dairy science 79:2104-2111.
42. Capra M, Binetti A, Mercanti D, Quiberoni A, Reinheimer J. 2009. Diversity among *Lactobacillus paracasei* phages isolated from a probiotic dairy product plant. Journal of applied microbiology 107:1350-1357.
43. Teuber M. 1993. Lactic acid bacteria. Biotechnology set second edition:325-366.
44. Mahony J, Moscarelli A, Kelleher P, Lugli GA, Ventura M, Settanni L, van Sinderen D. 2017. Phage biodiversity in artisanal cheese wheys reflects the complexity of the fermentation process. Viruses 9:45.
45. Hinrichs J. 2001. Incorporation of whey proteins in cheese. International dairy journal 11:495-503.
46. Canchaya C, Proux C, Fournous G, Bruttin A, Brüssow H. 2003. Prophage genomics. Microbiology and molecular biology reviews 67:238-276.
47. Durmaz E, Miller MJ, Azcarate-Peril MA, Toon SP, Klaenhammer TR. 2008. Genome sequence and characteristics of Lrm1, a prophage from industrial *Lactobacillus rhamnosus* strain M1. Applied and environmental microbiology 74:4601-4609.
48. Madera C, García P, Rodríguez A, Suárez JE, Martínez B. 2009. Prophage induction in *Lactococcus lactis* by the bacteriocin Lactococcin 972. International journal of food microbiology 129:99-102.
49. Lunde M, Aastveit AH, Blatny JM, Nes IF. 2005. Effects of diverse environmental conditions on ΦLC3 prophage stability in *Lactococcus lactis*. Applied and environmental microbiology 71:721-727.
50. Neve H, Berger A, Heller K. 1995. A method for detecting and enumerating airborne virulent bacteriophages of dairy starter cultures. Kieler milchwirtschaftliche forschungsberichte 47:193-207.

51. Neve H, Kemper U, Geis A, Heller K. 1994. Monitoring and characterization of lactococcal bacteriophages in a dairy plant. Kieler milchwirtschaftliche forschungsberichte (Germany).
52. Neve H, Laborius A, Heller K. 2003. Testing of the applicability of battery-powered portable microbial air samplers for detection and enumeration of airborne *Lactococcus lactis* dairy bacteriophages. Kieler milchwirtschaftliche forschungsberichte 55:301-315.
53. Verreault D, Gendron L, Rousseau GM, Veillette M, Massé D, Lindsley WG, Moineau S, Duchaine C. 2011. Detection of airborne lactococcal bacteriophages in cheese manufacturing plants. Applied and environmental microbiology 77:491-497.
54. Rousseau GM, Moineau S. 2009. Evolution of *Lactococcus lactis* phages within a cheese factory. Applied and environmental microbiology 75:5336-5344.
55. Lavelle K, Murphy J, Fitzgerald B, Lugli GA, Zomer A, Neve H, Ventura M, Franz CM, Cambillau C, van Sinderen D. 2018. A decade of *Streptococcus thermophilus* phage evolution in an Irish dairy plant. Applied and environmental microbiology:AEM. 02855-17.
56. Bron PA, Kleerebezem M. 2011. Engineering lactic acid bacteria for increased industrial functionality. Bioengineered bugs 2:80-87.
57. Deveau H, Labrie SJ, Chopin MC, Moineau S. 2006. Biodiversity and classification of lactococcal phages. Applied and environmental microbiology 72.
58. Maniloff J, Ackermann H-W. 1998. Taxonomy of bacterial viruses: establishment of tailed virus genera and the other *Caudovirales*. Archives of virology 143:2051-2063.
59. Brussow H, Hendrix RW. 2002. Phage genomics: small is beautiful. Cell 108:13-6.
60. Ackermann H-W. 1998. Tailed bacteriophages: the order *Caudovirales*. Advances in virus research 51:135-201.
61. Plisson C, White HE, Auzat I, Zafarani A, Sao-Jose C, Lhuillier S, Tavares P, Orlova EV. 2007. Structure of bacteriophage SPP1 tail reveals trigger for DNA ejection. The EMBO journal 26:3720-8.

62. Kostyuchenko VA, Leiman PG, Chipman PR, Kanamaru S, van Raaij MJ, Arisaka F, Mesyanzhinov VV, Rossmann MG. 2003. Three-dimensional structure of bacteriophage T4 baseplate. Nature structural and molecular biology 10:688-93.
63. Lander GC, Khayat R, Li R, Prevelige PE, Potter CS, Carragher B, Johnson JE. 2009. The P22 tail machine at subnanometer resolution reveals the architecture of an infection conduit. Structure 17:789-99.
64. Lander GC, Tang L, Casjens SR, Gilcrease EB, Prevelige P, Poliakov A, Potter CS, Carragher B, Johnson JE. 2006. The structure of an infectious P22 virion shows the signal for headful DNA packaging. Science 312:1791-5.
65. Braun Jr V, Hertwig S, Neve H, Geis A, Teuber M. 1989. Taxonomic differentiation of bacteriophages of *Lactococcus lactis* by electron microscopy, DNA-DNA hybridization, and protein profiles. Microbiology 135:2551-2560.
66. Ackermann H-W, Cantor ED, Jarvis AW, Lembke J, Mayo JA. 1984. New species definitions in phages of Gram-positive cocci. Intervirology 22:181-190.
67. Ackermann H-W, DuBow M, Jarvis A, Jones L, Krylov V, Maniloff J, Rocourt J, Safferman R, Schneider J, Seldin L. 1992. The species concept and its application to tailed phages. Archives of virology 124:69-82.
68. Jarvis AW, Fitzgerald GF, Mata M, Mercenier A, Neve H, Powell IB, Ronda C, Saxelin M, Teuber M. 1991. Species and type phages of lactococcal bacteriophages. Intervirology 32:2-9.
69. Labrie S, Moineau S. 2002. Complete genomic sequence of bacteriophage ul36: demonstration of phage heterogeneity within the P335 quasi-species of lactococcal phages. Virology 296:308-320.
70. Jarvis AW. 1995. Relationships by DNA homology between lactococcal phages 7–9, P335 and New Zealand industrial lactococcal phages. International dairy journal 5:355-366.
71. Moineau S, Walker SA, Vedamuthu ER, Vandenbergh PA. 1995. Cloning and sequencing of LlaDCHI [corrected] restriction/modification genes from *Lactococcus lactis* and relatedness of this system to the *Streptococcus pneumoniae* DpnII system. Applied and environmental microbiology 61:2193-2202.

72. Relano P, Mata M, Bonneau M, Ritzenthaler P. 1987. Molecular characterization and comparison of 38 virulent and temperate bacteriophages of *Streptococcus lactis*. Microbiology 133:3053-3063.
73. del Rio B, Binetti A, Martin M, Fernandez M, Magadan A, Alvarez M. 2007. Multiplex PCR for the detection and identification of dairy bacteriophages in milk. Food microbiology 24:75-81.
74. Labrie S, Moineau S. 2000. Multiplex PCR for detection and identification of lactococcal bacteriophages. Applied and environmental microbiology 66.
75. Mahanivong C, Boyce JD, Davidson BE, Hillier AJ. 2001. Sequence analysis and molecular characterization of the *Lactococcus lactis* temperate bacteriophage BK5-T. Applied and environmental microbiology 67:3564-76.
76. Trotter M, McAuliffe O, Callanan M, Edwards R, Fitzgerald GF, Coffey A, Ross RP. 2006. Genome analysis of the obligately lytic bacteriophage 4268 of *Lactococcus lactis* provides insight into its adaptable nature. Gene 366:189-199.
77. Mahony J, Oliveira J, Collins B, Hanemaaijer L, Lugli GA, Neve H, Ventura M, Kouwen TR, Cambillau C, van Sinderen D. 2017. Genetic and functional characterisation of the lactococcal P335 phage-host interactions. BMC genomics 18:146.
78. Seegers JF, Mc Grath S, O'Connell-Motherway M, Arendt EK, van de Guchte M, Creaven M, Fitzgerald GF, van Sinderen D. 2004. Molecular and transcriptional analysis of the temperate lactococcal bacteriophage Tuc2009. Virology 329:40-52.
79. Labrie SJ, Josephsen J, Neve H, Vogensen FK, Moineau S. 2008. Morphology, genome sequence, and structural proteome of type phage P335 from *Lactococcus lactis*. Applied and environmental microbiology 74:4636-4644.
80. Brøndsted L, Østergaard S, Pedersen M, Hammer K, Vogensen FK. 2001. Analysis of the complete DNA sequence of the temperate bacteriophage TP901-1: evolution, structure, and genome organization of lactococcal bacteriophages. Virology 283:93-109.
81. Blatny JM, Godager L, Lunde M, Nes IF. 2004. Complete genome sequence of the *Lactococcus lactis* temperate phage φLC3: comparative analysis of φLC3 and its relatives in lactococci and streptococci. Virology 318:231-244.

82. van Sinderen D, Karsens H, Kok J, Terpstra P, Ruiters MH, Venema G, Nauta A. 1996. Sequence analysis and molecular characterization of the temperate lactococcal bacteriophage r1t. Molecular microbiology 19:1343-1355.
83. Mahony J, Martel B, Tremblay DM, Neve H, Heller KJ, Moineau S, Sinderen D. 2013. Identification of a new P335 subgroup through molecular analysis of lactococcal phages Q33 and BM13. Applied and environmental microbiology 79(14), 4401-4409.
84. Coordinators NR. 2016. Database resources of the national center for biotechnology information. Nucleic acids research 44:D7.
85. Crutz-Le Coq AM, Cesselin B, Commissaire J, Anba J. 2002. Sequence analysis of the lactococcal bacteriophage bIL170: insights into structural proteins and HNH endonucleases in dairy phages. Microbiology 148:985-1001.
86. Chandry PS, Moore SC, Boyce JD, Davidson BE, Hillier AJ. 1997. Analysis of the DNA sequence, gene expression, origin of replication and modular structure of the *Lactococcus lactis* lytic bacteriophage sk1. Molecular microbiology 26:49-64.
87. Mahony J, Deveau H, Mc Grath S, Ventura M, Canchaya C, Moineau S, Fitzgerald GF, van Sinderen D. 2006. Sequence and comparative genomic analysis of lactococcal bacteriophages jj50, 712 and P008: evolutionary insights into the 936 phage species. FEMS Microbiology letters 261:253-61.
88. Lubbers MW, Waterfield NR, Beresford TP, Le Page RW, Jarvis AW. 1995. Sequencing and analysis of the prolate-headed lactococcal bacteriophage c2 genome and identification of the structural genes. Applied and environmental microbiology 61:4348-56.
89. Schouler C, Ehrlich SD, Chopin MC. 1994. Sequence and organization of the lactococcal prolate-headed bIL67 phage genome. Microbiology 140 (Pt 11):3061-9.
90. Samson JE, Moineau S. 2010. Characterization of *Lactococcus lactis* phage 949 and comparison with other lactococcal phages. Applied and environmental microbiology 76:6843-52.
91. Szczepanska AK, Hejnowicz MS, Kolakowski P, Bardowski J. 2007. Biodiversity of *Lactococcus lactis* bacteriophages in Polish dairy environment. Acta biochimica polonica 54:151-8.

92. Villion M, Chopin MC, Deveau H, Ehrlich SD, Moineau S, Chopin A. 2009. P087, a lactococcal phage with a morphogenesis module similar to an *Enterococcus faecalis* prophage. Virology 388:49-56.
93. Dupuis ME, Moineau S. 2010. Genome organization and characterization of the virulent lactococcal phage 1358 and its similarities to *Listeria* phages. Applied and environmental microbiology 76:1623-32.
94. Garneau JE, Tremblay DM, Moineau S. 2008. Characterization of 1706, a virulent phage from *Lactococcus lactis* with similarities to prophages from other *Firmicutes*. Virology 373:298-309.
95. Fortier LC, Bransi A, Moineau S. 2006. Genome sequence and global gene expression of Q54, a new phage species linking the 936 and c2 phage species of *Lactococcus lactis*. Journal of bacteriology 188:6101-14.
96. Kotsonis SE, Powell IB, Pillidge CJ, Limsowtin GK, Hillier AJ, Davidson BE. 2008. Characterization and genomic analysis of phage asccφ28, a phage of the family *Podoviridae* infecting *Lactococcus lactis*. Applied and environmental microbiology 74:3453-3460.
97. Chopin A, Deveau H, Ehrlich SD, Moineau S, Chopin MC. 2007. KSY1, a lactococcal phage with a T7-like transcription. Virology 365:1-9.
98. Popa MP, Mckelvey TA, Hempel J, Hendrix RW. 1991. Bacteriophage HK97 structure: wholesale covalent cross-linking between the major head shell subunits. Journal of virology 65:3227-3237.
99. Bebeacua C, Lai L, Vegge CS, Brondsted L, van Heel M, Veesler D, Cambillau C. 2013. Visualizing a complete *Siphoviridae* member by single-particle electron microscopy: the structure of lactococcal phage TP901-1. Journal of Virology 87:1061-8.
100. Collins B, Bebeacua C, Mahony J, Blangy S, Douillard FP, Veesler D, Cambillau C, van Sinderen D. 2013. Structure and functional analysis of the host recognition device of lactococcal phage Tuc2009. Journal of Virology 87:8429-40.
101. Veesler D, Johnson JE. 2012. Virus maturation. Annual review of biophysics 41:473-96.
102. Thuman-Commike PA, Greene B, Jakana J, Prasad VB, King J, Prevelige Jr PE, Chiu W. 1996. Three-dimensional structure of scaffolding-containing

phage P22 procapsids by electron cryo-microscopy. Journal of molecular biology 260:85-98.

103. Aksyuk AA, Rossmann MG. 2011. Bacteriophage assembly. Viruses 3:172-203.

104. Earnshaw WC, Casjens SR. 1980. DNA packaging by the double-stranded DNA bacteriophages. Cell 21:319-331.

105. Black LW. 1989. DNA packaging in dsDNA bacteriophages. Annual reviews in microbiology 43:267-292.

106. Catalano CE, Cue D, Feiss M. 1995. Virus DNA packaging: the strategy used by phage λ. Molecular microbiology 16:1075-1086.

107. Baumann RG, Mullaney J, Black LW. 2006. Portal fusion protein constraints on function in DNA packaging of bacteriophage T4. Molecular microbiology 61:16-32.

108. Casjens S, Wyckoff E, Hayden M, Sampson L, Eppler K, Randall S, Moreno ET, Serwer P. 1992. Bacteriophage P22 portal protein is part of the gauge that regulates packing density of intravirion DNA. Journal of molecular biology 224:1055-1074.

109. Sun S, Kondabagil K, Draper B, Alam TI, Bowman VD, Zhang Z, Hegde S, Fokine A, Rossmann MG, Rao VB. 2008. The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. Cell 135:1251-1262.

110. Isidro A, Henriques AO, Tavares P. 2004. The portal protein plays essential roles at different steps of the SPP1 DNA packaging process. Virology 322:253-263.

111. Jardine PJ, Anderson DL. 2006. DNA packaging in double-stranded DNA phages. The bacteriophages 2:49-65.

112. Fujisawa H, Morita M. 1997. Phage DNA packaging. Genes to cells 2:537-545.

113. Feiss M, Fisher R, Crayton M, Egner C. 1977. Packaging of the bacteriophage λ chromosome: Effect of chromosome length. Virology 77:281-293.

114. Hashimoto C, Fujisawai H. 1988. Packaging and transduction of non-T3 DNA by bacteriophage T3. Virology 166:432-439.

115. Arendt EK, Daly C, Fitzgerald GF, Van de Guchte M. 1994. Molecular characterization of lactococcal bacteriophage Tuc2009 and identification and

analysis of genes encoding lysin, a putative holin, and two structural proteins. Applied and environmental microbiology 60:1875-1883.

116. Hsiao C, Black L. 1977. DNA packaging and the pathway of bacteriophage T4 head assembly. Proceedings of the national academy of sciences of the United States of America 74:3652-3656.

117. Donate LE, Herranz L, Secilla JP, Carazo J, Fujisawa H, Carrascosa J. 1988. Bacteriophage T3 connector: three-dimensional structure and comparison with other viral head-tail connecting regions. Journal of molecular biology 201:91-100.

118. Cue D, Feiss M. 1997. Genetic evidence that recognition of cosQ the signal for termination of phage λ DNA packaging, depends on the extent of head filling. Genetics 147:7-17.

119. Tsui L, Hendrix RW. 1980. Head-tail connector of bacteriophage lambda. Journal of molecular biology 142:419-438.

120. Maxwell KL, Yee AA, Arrowsmith CH, Gold M, Davidson AR. 2002. The solution structure of the bacteriophage λ head–tail joining protein, gpFII. Journal of molecular biology 318:1395-1404.

121. Maxwell KL, Yee AA, Booth V, Arrowsmith CH, Gold M, Davidson AR. 2001. The solution structure of bacteriophage λ protein W, a small morphogenetic protein possessing a novel fold1. Journal of molecular biology 308:9-14.

122. Vegge CS, Brøndsted L, Neve H, Mc Grath S, van Sinderen D, Vogensen FK. 2005. Structural characterization and assembly of the distal tail structure of the temperate lactococcal bacteriophage TP901-1. Journal of bacteriology 187:4187-4197.

123. Veesler D, Spinelli S, Mahony J, Lichiere J, Blangy S, Bricogne G, Legrand P, Ortiz-Lombardia M, Campanacci V, van Sinderen D, Cambillau C. 2012. Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. Proceedings of the national academy of sciences of the United States of America 109:8954-8.

124. Lhuillier S, Gallopin M, Gilquin B, Brasiles S, Lancelot N, Letellier G, Gilles M, Dethan G, Orlova EV, Couprie J, Tavares P, Zinn-Justin S. 2009. Structure of bacteriophage SPP1 head-to-tail connection reveals mechanism for viral

DNA gating. Proceedings of the national academy of sciences of the United States of America 106:8507-12.

125. Lebedev AA, Krause MH, Isidro AL, Vagin AA, Orlova EV, Turner J, Dodson EJ, Tavares P, Antson AA. 2007. Structural framework for DNA translocation via the viral portal protein. The EMBO journal 26:1984-94.

126. Goulet A, Lai-Kee-Him J, Veesler D, Auzat I, Robin G, Shepherd DA, Ashcroft AE, Richard E, Lichiere J, Tavares P, Cambillau C, Bron P. 2011. The opening of the SPP1 bacteriophage tail, a prevalent mechanism in Gram-positive-infecting siphophages. Journal of biological chemistry 286:25397-405.

127. Veesler D, Robin G, Lichiere J, Auzat I, Tavares P, Bron P, Campanacci V, Cambillau C. 2010. Crystal structure of bacteriophage SPP1 distal tail protein (gp19.1): a baseplate hub paradigm in Gram-positive infecting phages. Journal of biological chemistry 285:36666-73.

128. Veesler D, Cambillau C. 2011. A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. Microbiology and molecular biology reviews 75:423-433.

129. Pell LG, Liu A, Edmonds L, Donaldson LW, Howell PL, Davidson AR. 2009. The X-ray crystal structure of the phage λ tail terminator protein reveals the biologically relevant hexameric ring structure and demonstrates a conserved mechanism of tail termination among diverse long-tailed phages. Journal of molecular biology 389:938-951.

130. Sciara G, Bebeacua C, Bron P, Tremblay D, Ortiz-Lombardia M, Lichière J, Van Heel M, Campanacci V, Moineau S, Cambillau C. 2010. Structure of lactococcal phage p2 baseplate and its mechanism of activation. Proceedings of the national academy of sciences of the United States of America 107:6852-6857.

131. Stockdale SR, Collins B, Spinelli S, Douillard FP, Mahony J, Cambillau C, van Sinderen D. 2015. Structure and assembly of TP901-1 virion unveiled by mutagenesis. PLOS One 10:e0131676.

132. Spinelli S, Campanacci V, Blangy S, Moineau S, Tegoni M, Cambillau C. 2006. Modular structure of the receptor binding proteins of *Lactococcus lactis* phages. The RBP structure of the temperate phage TP901-1. Journal of biological chemistry 281:14256-62.

133. Baptista C, Santos MA, São-José C. 2008. Phage SPP1 reversible adsorption to *Bacillus subtilis* cell wall teichoic acids accelerates virus recognition of membrane receptor YueB. Journal of bacteriology 190:4989-4996.
134. Valyasevi R, Sandine W, Geller B. 1991. A membrane protein is required for bacteriophage c2 infection of *Lactococcus lactis* subsp. *lactis* C2. Journal of bacteriology 173:6095-6100.
135. Mahony J, Van Sinderen D. 2012. Structural aspects of the interaction of dairy phages with their host bacteria. Viruses 4:1410-1424.
136. Millen AM, Romero DA. 2016. Genetic determinants of lactococcal C2 viruses for host infection and their role in phage evolution. Journal of general virology 97:1998-2007.
137. Bebeacua C, Bron P, Lai L, Vegge CS, Brondsted L, Spinelli S, Campanacci V, Veesler D, van Heel M, Cambillau C. 2010. Structure and molecular assignment of lactococcal phage TP901-1 baseplate. Journal of biological chemistry 285:39079-86.
138. Stockdale SR, Mahony J, Courtin P, Chapot-Chartier MP, van Pijkeren JP, Britton RA, Neve H, Heller KJ, Aideh B, Vogensen FK, van Sinderen D. 2013. The lactococcal phages Tuc2009 and TP901-1 incorporate two alternate forms of their tail fiber into their virions for infection specialization. Journal of biological chemistry 288:5581-90.
139. Mc Grath S, Neve H, Seegers JF, Eijlander R, Vegge CS, Brondsted L, Heller KJ, Fitzgerald GF, Vogensen FK, van Sinderen D. 2006. Anatomy of a lactococcal phage tail. Journal of bacteriology 188:3972-82.
140. Monteville MR, Ardestani B, Geller BL. 1994. Lactococcal bacteriophages require a host cell wall carbohydrate and a plasma membrane protein for adsorption and ejection of DNA. Applied and environmental microbiology 60:3204-3211.
141. Rakhuba D, Kolomiets E, Dey ES, Novik G. 2010. Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. Polish journal of microbiology 59:145-155.
142. Farenc C, Spinelli S, Vinogradov E, Tremblay D, Blangy S, Sadovskaya I, Moineau S, Cambillau C. 2014. Molecular insights on the recognition of a *Lactococcus lactis* cell wall pellicle by the phage 1358 receptor binding protein. Journal of Virology 88:7005-15.

143. Mahony J, Kot W, Murphy J, Ainsworth S, Neve H, Hansen LH, Heller KJ, Sørensen SJ, Hammer K, Cambillau C. 2013. Investigation of the relationship between *Lactococcal* host cell wall polysaccharide genotype and 936 phage receptor binding protein phylogeny. Applied and environmental microbiology 79:4385-4392.
144. Mahony J, Randazzo W, Neve H, Settanni L, van Sinderen D. 2015. Lactococcal 949 group phages recognize a carbohydrate receptor on the host cell surface. Applied and environmental microbiology 81:3299-305.
145. Sijtsma L, Wouters J, Hellingwerf K. 1990. Isolation and characterization of lipoteichoic acid, a cell envelope component involved in preventing phage adsorption, from *Lactococcus lactis* subsp. *cremoris* SK110. Journal of bacteriology 172:7126-7130.
146. Xia G, Corrigan RM, Winstel V, Goerke C, Gründling A, Peschel A. 2011. Wall teichoic acid-dependent adsorption of *Staphylococcal siphovirus* and *myovirus*. Journal of bacteriology 193:4006-4009.
147. Luria S, Steiner DL. 1954. The role of calcium in the penetration of bacteriophage T5 into its host. Journal of bacteriology 67:635.
148. Mahony J, Tremblay DM, Labrie SJ, Moineau S, van Sinderen D. 2015. Investigating the requirement for calcium during lactococcal phage infection. International journal of food microbiology 201:47-51.
149. Luckey M, Nikaido H. 1980. Diffusion of solutes through channels produced by phage lambda receptor protein of *Escherichia coli*: Inhibition by higher oligosaccharides of maltose series. Biochemical and biophysical research communications 93:166-171.
150. Samsonov VV, Samsonov VV, Sineoky SP. 2002. DcrA and dcrB *Escherichia coli* genes can control DNA injection by phages specific for BtuB and FhuA receptors. Research in microbiology 153:639-646.
151. Cumby N, Reimer K, Mengin-Lecreulx D, Davidson AR, Maxwell KL. 2015. The phage tail tape measure protein, an inner membrane protein and a periplasmic chaperone play connected roles in the genome injection process of *E. coli* phage HK97. Molecular microbiology 96:437-447.
152. Bohannan BJ, Lenski RE. 2000. Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. Ecology letters 3:362-377.

153. Lenski RE. 1984. Coevolution of bacteria and phage: are there endless cycles of bacterial defenses and phage counterdefenses? Journal of theoretical biology 108:319-325.
154. Kanamaru S, Leiman PG, Kostyuchenko VA, Chipman PR, Mesyanzhinov VV, Arisaka F, Rossmann MG. 2002. Structure of the cell-puncturing device of bacteriophage T4. Nature 415:553.
155. Letellier L, Boulanger P. 1989. Involvement of ion channels in the transport of phage DNA through the cytoplasmic membrane of *E. coli*. Biochimie 71:167-174.
156. Boulanger P, Letellier L. 1988. Characterization of ion channels involved in the penetration of phage T4 DNA into *Escherichia coli* cells. Journal of biological chemistry 263:9767-9775.
157. Garcia LR, Molineux IJ. 1996. Transcription-independent DNA translocation of bacteriophage T7 DNA into *Escherichia coli*. Journal of bacteriology 178:6921-6929.
158. Molineux IJ. 2001. No syringes please, ejection of phage T7 DNA from the virion is enzyme driven. Molecular microbiology 40:1-8.
159. Lanni YT. 1968. First-step-transfer deoxyribonucleic acid of bacteriophage T5. Bacteriological reviews 32:227.
160. Letellier L, Plançon L, Bonhivers M, Boulanger P. 1999. Phage DNA transport across membranes. Research in microbiology 150:499-505.
161. Letellier L, Boulanger P, de Frutos M, Jacquot P. 2003. Channeling phage DNA through membranes: from in vivo to in vitro. Research in microbiology 154:283-287.
162. González-Huici V, Salas M, Hermoso JM. 2004. The push–pull mechanism of bacteriophage Ø29 DNA injection. Molecular microbiology 52:529-540.
163. Bertani G. Lysogenic versus lytic cycle of phage multiplication, p 65-70. *In* (ed), Cold Spring Harbor Laboratory Press,
164. Wang N, Deaton J, Young R. 2003. Sizing the holin lesion with an endolysin-β-galactosidase fusion. Journal of bacteriology 185:779-787.
165. White R, Chiba S, Pang T, Dewey JS, Savva CG, Holzenburg A, Pogliano K, Young R. 2011. Holin triggering in real time. Proceedings of the national academy of sciences of the United States of America 108:798-803.

166. García P, García J, García E, Sánchez-Puelles J, López R. 1990. Modular organization of the lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. Gene 86:81-88.
167. Mills S, Griffin C, O'Sullivan O, Coffey A, McAuliffe O, Meijer W, Serrano L, Ross R. 2011. A new phage on the 'Mozzarella' block: bacteriophage 5093 shares a low level of homology with other *Streptococcus thermophilus* phages. International dairy journal 21:963-969.
168. Ainsworth S, Zomer A, Mahony J, van Sinderen D. 2013. Lytic infection of *Lactococcus lactis* by bacteriophages Tuc2009 and c2 trigger alternative transcriptional host responses. Applied and environmental microbiology:AEM. 01197-13.
169. Fallico V, Ross RP, Fitzgerald GF, McAuliffe O. 2011. Genetic response to bacteriophage infection in *Lactococcus lactis* reveals a four-strand approach involving induction of membrane stress proteins, D-alanylation of the cell wall, maintenance of proton motive force and energy conservation. Journal of virology:JVI. 00275-11.
170. Lwoff A. 1953. Lysogeny. Bacteriological reviews 17:269.
171. Boye E, Løbner-Olesen A. 1990. The role of dam methyltransferase in the control of DNA replication in *E. coli*. Cell 62:981-989.
172. Julio SM, Heithoff DM, Provenzano D, Klose KE, Sinsheimer RL, Low DA, Mahan MJ. 2001. DNA adenine methylase is essential for viability and plays a role in the pathogenesis of *Yersinia pseudotuberculosis* and *Vibrio cholerae*. Infection and immunity 69:7610-7615.
173. Mohapatra SS, Fioravanti A, Biondi EG. 2014. DNA methylation in *Caulobacter* and other *Alphaproteobacteria* during cell cycle progression. Trends in microbiology 22:528-535.
174. Marinus MG, Casadesus J. 2009. Roles of DNA adenine methylation in host–pathogen interactions: mismatch repair, transcriptional regulation, and more. FEMS microbiology reviews 33:488-503.
175. Groth AC, Calos MP. 2004. Phage integrases: biology and applications. Journal of molecular biology 335:667-678.
176. Lillehaug D, Birkeland N. 1993. Characterization of genetic elements required for site-specific integration of the temperate lactococcal bacteriophage phi

LC3 and construction of integration-negative phi LC3 mutants. Journal of bacteriology 175:1745-1755.

177. Van de Guchte M, Daly C, Fitzgerald GF, Arendt EK. 1994. Identification of int and attP on the genome of lactococcal bacteriophage Tuc2009 and their use for site-specific plasmid integration in the chromosome of Tuc2009-resistant *Lactococcus lactis* MG1363. Applied and environmental microbiology 60:2324-2329.

178. Johnson AD, Poteete AR, Lauer G, Sauer RT, Ackers GK, Ptashne M. 1981. Lambda repressor and cro-components of an efficient molecular switch. Nature 294:217-23.

179. Sauer R, Ross M, Ptashne M. 1982. Cleavage of the lambda and P22 repressors by recA protein. Journal of biological chemistry 257:4458-4462.

180. Durmaz E, Madsen SM, Israelsen H, Klaenhammer TR. 2002. *Lactococcus lactis* lytic bacteriophages of the P335 group are inhibited by overexpression of a truncated CI repressor. Journal of bacteriology 184:6532-6544.

181. García P, Ladero V, Alonso JC, Suárez JE. 1999. Cooperative interaction of CI protein regulates lysogeny of *Lactobacillus casei* by bacteriophage A2. Journal of virology 73:3920-3929.

182. Dhaese P, Seurinck J, De Smet B, Van Montagu M. 1985. Nucleotide sequence and mutational analysis of an immunity repressor gene from *Bacillus subtilis* temperate phage ø105. Nucleic acids research 13:5441-5455.

183. Ladero V, García P, Bascarán V, Herrero M, Alvarez MA, Suárez JE. 1998. Identification of the repressor-encoding gene of the *Lactobacillus* bacteriophage A2. Journal of bacteriology 180:3474-3476.

184. Newton GJ, Daniels C, Burrows LL, Kropinski AM, Clarke AJ, Lam JS. 2001. Three-component-mediated serotype conversion in *Pseudomonas aeruginosa* by bacteriophage D3. Molecular microbiology 39:1237-1247.

185. Vostrov AA, Vostrukhina OA, Svarchevsky AN, Rybchin VN. 1996. Proteins responsible for lysogenic conversion caused by coliphages N15 and phi80 are highly homologous. Journal of bacteriology 178:1484-1486.

186. Uc-Mass A, Loeza EJ, de la Garza M, Guarneros G, Hernández-Sánchez J, Kameyama L. 2004. An orthologue of the cor gene is involved in the exclusion of temperate lambdoid phages. Evidence that Cor inactivates FhuA receptor functions. Virology 329:425-433.

187. Bondy-Denomy J, Davidson AR. 2014. When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness. Journal of microbiology 52:235-242.
188. Cumby N, Edwards AM, Davidson AR, Maxwell KL. 2012. The bacteriophage HK97 gp15 moron element encodes a novel superinfection exclusion protein. Journal of bacteriology 194:5012-5019.
189. Hofer B, Ruge M, Dreiseikelmann B. 1995. The superinfection exclusion gene (*sieA*) of bacteriophage P22: identification and overexpression of the gene and localization of the gene product. Journal of bacteriology 177:3080-3086.
190. Maillou J, Dreiseikelmann B. 1990. The *sim* gene of *Escherichia coli* phage P1: nucleotide sequence and purification of the processed protein. Virology 175:500-507.
191. Mahony J, McGrath S, Fitzgerald GF, van Sinderen D. 2008. Identification and characterization of lactococcal-prophage-carried superinfection exclusion genes. Applied and environmental microbiology 74:6206-6215.
192. Sun X, Göhler A, Heller KJ, Neve H. 2006. The *ltp* gene of temperate Streptococcus thermophilus phage TP-J34 confers superinfection exclusion to *Streptococcus thermophilus* and *Lactococcus lactis*. Virology 350:146-157.
193. Bruttin A, Desiere F, Lucchini S, Foley S, Brüssow H. 1997. Characterization of the lysogeny DNA module from the temperate *Streptococcus thermophilus* bacteriophage φSfi21. Virology 233:136-148.
194. Casjens S. 2003. Prophages and bacterial genomics: what have we learned so far? Molecular microbiology 49:277-300.
195. Kelleher P. 2017. Comparative and functional genomic analysis of dairy lactococci: PhD thesis. University College Cork.
196. Silhavy TJ, Kahne D, Walker S. 2010. The bacterial cell envelope. Cold spring harbor perspectives in biology 2:a000414.
197. Mitchell P. 1961. Approaches to the analysis of specific membrane transport. Biological structure and function 2:581-599.
198. Heppel LA. 1967. Selective release of enzymes from bacteria. Science 156:1451-1455.
199. Miura T, Mizushima S. 1968. Separation by density gradient centrifugation of two types of membranes from spheroplast membrane of *Escherichia coli* K12. Biochimica et biophysica acta (BBA)-biomembranes 150:159-161.

200. Pollack JH, Neuhaus FC. 1994. Changes in wall teichoic acid during the rod-sphere transition of *Bacillus subtilis* 168. Journal of bacteriology 176:7252-7259.

201. Boylan R, Mendelson N, Brooks D, Young F. 1972. Regulation of the bacterial cell wall: analysis of a mutant of *Bacillus subtilis* defective in biosynthesis of teichoic acid. Journal of bacteriology 110:281-290.

202. Strominger JL, Park JT, Thompson RE. 1959. Composition of the cell wall of *Staphylococcus aureus*: its relation to the mechanism of action of penicillin. Journal of biological chemistry 234:3263-3268.

203. Schneewind O, Fowler A, Faull KF. 1995. Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. Science 268:103-106.

204. Delcour J, Ferain T, Deghorain M, Palumbo E, Hols P. 1999. The biosynthesis and functionality of the cell-wall of lactic acid bacteria, p159-184, Lactic acid bacteria: genetics, metabolism and applications. Springer.

205. Chapot-Chartier MP, Vinogradov E, Sadovskaya I, Andre G, Mistou MY, Trieu-Cuot P, Furlan S, Bidnenko E, Courtin P, Pechoux C, Hols P, Dufrene YF, Kulakauskas S. 2010. Cell surface of *Lactococcus lactis* is covered by a protective polysaccharide pellicle. Journal of biological chemistry 285:10464-71.

206. Glauert AM, Thornley MJ. 1969. The topography of the bacterial cell wall. Annual reviews in microbiology 23:159-198.

207. Osborn M, Gander J, Parisi E, Carson J. 1972. Mechanism of assembly of the outer membrane of *Salmonella typhimurium* isolation and characterization of cytoplasmic and outer membrane. Journal of biological chemistry 247:3962-3972.

208. Randall-Hazelbauer L, Schwartz M. 1973. Isolation of the bacteriophage lambda receptor from *Escherichia coli*. Journal of bacteriology 116:1436-1446.

209. Hancock R, Braun V. 1976. Nature of the energy requirement for the irreversible adsorption of bacteriophages T1 and phi80 to *Escherichia coli*. Journal of bacteriology 125:409-415.

210. Roof S, Allard J, Bertrand K, Postle K. 1991. Analysis of *Escherichia coli* TonB membrane topology by use of PhoA fusions. Journal of bacteriology 173:5554-5557.

211. Brogden KA. 2005. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? Nature reviews microbiology 3:238.
212. Kamio Y, Nikaido H. 1976. Outer membrane of *Salmonella typhimurium*: accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium. Biochemistry 15:2561-2570.
213. Sankaran K, Wu HC. 1994. Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. Journal of biological chemistry 269:19701-19706.
214. Miyadai H, Tanaka-Masuda K, Matsuyama S-i, Tokuda H. 2004. Effects of lipoprotein overproduction on the induction of DegP (HtrA) involved in quality control in the *Escherichia coli* periplasm. Journal of biological chemistry 279:39807-39813.
215. Cowan S, Schirmer T, Rummel G, Steiert M, Ghosh R, Pauptit R, Jansonius J, Rosenbusch J. 1992. Crystal structures explain functional properties of two *E. coli* porins. Nature 358:727.
216. Schirmer T, Keller TA, Wang Y-F, Rosenbusch JP. 1995. Structural basis for sugar translocation through maltoporin channels at 3.1 A resolution. Science 267:512-514.
217. Arora A, Rinehart D, Szabo G, Tamm LK. 2000. Refolded outer membrane protein A of *Escherichia coli* forms ion channels with two conductance states in planar lipid bilayers. Journal of biological chemistry 275:1594-1600.
218. Nikaido H. 2003. Molecular basis of bacterial outer membrane permeability revisited. Microbiology and molecular biology reviews 67:593-656.
219. Snijder H, Ubarretxena-Belandia I, Blaauw M, Kalk K, Verheij H, Egmond M, Dekker N, Dijkstra B. 1999. Structural evidence for dimerization-regulated activation of an integral membrane phospholipase. Nature 401:717.
220. Vandeputte-Rutten L, Kramer RA, Kroon J, Dekker N, Egmond MR, Gros P. 2001. Crystal structure of the outer membrane protease OmpT from *Escherichia coli* suggests a novel catalytic site. The EMBO journal 20:5033-5039.
221. Hwang PM, Choy W-Y, Lo EI, Chen L, Forman-Kay JD, Raetz CR, Privé GG, Bishop RE, Kay LE. 2002. Solution structure and dynamics of the outer

membrane enzyme PagP by NMR. Proceedings of the national academy of sciences of the United States of America 99:13560-13565.

222. Raetz CR, Whitfield C. 2002. Lipopolysaccharide endotoxins. Annual review of biochemistry 71:635-700.

223. Vollmer W. 2008. Structural variation in the glycan strands of bacterial peptidoglycan. FEMS microbiology reviews 32:287-306.

224. Gan L, Chen S, Jensen GJ. 2008. Molecular organization of Gram-negative peptidoglycan. Proceedings of the national academy of sciences of the United States of America 105:18953-18957.

225. Braun V. 1975. Covalent lipoprotein from the outer membrane of *Escherichia coli*. Biochimica et Biophysica Acta (BBA)-reviews on biomembranes 415:335-377.

226. Yem DW, Wu HC. 1978. Physiological characterization of an *Escherichia coli* mutant altered in the structure of murein lipoprotein. Journal of bacteriology 133:1419-1426.

227. Mullineaux CW, Nenninger A, Ray N, Robinson C. 2006. Diffusion of green fluorescent protein in three cell environments in *Escherichia coli*. Journal of bacteriology 188:3442-3448.

228. De Duve C, Wattiaux R. 1966. Functions of lysosomes. Annual review of physiology 28:435-492.

229. Jansson M. 1988. Phosphate uptake and utilization by bacteria and algae, p 177-189, Phosphorus in freshwater ecosystems. Springer.

230. Willsky GR, Malamy MH. 1976. Control of the synthesis of alkaline phosphatase and the phosphate-binding protein in *Escherichia coli*. Journal of bacteriology 127:595-609.

231. Jean-Pierre B, Jean-Marie L. 2007. Osmoregulation in the periplasm, p 325-341, The periplasm. American Society of Microbiology.

232. Miller SI, Salama NR. 2018. The Gram-negative bacterial periplasm: Size matters. PLOS Biology 16:e2004935.

233. Chapman GB, Hillier J. 1953. Electron microscopy of ultra-thin sections of bacteria I.: cellular division in *Bacillus cereus*. Journal of bacteriology 66:362.

234. Mahony J, Ainsworth S, Stockdale S, van Sinderen D. 2012. Phages of lactic acid bacteria: the role of genetics in understanding phage-host interactions and their co-evolutionary processes. Virology 434:143-150.

235. Ventura M, Callegari M, Morelli L. 1999. Surface layer variations affecting phage adsorption on seven *Lactobacillus helveticus* strains. Annali di microbiologia ed enzimologia 49:45-54.
236. Dijkstra AJ, Keck W. 1996. Peptidoglycan as a barrier to transenvelope transport. Journal of bacteriology 178:5555.
237. Demchick P, Koch AL. 1996. The permeability of the wall fabric of *Escherichia coli* and *Bacillus subtilis*. Journal of bacteriology 178:768-773.
238. Araki Y, oba S, ARAKI S, ITO E. 1980. Enzymatic deacetylation of N-acetylglucosamine residues in cell wall peptidoglycan. The journal of biochemistry 88:469-479.
239. Zipperle Jr GF, Ezzell Jr JW, Doyle RJ. 1984. Glucosamine substitution and muramidase susceptibility in *Bacillus anthracis*. Canadian journal of microbiology 30:553-559.
240. Logardt I, Neujahr HY. 1975. Lysis of modified walls from *Lactobacillus fermentum*. Journal of bacteriology 124:73-77.
241. Amano Ki, Hayashi H, Araki Y, Ito E. 1977. The action of lysozyme on peptidoglycan with N-unsubstituted glucosamine residues. The FEBS journal 76:299-307.
242. Kawagishi S, Araki Y, Ito E. 1980. *Bacillus cereus* autolytic endoglucosaminidase active on cell wall peptidoglycan with N-unsubstituted glucosamine residues. Journal of bacteriology 141:137-143.
243. Kleppe G, Vasstrand E, Jensen Hb. 1981. The specificity requirements of bacteriophage T4 lysozyme. The FEBS journal 119:589-593.
244. Chang JD, Wallace AG, Foster EE, Kim SJ. 2018. Peptidoglycan compositional analysis of *Enterococcus faecalis* biofilm by stable isotope labeling by amino acids in a bacterial culture. Biochemistry 57:1274-1283.
245. Billot-Klein D, Legrand R, Schoot B, van Heijenoort J, Gutmann L. 1997. Peptidoglycan structure of *Lactobacillus casei*, a species highly resistant to glycopeptide antibiotics. Journal of bacteriology 179:6208-6212.
246. Coyette J, Ghuysen JM. 1970. Structure of the walls of *Lactobacillus acidophilus* strain 63 AM. Biochemistry 9:2935-2943.
247. Clarke AJ, Dupont C. 1992. O-acetylated peptidoglycan: its occurrence, pathobiological significance, and biosynthesis. Canadian journal of microbiology 38:85-91.

248. Laaberki M-H, Pfeffer J, Clarke AJ, Dworkin J. 2011. O-Acetylation of peptidoglycan is required for proper cell separation and S-layer anchoring in *Bacillus anthracis*. Journal of biological chemistry 286:5278-5288.
249. Armstrong J, Baddiley J, Buchanan J, Carss B, Greenberg G. 1958. Isolation and structure of ribitol phosphate derivatives (teichoic acids) from bacterial cell walls. Journal of the chemical society (resumed):4344-4354.
250. Brown S, Santa Maria Jr JP, Walker S. 2013. Wall teichoic acids of Gram-positive bacteria. Annual review of microbiology 67:313-336.
251. Fischer W. 1990. Bacterial phosphoglycolipids and lipoteichoic acids, p 123-234, Glycolipids, phosphoglycolipids, and sulfoglycolipids. Springer.
252. Fischer W. 1994. Lipoteichoic acids and lipoglycans, p 199-215, New comprehensive biochemistry, vol 27. Elsevier.
253. Archibald A, Hancock I, Harwood C. 1993. Cell wall structure, synthesis, and turnover, p 381-410, *Bacillus subtilis* and other Gram-positive bacteria. American society of microbiology.
254. Kojima N, Araki Y, Ito E. 1985. Structural studies on the linkage unit of ribitol teichoic acid of *Lactobacillus plantarum*. The FEBS journal 148:29-34.
255. Vinogradov E, Sadovskaya I, Courtin P, Kulakauskas S, Grard T, Mahony J, van Sinderen D, Chapot-Chartier M-P. 2018. Determination of the cell wall polysaccharide and teichoic acid structures from *Lactococcus lactis* IL1403. Carbohydrate research 462:39-44.
256. Neuhaus FC. 1985. Inter-chain transacylation of D-alanine ester residues of lipoteichoic acid: a unique mechanism of membrane communication. Portland press limited.
257. Valyasevi R, Sandine W, Geller B. 1990. The bacteriophage kh receptor of *Lactococcus lactis* subsp. *cremoris* KH is the rhamnose of the extracellular wall polysaccharide. Applied and environmental microbiology 56:1882-1889.
258. Ward J. 1981. Teichoic and teichuronic acids: biosynthesis, assembly, and location. Microbiological reviews 45:211.
259. Lang W, Glassey K, Archibald A. 1982. Influence of phosphate supply on teichoic acid and teichuronic acid content of *Bacillus subtilis* cell walls. Journal of bacteriology 151:367-375.

260. Neuhaus FC, Baddiley J. 2003. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in Gram-positive bacteria. Microbiology and molecular biology reviews 67:686-723.
261. Wicken A, Knox K. 1970. Studies on the group F antigen of lactobacilli: isolation of a teichoic acid-lipid complex from *Lactobacillus fermenti* NCTC 6991. Microbiology 60:293-301.
262. Young F. 1967. Requirement of glucosylated teichoic acid for adsorption of phage in *Bacillus subtilis* 168. Proceedings of the national academy of sciences of the United States of America 58:2377-2384.
263. Sijtsma L, Sterkenburg A, Wouters JT. 1988. Properties of the cell walls of *Lactococcus lactis* subsp. *cremoris* SK110 and SK112 and their relation to bacteriophage resistance. Applied and environmental microbiology 54:2808-11.
264. Sijtsma L, Jansen N, Hazeleger WC, Wouters JT, Hellingwerf KJ. 1990. Cell surface characteristics of bacteriophage-resistant *Lactococcus lactis* subsp. *cremoris* SK110 and its bacteriophage-sensitive variant SK112. Applied and environmental microbiology 56:3230-3.
265. Grant W. 1979. Cell wall teichoic acid as a reserve phosphate source in *Bacillus subtilis*. Journal of bacteriology 137:35-43.
266. Hughes A, Hancock I, Baddiley J. 1973. The function of teichoic acids in cation control in bacterial membranes. Biochemical journal 132:83.
267. Höltje J, Tomasz A. 1975. Lipoteichoic acid: a specific inhibitor of autolysin activity in *Pneumococcus*. Proceedings of the national academy of sciences of the United States of America 72:1690-1694.
268. Yamada M, Hirose A, Matsuhashi M. 1975. Association of lack of cell wall teichuronic acid with formation of cell packets of *Micrococcus lysodeikticus* (*luteus*) mutants. Journal of bacteriology 123:678-686.
269. Bierbaum G, Sahl H. 1987. Autolytic system of *Staphylococcus simulans* 22: influence of cationic peptides on activity of N-acetylmuramoyl-L-alanine amidase. Journal of bacteriology 169:5452-5458.
270. Gopal PK, Crow VL. 1993. Characterization of loosely associated material from the cell surface of *Lactococcus lactis* subsp. *cremoris* E8 and its phage-resistant variant strain 398. Applied and environmental microbiology 59:3177-3182.

271. Whitfield C. 1988. Bacterial extracellular polysaccharides. Canadian journal of microbiology 34:415-420.
272. Wicken A, Ayres A, Campbell L, Knox K. 1983. Effect of growth conditions on production of rhamnose-containing cell wall and capsular polysaccharides by strains of *Lactobacillus casei* subsp. *rhamnosus*. Journal of bacteriology 153:84-92.
273. Hall EA, Knox K. 1965. Properties of the polysaccharide and mucopeptide components of the cell wall of *Lactobacillus casei*. Biochemical journal 96:310.
274. Wolin M, Douglas LJ. 1971. Cell wall polymers and phage lysis of *Lactobacillus plantarum*. Biochemistry 10:1551-1555.
275. Schleifer K, Kilpper-Bälz R. 1987. Molecular and chemotaxonomic approaches to the classification of streptococci, enterococci and lactococci: a review. Systematic and applied microbiology 10:1-19.
276. Ishibashi K, Takesue S, Watanabe K, Oishi K. 1982. Use of lectins to characterize the receptor sites for bacteriophage PL-1 of *Lactobacillus casei*. Microbiology 128:2251-2259.
277. Ainsworth S, Sadovskaya I, Vinogradov E, Courtin P, Guerardel Y, Mahony J, Grard T, Cambillau C, Chapot-Chartier MP, van Sinderen D. 2014. Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. MBio 5:e00880-14.
278. Legrand P, Collins B, Blangy S, Murphy J, Spinelli S, Gutierrez C, Richet N, Kellenberger C, Desmyter A, Mahony J, van Sinderen D, Cambillau C. 2016. The atomic structure of the phage Tuc2009 baseplate tripod suggests that host recognition involves two different carbohydrate binding modules. MBio 7:e01781-15.
279. Dramsi S, Magnet S, Davison S, Arthur M. 2008. Covalent attachment of proteins to peptidoglycan. FEMS microbiology reviews 32:307-320.
280. Scott JR, Barnett TC. 2006. Surface proteins of Gram-positive bacteria and how they get there. Annual Review of microbiology 60:397-423.
281. Sleytr UB, Messner P, Pum D, Sára M. 1993. Crystalline bacterial cell surface layers. Molecular microbiology 10:911-916.
282. Toba T, Virkola R, Westerlund B, Bjorkman Y, Sillanpaa J, Vartio T, Kalkkinen N, Korhonen TK. 1995. A collagen-binding S-layer protein in

Lactobacillus crispatus. Applied and environmental microbiology 61:2467-2471.

283. Masuda K, Kawata T. 1983. Distribution and chemical characterization of regular arrays in the cell walls of strains of the genus *Lactobacillus*. FEMS microbiology letters 20:145-150.
284. Messner P, Sleytr UB. 1992. Crystalline bacterial cell-surface layers, p 213-275, Advances in microbial physiology, vol 33. Elsevier.
285. Lortal S. 1993. Crystalline surface-layers of the genus *Lactobacillus*, p 57-65, Advances in bacterial paracrystalline surface layers. Springer.
286. Letellier L, Boulanger P, Plançon L, Jacquot P, Santamaria M. 2004. Main features on tailed phage, host recognition and DNA uptake. Frontiers in bioscience 9:1228-1339.
287. Roessner CA, Ihler GM. 1986. Formation of transmembrane channels in liposomes during injection of lambda DNA. Journal of biological chemistry 261:386-390.
288. Feucht A, Schmid A, Benz R, Schwarz H, Heller K. 1990. Pore formation associated with the tail-tip protein pb2 of bacteriophage T5. Journal of biological chemistry 265:18561-18567.
289. Boulanger P, le Maire M, Bonhivers M, Dubois S, Desmadril M, Letellier L. 1996. Purification and structural and functional characterization of FhuA, a transporter of the *Escherichia coli* outer membrane. Biochemistry 35:14216-14224.
290. Trojet SN, Caumont-Sarcos A, Perrody E, Comeau AM, Krisch H. 2011. The gp38 adhesins of the T4 superfamily: a complex modular determinant of the phage's host specificity. Genome biology and evolution 3:674-686.
291. Boulanger P, Letellier L. 1992. Ion channels are likely to be involved in the two steps of phage T5 DNA penetration into *Escherichia coli* cells. Journal of biological chemistry 267:3168-3172.
292. São-José C, Baptista C, Santos MA. 2004. *Bacillus subtilis* operon encoding a membrane receptor for bacteriophage SPP1. Journal of bacteriology 186:8337-8346.
293. São-José C, Lhuillier S, Lurz R, Melki R, Lepault J, Santos MA, Tavares P. 2006. The ectodomain of the viral receptor YueB forms a fiber that triggers

ejection of bacteriophage SPP1 DNA. Journal of biological chemistry 281:11464-11470.

294. Lurz R, Orlova EV, Günther D, Dube P, Dröge A, Weise F, van Heel M, Tavares P. 2001. Structural organisation of the head-to-tail interface of a bacterial virus. Journal of molecular biology 310:1027-1037.

295. Orlova EV, Gowen B, Dröge A, Stiege A, Weise F, Lurz R, van Heel M, Tavares P. 2003. Structure of a viral DNA gatekeeper at 10 Å resolution by cryo-electron microscopy. The EMBO journal 22:1255-1262.

296. Tavares P, Lurz R, Stiege A, Rückert B, Trautner TA. 1996. Sequential headful packaging and fate of the cleaved DNA ends in bacteriophage SPP1. Journal of molecular biology 264:954-967.

297. Räisänen L, Schubert K, Jaakonsaari T, Alatossava T. 2004. Characterization of lipoteichoic acids as *Lactobacillus delbrueckii* phage receptor components. Journal of bacteriology 186:5529-5532.

298. Young F, Smith C, Reilly B. 1969. Chromosomal location of genes regulating resistance to bacteriophage in *Bacillus subtilis*. Journal of bacteriology 98:1087-1097.

299. Wendlinger G, Loessner MJ, Scherer S. 1996. Bacteriophage receptors on *Listeria monocytogenes* cells are the N-acetylglucosamine and rhamnose substituents of teichoic acids or the peptidoglycan itself. Microbiology 142:985-992.

300. Spears PA, Havell EA, Hamrick TS, Goforth JB, Levine AL, Abraham ST, Heiss C, Azadi P, Orndorff PE. 2016. *Listeria monocytogenes* wall teichoic acid decoration in virulence and cell-to-cell spread. Molecular microbiology 101:714-730.

301. Spinelli S, Campanacci V, Blangy S, Moineau S, Tegoni M, Cambillau C. 2006. Modular structure of the receptor binding proteins of *Lactococcus lactis* phages: the RBP structure of the temperate phage TP901-1. Journal of Biological Chemistry 281:14256-14262.

302. Tremblay DM, Tegoni M, Spinelli S, Campanacci V, Blangy S, Huyghe C, Desmyter A, Labrie S, Moineau S, Cambillau C. 2006. Receptor-binding protein of Lactococcus lactis phages: identification and characterization of the saccharide receptor-binding site. Journal of bacteriology 188:2400-2410.

303. Ricagno S, Campanacci V, Blangy S, Spinelli S, Tremblay D, Moineau S, Tegoni M, Cambillau C. 2006. Crystal structure of the receptor-binding protein head domain from *Lactococcus lactis* phage bIL170. Journal of virology 80:9331-9335.

304. McCabe O, Spinelli S, Farenc C, Labbé M, Tremblay D, Blangy S, Oscarson S, Moineau S, Cambillau C. 2015. The targeted recognition of *Lactococcus lactis* phages to their polysaccharide receptors. Molecular microbiology 96:875-886.

305. Dupont K, Janzen T, Vogensen FK, Josephsen J, Stuer-Lauridsen B. 2004. Identification of *Lactococcus lactis* genes required for bacteriophage adsorption. Applied and environmental microbiology 70:5825-32.

306. Sadovskaya I, Vinogradov E, Courtin P, Armalyte J, Meyrand M, Giaouris E, Palussière S, Furlan S, Péchoux C, Ainsworth S. 2017. Another brick in the wall: a rhamnan polysaccharide trapped inside peptidoglycan of *Lactococcus lactis*. MBio 8:e01303-17.

307. Andre G, Kulakauskas S, Chapot-Chartier M-P, Navet B, Deghorain M, Bernard E, Hols P, Dufrêne YF. 2010. Imaging the nanoscale organization of peptidoglycan in living *Lactococcus lactis* cells. Nature communications 1:27.

308. Chapot-Chartier MP, Vinogradov E, Sadovskaya I, Andre G, Mistou MY, Trieu-Cuot P, Furlan S, Bidnenko E, Courtin P, Pechoux C. 2010. Cell surface of *Lactococcus lactis* is covered by a protective polysaccharide pellicle. Journal of biological chemistry 285:10464-71.

309. Ainsworth S, Sadovskaya I, Vinogradov E, Courtin P, Guerardel Y, Mahony J, Grard T, Cambillau C, Chapot-Chartier MP, Sinderen D. 2014. Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. MBio 5:e00880-14.

310. Farenc C, Spinelli S, Vinogradov E, Tremblay D, Blangy S, Sadovskaya I, Moineau S, Cambillau C. 2014. Molecular insights on the recognition of a *Lactococcus lactis* cell wall pellicle by the phage 1358 receptor binding protein. Journal of virology 88:7005-15.

311. Vinogradov E, Sadovskaya I, Grard T, Murphy J, Mahony J, Chapot-Chartier M-P, van Sinderen D. 2018. Structural studies of the cell wall polysaccharide from *Lactococcus lactis* UC509.9. Carbohydrate research 461:25-31.

312. Ostergaard Breum S, Neve H, Heller KJ, Vogensen FK. 2007. Temperate phages TP901-1 and phiLC3, belonging to the P335 species, apparently use different pathways for DNA injection in *Lactococcus lactis* subsp. *cremoris* 3107. FEMS Microbiology letters 276:156-64.

313. Labrie SJ, Moineau S. 2007. Abortive infection mechanisms and prophage sequences significantly influence the genetic makeup of emerging lytic lactococcal phages. Journal of bacteriology 189:1482-1487.

314. Kelleher P, Mahony J, Schweinlin K, Neve H, Franz CM, van Sinderen D. 2018. Assessing the functionality and genetic diversity of lactococcal prophages. International journal of food microbiology 272:29-40.

Chapter II

Complete genome sequence of *Lactococcus lactis* subsp. *cremoris* 3107, host for the model lactococcal P335 bacteriophage TP901-1

Part of this chapter has been published in Microbiology Resource Announcements: Complete genome sequence of *Lactococcus lactis* subsp. *cremoris* 3107, host for the model lactococcal P335 bacteriophage TP901-1. **Andrea Erazo, Jennifer Mahony, Francesca Bottacini, Philip Kelleher and Douwe van Sinderen**

Note: REBASE analysis and assignment of methylome motifs was performed by Dr. Richard J. Roberts of New England Biolabs, USA. Genome assembly was performed by Dr. Jennifer Mahony and Dr. Francesca Bottacini, School of Microbiology and APC Microbiome Ireland, University College Cork, Ireland.

Chapter II contents

2 Abstract

In the current study, the complete genome sequence of *Lactococcus lactis* subsp. *cremoris* (*L. lactis*) 3107, a dairy starter strain and host for the model P335 bacteriophage TP901-1, is reported. The circular chromosome of *L. lactis* 3107 is among the smallest genomes of currently sequenced lactococcal strains. *L. lactis* 3107 harbours a complement of six plasmids which encode traits including lactose utilization and casein metabolism, reflecting its adaptation to the nutrient-rich dairy environment. Additionally, an improved method for electrotransformation of *L. lactis* 3107 and a multiplex PCR methodology for the identification of the plasmids of this dairy-associated strain are presented.

2.1 Introduction

The dairy fermentation industry relies on starter or adjunct cultures, predominantly composed of lactic acid bacteria (LAB), to produce high quality end products (1). Like all bacteria, individual strains of LAB are susceptible to infection by bacteriophages (phages), which are ubiquitous in the dairy environment, while they can be rather insensitive to pasteurization treatments (2). Therefore, phage infection of starter and adjunct strains during the fermentation process is a major concern for the dairy fermentation industry as this may negatively impact on the final product quality and production regimes, thereby causing significant economic losses (3). Among the LAB, strains of the industrially significant species *Lactococcus lactis* are particularly susceptible to infection by phages (4, 5). For millennia *L. lactis* has been safely employed as a starter culture in dairy fermentations, and for this reason it has been granted "GRAS" (Generally Regarded as Safe) status (6).

The advent of so-called next generation sequencing (NGS) technologies has transformed life sciences, making genome sequence analysis easily accessible. This, for example, has facilitated the availability of some hundreds lactococcal genomes in the publically accessible NCBI (National Centre for Biotechnology Information) database. In the current study we applied Single Molecule Real Time (SMRT) (7, 8) and Illumina sequencing (9) to sequence the complete genome of *L. lactis* subsp. *cremoris* 3107, which is host for the model lysogenic P335 phage TP901-1 (10).

SMRT sequencing, which was developed by Pacific Biosciences (7, 8), is one of the latest sequencing technologies and is currently the only technology which facilitates the detection of base modifications with primary sequencing analysis. Consequently, SMRT sequencing can be applied to determine the methylome of strains allowing the study of genome methylation by DNA methyltransferases (MTases). MTases are encoded by prokaryotes to confer base modifications known to be involved in a variety of processes including cell cycle regulation, DNA repair and pathogenesis (11). Studies have demonstrated that invasion of a bacterial cell by foreign DNA (such as a phage) may be prevented by restriction-modification (R-M) systems. A typical R-M system consists of an MTase linked to a cognate restriction endonuclease (REase) activity, where the MTase methylates 'self' DNA, while the REase targets invading non-methylated DNA (12-15), thereby preventing phage propagation. R-M systems have been classified into four different types (Types I - IV)

based on sub-unit composition, ATP (or GTP) requirements and cleavage mechanisms (16).

In order to study the molecular events that underpin TP901-1 interactions with its *L. lactis* 3107 host, the cell wall polysaccharide (CWPS) biosynthesis operon of this strain was previously analysed (17). The *L. lactis* 3107 genomic region that is responsible for CWPS biosynthesis belongs to the so-called type C genotype, with its variable region (subtype C_2) being responsible for the synthesis of a surface-exposed polysaccharide pellicle (PSP) component which acts as the TP901-1 phage receptor (17). Further studies focusing on the interactions of TP901-1 with its *L. lactis* 3107 host have been limited as this strain is poorly transformable rendering it recalcitrant to genetic manipulation (18, 19). Here, the *L. lactis* 3107 genome was analysed in terms of its general genome characteristics, the traits imparted on the strain by its plasmids, its methylome, and impact of methylation on the transformability of the strain. Additionally, an optimised electrotransformation protocol was developed and described for this strain.

2.2 Materials and Methods

2.2.1 *L. lactis* 3107 growth conditions and plasmid

L. lactis 3107 was routinely cultured at 30 °C in M17 broth/agar (Oxoid, U.K.) supplemented with 0.5 % glucose (v/v) (GM17 broth/agar) without agitation. The pNZ44 vector expressing a chloramphenicol resistance (Cm^r) gene (20) was employed in electrotransformation assays.

2.2.2 *L. lactis* 3107 genome sequencing

For Pacbio sequencing, *L. lactis* 3107 cell pellets (from an overnight culture, see below) containing 10^9 colony forming units were provided to GATC Biotech Ltd. (Germany) to perform chromosomal DNA extraction, library construction and Single Molecule Real Time (SMRT) sequencing on a Pacific Biosciences RS (run 1) and RSII (run 2) sequencing platform. The library for PacBio sequencing was prepared using the SMRTbell Template Prep Kit with 8-12 kb inserts according to the manufacturer's instructions (Pacific Biosciences, Menlo Park, CA, USA). *De novo* genome assembly of the SMRT sequencing data was performed using the RS_HGAP_Assembly.2 protocol (default parameters) implemented in the Pacific Biosciences SMRT Analysis Portal (version 2.3.1). Quality filtering was performed automatically during assembly by the SMRT portal P-filter module. Two SMRT cells were used for PacBio sequencing to achieve an initial assembly of the derived 52,984 filtered reads into 23 contigs with an N_{50} contig length of 337,497 bp and an average coverage of 60.37-fold.

For Illumina-based sequencing 5 µg chromosomal DNA from *L. lactis* 3107 was extracted using phenol-chloroform based extractions as previously described (21) following overnight growth of the strain at 30 °C in M17 broth (Oxoid, U.K.) supplemented with 0.5 % glucose (v/v). Genomic libraries were constructed using the TruSeq DNA PCR-Free LT Kit (Illumina®) and 2.5 µg of genomic DNA, which was fragmented with a Bioruptor NGS ultrasonicator (Diagenode, USA) followed by size evaluation using Tape Station 2200 (Agilent Technologies). Library samples were loaded into a Flow Cell V3 600 cycles (illumina®) and draft genome Illumina sequencing was performed on a MiSeq genomic platform (Illumina, UK) at GenProbio srl. (Parma, Italy). The Illumina reads were mapped on corresponding PacBio scaffolds to provide confidence in the generated sequence quality and to resolve base

conflicts using Bowtie2 v2.2.7 achieving a mapping coverage of 847-fold for the chromosome and on average 4,776-fold for the plasmids. Remaining low quality regions or sequence conflicts were resolved by primer walking and Sanger sequencing of PCR products (Eurofins MWG Operon, Germany).

2.2.3 General feature predictions

Following final genome assembly, putative protein-encoding genes were identified using the prediction software Prodigal (version 2.0) (22). Protein-encoding genes were automatically annotated using a BLASTP v2.2.26 (cut-off *E*-value of 0.0001) sequence alignments against the non-redundant protein (nr) database curated by NCBI (ftp://ftp.ncbi.nih.gov/blast/db/). Following automatic annotation, the obtained Open Reading Frames (ORFs) were manually inspected and refined using the genome visualisation tool Artemis v16 (23). Finally, ORF annotations were refined further where necessary using alternative functional searches using PFAM database (24), and the Clusters of Orthologous Groups (COG) database (25). Predicted transfer RNA (tRNA) and ribosomal RNA (rRNA) genes were identified using tRNA-scan-SE v1.4 (http://lowelab.ucsc.edu/tRNAscan-SE/) and RNAmmer v1.2 (http://www.cbs.dtu.dk/services/RNAmmer/), respectively. Using Artemis v16, the predicted RNA-specifying loci were manually added to the genome annotation. PHASTER (26) analysis of the *L. lactis* 3107 genome was also performed to identify potential prophages.

Additionally, DNA base modification analysis was performed by means of SMRT sequencing. This was performed using the RS_Modification_and_Motif_Analysis.1 protocol and the finished genome assembly as a reference file.

2.2.4 Nucleotide sequence accession numbers and SRA accession number

The complete chromosome and plasmid complement of *L. lactis* subsp. *cremoris* 3107 were deposited in GenBank under accession no. CP031538 (chromosome), CP031539 (p3107A), CP031540 (p3107B), CP031541 (p3107C), CP031542 (p3107D), CP031543 (p3107E), and CP031544 (p3107F). The SMRT and Illumina raw reads have been deposited in Sequence Read Archive (SRA) under accession no. PRJNA438435.

2.2.5 Multiplex PCR targeting *L. lactis* 3107 plasmids

Primers employed in the multiplex PCR system targeting each of the six *L. lactis* 3107 plasmids are presented in Table 2.1. Each primer pair was designed to generate differently sized products as detailed in Table 2.1. Primer characteristics were analysed using the online tool OligoCalc: Oligonucleotide Properties Calculator (http://biotools.nubic.northwestern.edu/OligoCalc.html).

Multiplex-PCR was performed on an *L. lactis* 3107 colony using Thermo Scientific Phusion Green Hot Start II High-Fidelity PCR Master Mix (Fisher Scientific, Ireland) using the following PCR conditions: initial denaturation at 98 °C x 30 sec, followed by 30 cycles of denaturation at 98 °C x 10 sec, annealing at 50 °C x 30 sec and extension at 72 °C x 90 sec; and a final extension of 72 °C x 10 min. Following amplification, agarose gel electrophoresis using a 1 % agarose gel was performed using a standard procedure (27) to visualize generated amplicons, thus confirming the presence of all six 3107 plasmids which was indicated by a product of the appropriate size corresponding to each plasmid.

Table 2.1: Multiplex-PCR primers for *Lactococcus lactis* subsp. *cremoris* 3107 plasmids.

Target	Primer Name	Primer Sequence (5' to 3')	Annealing Temp. (ºC)	Product Size (bp)
L. lactis 3107 plasmid				
p3107A	p3107A_F	GGAAGGTGGCAGAGCATA	50.3	857
	p3107A_R	CGTGTCCTGGTTTGCTAT	48.0	
p3107B	p3107B_F	GGGAGTGAACCAGTAACA	48.0	740
	p3107B_R	CCAGGAGAAGAAGTACGA	48.0	
p3107C	p3107C_F	CGTCCCAATACCAACTGT	48.0	417
	p3107C_R	GCGGTTCGTTTCTTCTCA	48.0	
p3107D	p3107D_F	GCGTATCTATGGCTGTCA	48.0	228
	p3107D_R	GCGTCCTTTGATTCATGAG	48.9	
p3107E	p3107E_F	GCCGTGTCTTGCATTGAT	48.0	1,012
	p3107E_R	GGCTTGCACCTTTACCTC	50.3	
p3107F	p3107F_F	CCGAAGAGCCACCGAAAG	52.6	616
	p3107F_R	GCTTGTAGCTGTTGTCGT	48.0	

2.2.6 Electrotransformation of *L. lactis* 3107

This procedure was performed as described previously (28, 29) with the following modifications: 50 mL of M17 (Oxoid, U.K.) broth supplemented with 0.5 M sucrose (w/v), 0.5 % glucose (v/v) and 1 % glycine (v/v) was inoculated with 6 % fresh overnight culture. This culture was incubated at 30 °C until an approximate OD_{600nm} of 0.6 – 0.8 was reached. Following centrifugation at 1900 × *g* for 15 minutes at 4 °C, the pellet was washed twice in an ice-cold wash solution composed of 0.5 M sucrose (w/v) and 10 % glycerol (v/v). The obtained cell pellet was re-suspended in 500 µl wash solution, aliquoted and used immediately or frozen at -80 °C. 45 µl of fresh or thawed competent cell suspension was mixed with 5 µl of pNZ44 plasmid DNA and electroporated with an Electro Cell Manipulator® Precision Pulse™ (BTX® Harvard Apparatus, Massachusetts, U.S.A.) at 2 kV, 200 Ω and 25 µF in an ice-cold 0.2 cm electroporation cuvette (Cell Projects Distributor, Kent, U.K.). Cells were re-suspended in 950 µl ice-cold recovery media composed of GM17 broth supplemented with 20 mM $MgCl_2$ and 2 mM $CaCl_2$. Following this, the mixture was incubated at 30 °C for 2.5 hours. The recovered cells were centrifuged at 1,900 × *g* for 1 minute, resuspended in 100 µl of the recovery media, spread plated on GM17 agar supplemented with 2.5 mM $CaCl_2$, 2.5 mM $MgCl_2$ and 2.5 µg ml^{-1} Cm for the selection of pNZ44, and were then incubated anaerobically at 30 °C for 48 hours. The presence of pNZ44 in the generated transformants was confirmed by PCR amplification of a 500 bp specific to the p44 lactococcal promoter and multiple cloning site (MCS) sequence of pNZ44 using pNZ44_F 5'-CTAATGTCACTAACCTGCCCCG-3' and pNZ44_R 5'-GCTTTATCAACTGCTGCT-3' primers. Following this, selected colonies were cultured overnight in GM17 broth supplemented with 5 µg ml^{-1} Cm. Additionally, a multiplex-PCR using the primers presented in Table 2.1 was performed on transformants to confirm the integrity of Cm^r colonies as *L. lactis* 3107 derivatives.

2.3 Results

2.3.1 General chromosome and plasmid features of *L. lactis* 3107

In this study the *L. lactis* 3107 genome was sequenced (30). The complete genome sequence of this dairy starter strain consists of a single circular chromosome of 2,378,982 bp (35.82 % G+C content) in length, while the strain also harbours six plasmids ranging in size from 2,232 – 60,216 bp (Table 2.2). General feature analysis of the chromosome of *L. lactis* 3107 identified a total of 2,380 predicted CDS (Coding Sequences) of which 101 (4.24 %) are pseudogenes, which is comparable to the number of pseudogenes present in other *L. lactis* subsp. *cremoris* strains (31). Further genome analysis demonstrated that 80.2 % of the CDS can be functionally assigned using BLAST (Basic Local Alignment Search Tool) based on *in silico* predictions, while the remaining 19.8 % are assigned as hypothetical proteins. The genome of *L. lactis* 3107 contains 164 insertion sequence (IS) elements/transposase-encoding genes, including 20 copies of IS*712H* and 31 copies of IS*982B*. Prophage analysis using PHASTER (26) identified two intact prophages with lengths of 42.6 and 48.8 Kb, and four (apparently) incomplete prophages ranging in size from 11 to 36.8 Kb. Methylome analysis of the genome of this strain did not reveal any methylated motifs, although restriction-modification (R-M) associated genes are present in the genome of 3107. These results have been published in the REBASE Genomes database (http://tools.neb.com/genomes/view.php?genome_id=35512) and are presented in Table 2.3.

Analysis of the six plasmids harboured by *L. lactis* 3107 was also performed. This extra-chromosomal DNA represents a total of 161,686 bp and is predicted to represent 172 CDSs of which 29 (16.86 %) are annotated as hypothetical proteins. The G+C content of each plasmid ranges from 31.63 to 37.63 %, which is comparable to G+C content of previously reported *L. lactis* plasmids (32). Analysis of the plasmid-encoded genes revealed that plasmids p3107A, p3107B, p3107C and p3107E harbour homologues of the replication gene *repB*. The 3107 plasmid-encoded RepB proteins show 58 – 100 % amino acid identity to the replication protein RepB, which has been associated with theta-type replication (33), a replication mechanisms used by the majority of lactococcal plasmids (32, 34). The two smallest plasmids, p3107D and p3107F, do not harbour a *repB* homologue, which may not be surprising as small plasmids (< 10 Kb) typically replicate via the rolling circle replication (RCR)

mechanism (32). The second largest plasmid p3107A encodes (predicted) traits such as lactose and casein metabolism, an oligopeptide permease system, a mobilisation protein MobA, a universal stress protein UspA, and aminoglycoside antibiotic resistance. The largest plasmid p3107B encodes putative functions such as an R-M type I system, mobilisation proteins MobC and MobD, as well as cold-shock response proteins. Plasmids p3107C and p3107E are predicted to encode supplementary mineral uptake systems including Mg^{2+} and Co^{2+} transporters. In addition, a homologue of this system (L3107_1068) is present on the 3107 chromosome. Furthermore, p3107C encodes the mobilisation protein MobC. The smallest plasmids p3107F and p3107D encode two and three proteins, respectively, putatively involved in plasmid replication.

2.3.2 Multiplex PCR targeting *L. lactis* 3107 plasmids

In order to accurately identify the presence of the six plasmids harboured by *L. lactis* 3107, to be used, for example, for verification purposes (for *L. lactis* 3107 transformants), a multiplex PCR was developed which targets each of these six plasmids (Table 2.1). Each set of primers has a similar annealing temperature and was designed to amplify a specific DNA region of a distinctive size in each plasmid. Indeed, Figure 2.1 shows that DNA regions of the expected size were amplified where each fragment represents the presence of each of the six plasmids of *L. lactis* 3107.

Table 2.2: General genome and plasmid features of *L. lactis* 3107.

***L. lactis* subsp. *cremoris* 3107**	**Length (bp)**	**CDS**	**tRNA features**	**rRNA features**	**Pseudo genes**	**IS elements/ transposases**	**Prophage(s)**	**GC %**	**Pacbio genome coverage (fold)**	**Illumina genome coverage (fold)**
genome	2,378,982	2,279	60	19	101	215	2 Int 4 Inc	35.82	65.01	846.88
p3107A	50,160	52	-	-	-	12	-	35.64	228.24	1633.72
p3107B	60,216	69	1	-	-	19	-	33.38	230.17	1550.52
p3107C	26,709	24	-	-	-	9	-	37.63	455.12	2467.09
p3107D	2,232	3	-	-	-	-	-	33.56	615.77	11381.27
p3107E	18,170	22	-	-	-	5	-	34.81	491.65	4443.82
p3107F	4,199	2	-	-	-	-	-	31.63	332.51	7180.93

Int: predicted complete prophage, Inc: predicted incomplete/partial prophage, -: not present

Table 2.3: Restriction-Modification (R-M) associated genes present in the genome of *L. lactis* 3107.

R-M System	ORF L3107_		Predicted protein	Predicted recognized sequence
Type I				
	1148	R	Type I restriction subunit (pseudogene)	Unknown
	1150	S	Type I specificity subunit	Unknown
	1153	S	Type I specificity subunit	Unknown
	1154	M	Type I N6-adenine DNA methyltransferase	Unknown
	2087	S	Type I specificity subunit	Unknown
Type II				
	0842	M	Type II N4-cytosine or N6-adenine DNA methyltransferase	Unknown

R: Restriction, M: Methylase, S: Specificity.

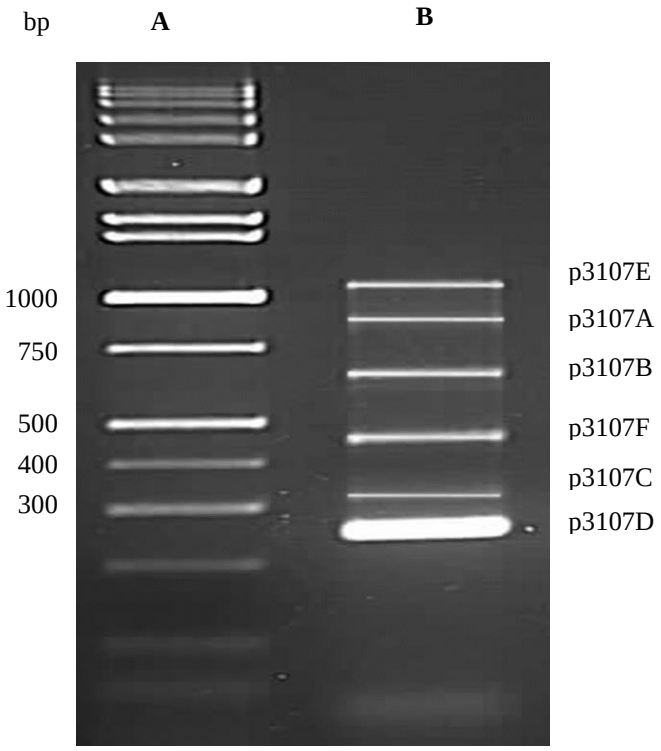


Figure 2.1: Agarose gel image of amplicons generated following multiplex-PCR targeting *L. lactis* 3107 plasmids. Lane A exhibits a DNA Ladder (relevant sizes are indicated; Fisher BioReagents, Ireland); Lane B exhibits six PCR products of varying sizes representing each of the six *L. lactis* 3107 plasmids including p3107E (1,012 bp), p3107A (amplicon size 857 bp), p3107B (amplicon size 740 bp), p3107F (amplicon size 616 bp), p3107C (amplicon size 417 bp) and p3107D (amplicon size 228 bp).

2.3.3 Optimization of electrotransformation efficacy of *L. lactis* 3107

In order to enhance the electrotransformation efficiency of *L. lactis* 3107, various parameters of previously described electrotransformation protocols (28, 29) were assessed (Table 2.4). Following modification of each parameter, the transformation efficiency of 3107 was determined using the high copy number vector pNZ44. First, analysis of growth of *L. lactis* 3107 in M17 broth supplemented with 0.5 % glucose (GM17) or lactose (LM17) (data not shown) was performed, and based on the superior growth of *L. lactis* 3107 on GM17, this medium was selected as the growth medium to prepare this strain for electrotransformation.

In order to enhance transformation efficacy, while maintaining sufficient cell viability, different concentrations of cell wall weakening agents (e.g. glycine, DL-threonine and penicillin) were tested individually and in combination with various sucrose concentrations to create osmotic stabilization (Table 2.4). Results showed that at 0.5 M sucrose, the highest glycine concentration tolerated by 3107 is 1 % and following electroporation, the highest efficiency of transformation ($5.60 \times 10^4 \pm 1.00 \times 10^3$ Cm^r colonies per µg of DNA transformed) was observed (Table 2.4), when 3107 is grown under these conditions. Glycine levels above and below 1 % in combination with sucrose molarities at and below 0.5 M were also tested, although this did not improve transformation efficiency (Table 2.4).

In addition to glycine, the cell wall weakening agent ampicillin was used in an attempt to further enhance cell wall permeability of the strain. However, exposure to ampicillin caused a 3-log reduction in the viability of the cells before electroporation (results not shown) suggesting that the already weakened cell wall due to the exposure to glycine cannot withstand further damage as it results in cell death. When DL-threonine was applied instead of glycine a negative effect on the efficiency of transformation of 3107 was observed when compared to the number of transformants obtained when competent cells were prepared using glycine (results not shown).

A 4 – 6 % inoculum of *L. lactis* 3107 overnight culture optimizes the growth of 3107 in the presence of 1 % glycine and 0.5 M sucrose without affecting the transformation efficiency of this strain (as $5.60 \times 10^4 \pm 1.00 \times 10^3$ Cm^r colonies per µg of DNA transformed were obtained). Results also indicated that an OD_{600nm} value between 0.6 and 0.8 represents the optimal range for the preparation of competent cells. A voltage of 2 kV yielded optimal transformation efficiencies compared to all other tested conditions (Table 2.4).

The impact of freezing the prepared competent cells at -80 °C prior to electroporation was also analysed. Results showed that storing 3107 competent cells at -80 °C negatively impacts cell viability and consequently reduces the efficiency of transformation.

The presence of active R-M systems has a negative effect on the introduction of foreign DNA into strains (3). Based on this and the presence of R-M associated genes in the genome of 3107, the impact of these potentially active systems on the transformation efficacy of 3107 was assessed. Isolation of previously introduced pNZ44 into *L. lactis* 3107 and re-introduction of this plasmid into *L. lactis* 3107 had no effect on the transformation efficiency of this strain, suggesting that the R-M systems are not active in this strain or that they do not impair the transformation of this plasmid.

Multiplex-PCR using the primers detailed in Table 2.1 performed on transformants confirmed the integrity of Cm^r colonies as *L. lactis* 3107 derivatives, thus demonstrating successful electrotransformation of *L. lactis* 3107.

Table 2.4: Different parameters tested to improve the electrotransformation efficiency of *L. lactis* 3107; optimal parameters in bold.

	Optimal parameters	Tested parameters										Previously used method for *L. lactis* 3107 (18)	Suggested method for *Lactococcus* strains (28)
Glycine (%)	**1**	0.6-0.8	> 1	0.8-1.2	1	1	1	1	1	1	1	1	1
Sucrose (M)	**0.5**	0.5	0.5	0.1-0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.2	0.5
3107 (%)	**4-6**	6	6	6	2	> 8	6	6	6	6	6	1	1
OD_{600nm}	**0.6-0.8**	0.6-0.8	N	0.6-0.8	N	0.6-0.8	0.5 -0.6	0.6-0.8	0.6-0.8	0.6-0.8	0.6-0.8	0.7	0.6
Ampicillin (µg mL^{-1})	-	-	-	-	-	-	-	5-20	-	-	-	-	-
kV	**2**	2	2	2	2	2	2	2	1.5-1.9	2.1-2.5	2	2	1.7
-80 °C	-	-	-	-	-	-	-	-	-	-	Y	Y	Y
No. of Cm^r colonies per µg of DNA transformed	**5.60×10^4 ± 1.00×10^3**	4.00×10^3 – 5.00×10^3 ± 1.00×10^3	-	4.00×10^2 – 5.00×10^3 ± 1.00×10^3	-	0	2.00×10^3 ± 1.00×10^3	3.00×10^3 ± 1.00×10^3	2.0×10^4 – 4.00×10^4 ± 1.00×10^3	1.20×10^4 – 2.00×10^4 ± 1.00×10^3	6.00×10^3 ± 1.00×10^3	$4.00 \times 10^2 \pm 4.00 \times 10^2$	Cells failed to reach necessary OD_{600nm}

N: necessary OD_{600nm} not reached, -: not applied, Y: competent cells frozen at -80 °C prior to electroporation

2.4 Discussion

Recent advances in NGS technologies have made it easier and faster to obtain a complete and high-quality sequence of a bacterial genome. Here, a combination of SMRT and Illumina sequencing technologies were applied to obtain the complete sequence of the *L. lactis* subsp. *cremoris* 3107 genome (30). The circular chromosome of this dairy-associated strain, which is the host for the model P335 phage TP901-1, is among the smallest genomes of currently sequenced lactococcal strains. The chromosomal features of this strain were assessed, revealing a genome which harbours a high number of pseudogenes and transposable elements. It also contains a number of apparently intact and incomplete prophages, which are common characteristics of *L. lactis* subsp. *cremoris* strains (31). The presence of such a high number of transposons and pseudogenes within this relatively small lactococcal chromosome suggests that the *L. lactis* 3107 genome has undergone significant genome decay while adapting to its milk environment. *L. lactis* 3107 is no exception in this respect, as various *L. lactis* subsp. *cremoris* strains isolated from the dairy environment have been reported to suffer from genome decay (4, 31, 36, 37). It has also been suggested that the high number of transposable elements and prophages present in *L. lactis* subsp. *cremoris* strains may be the outcome of industrial pressure (31). Thus, prophage-containing starter strains represent a risk factor to the dairy fermentation industry and, therefore, demands a comprehensive functional assessment before their utilization for the production of fermented dairy products.

L. lactis strains seem to have adapted to the dairy environment by the acquisition of additional metabolic capabilities and the loss of redundant functions. Studies (32, 38-41) have shown that dairy-associated *L. lactis* strains acquire specific plasmids, such as those encoding lactose and citrate metabolism for niche adaptation. Indeed, *L. lactis* 3107 harbours a complement of six plasmids, of which p3107A encodes lactose metabolism allowing this strain to rapidly ferment lactose (42), extracellular casein hydrolysis (43) and uptake of the resulting peptides (44). Plasmid p3107A also encodes intracellular peptidases to facilitate metabolism of casein-derived peptides, which may contribute to the formation of aroma compounds and flavour development in the final product (45). Finally, p3107A encodes a universal stress protein (UspA) which is present in all sequenced dairy plasmids (32) and has been shown to be expressed in the absence of carbon and nutrients, heat exposure, oxidation and in response to other environmental stress conditions (46). Homologues

of the p3107B plasmid-encoded cold-shock protein have also been shown to enhance survival of the strain during harsh dairy fermentation conditions. This protein protects the cell from rapid temperature downshift by decreasing enzyme activity, efficiency of transcription and translation, among other functions (47). Additionally, cold-shock proteins have been shown to provide tolerance to osmotic, oxidative, starvation, pH and ethanol stress (48). Genes encoding stress-associated proteins are presumed to be acquired due to the conditions that dairy-associated strains endure during fermentation including changes in pH, temperature and osmolarity (49). The 3107 chromosome as well as plasmids p3107C and p3107E encode bivalent metal ion uptake systems including Mg^{2+} and Co^{2+} transporters which have previously been described as potentially beneficial dairy adaptations (32, 50). The expression of such systems inhibits growth of strains exposed to $CoCl_2$ concentrations of > 2.0 mM, while its inactivation allows cells to grow above this concentration (51). Mg^{2+} and Co^{2+} transporter systems are present in some *L. lactis* strains (51). p3107A also encodes a putative aminoglycoside 3-N-acetyltransferase which has previously been associated with gentamicin antibiotic resistance (52, 53). Antibiotic resistance genes providing resistance to tetracycline have previously been identified on lactococcal plasmids from dairy strains (54, 55). This highlights the importance of genomic assessment and antibiotic testing as the increase in drug-resistant bacteria is a major health concern. Thus, before using dairy strains in food production antibiotic testing must be performed to avoid the undesirable spread of antibiotic resistance within the food chain. The plasmid complement of *L. lactis* 3107 harbours homologues of *mobA* (p3107A), *mobC* (p3107B/C) and *mobD* (p3107C) which have been shown to play a role in the mobilisation of plasmids in *L. lactis* (56, 57). During such plasmid transfer the plasmid receives a single stranded break by a nickase enzyme and the resulting linear ssDNA strand is passed to a recipient cell (58). The *mobA* and *mobD* genes are predicted to encode nickases, and *mobC* is predicted to encode a protein presumed to form a relaxosome with an associated nickase (56, 57). This type of naturally occurring DNA transfer process is advantageous in food grade applications to exert beneficial traits to industrial strains (56).

Even though the *L. lactis* 3107 chromosome is predicted to contain R-M type I and II associated genes, SMRT sequencing did not detect any methylated DNA bases within the host genome. This suggests an absence of active R-M systems in 3107, though it should be kept in mind that SMRT sequencing mainly detects N^6-

methyladenine base modifications (59). In addition, isolating pNZ44 DNA from a previously transformed *L. lactis* 3107 strain and re-introducing it into the wild type *L. lactis* 3107 did not have an effect on transformation efficiency, which indicates that no active R-M systems are present in *L. lactis* 3107 that target pNZ44. In support of this latter observation, the gene predicted to encode a Type I restriction subunit was identified as a pseudogene (Table 2.3).

Since *L. lactis* 3107 is the host strain for the model lactococcal P335 phage TP901-1, it is has featured in various phage-host interaction studies (17-19, 35). However, the lack of refined protocols for the efficient electrotransformation of this strain has restricted the genetic manipulation of the strain. Protocols for the electrotransformation of *L. lactis* subsp. *cremoris* strains (29) and other species of *Lactococcus* (28) are currently available. *L. lactis* 3107 displays a low efficiency of transformation using these methods (as shown in Table 2.4) (28, 29). Optimisation of the transformation protocol for this strain creates the possibility of adapting systems for the genetic manipulation of this dairy-associated strain.

In conclusion, we determined the complete genetic make-up of *L. lactis* 3107, host for the model P335 group phage TP901-1. Additionally, protocols for efficient transfer of DNA into 3107 and for multiplex-PCR plasmid targeting were developed. Thus, this study introduces new possibilities for future studies focused on the interactions of TP901-1 with its *L. lactis* 3107 host at a molecular level.

2.5 References

1. Teuber M. 1993. Lactic acid bacteria. Biotechnology set second edition:325-366.
2. Deveau H, Labrie SJ, Chopin M-C, Moineau S. 2006. Biodiversity and classification of lactococcal phages. Applied and environmental microbiology 72:4338-4346.
3. Coffey A, Ross RP. 2002. Bacteriophage-resistance systems in dairy starter strains: molecular analysis to application, p 303-321, Lactic acid bacteria: genetics, metabolism and applications. Springer.
4. Ainsworth S, Zomer A, de Jager V, Bottacini F, van Hijum SA, Mahony J, van Sinderen D. 2013. Complete genome of *Lactococcus lactis* subsp. *cremoris* UC509.9, host for a model lactococcal P335 bacteriophage. Genome Announcement 1:e00119-12.
5. Bolotin A, Wincker P, Mauger S, Jaillon O, Malarme K, Weissenbach J, Ehrlich SD, Sorokin A. 2001. The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. Genome research 11:731-53.
6. Wessels S, Axelsson L, Hansen EB, De Vuyst L, Laulund S, Lähteenmäki L, Lindgren S, Mollet B, Salminen S, von Wright A. 2004. The lactic acid bacteria, the food chain, and their regulation. Trends in food science and technology 15:498-505.
7. Gupta PK. 2008. Single-molecule DNA sequencing technologies for future genomics research. Trends in biotechnology 26:602-611.
8. McCarthy A. 2010. Third generation DNA sequencing: pacific biosciences' single molecule real time technology. Chemistry and biology 17:675-676.
9. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and environmental microbiology: 79 (17) 5112-5120.
10. Braun Jr V, Hertwig S, Neve H, Geis A, Teuber M. 1989. Taxonomic differentiation of bacteriophages of *Lactococcus lactis* by electron microscopy, DNA-DNA hybridization, and protein profiles. Microbiology 135:2551-2560.
11. Korlach J, Turner SW. 2012. Going beyond five bases in DNA sequencing. Current opinion in structural biology 22:251-261.

12. Boye E, Løbner-Olesen A. 1990. The role of dam methyltransferase in the control of DNA replication in *E. coli*. Cell 62:981-989.
13. Julio SM, Heithoff DM, Provenzano D, Klose KE, Sinsheimer RL, Low DA, Mahan MJ. 2001. DNA adenine methylase is essential for viability and plays a role in the pathogenesis of *Yersinia pseudotuberculosis* and *Vibrio cholerae*. Infection and immunity 69:7610-7615.
14. Mohapatra SS, Fioravanti A, Biondi EG. 2014. DNA methylation in *Caulobacter* and other *Alphaproteobacteria* during cell cycle progression. Trends in microbiology 22:528-535.
15. Marinus MG, Casadesus J. 2009. Roles of DNA adenine methylation in host–pathogen interactions: mismatch repair, transcriptional regulation, and more. FEMS microbiology reviews 33:488-503.
16. Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, Bitinaite J, Blumenthal RM, Degtyarev SK, Dryden DT, Dybvig K. 2003. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. Nucleic acids research 31:1805-1812.
17. Ainsworth S, Sadovskaya I, Vinogradov E, Courtin P, Guerardel Y, Mahony J, Grard T, Cambillau C, Chapot-Chartier MP, van Sinderen D. 2014. Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. MBio 5:e00880-14.
18. Østergaard S, Brøndsted L, Vogensen FK. 2001. Identification of a replication protein and repeats essential for DNA replication of the temperate lactococcal bacteriophage TP901-1. Applied and environmental microbiology 67:774-781.
19. Vegge CS, Vogensen FK, Mc Grath S, Neve H, van Sinderen D, Brøndsted L. 2006. Identification of the lower baseplate protein as the antireceptor of the temperate lactococcal bacteriophages TP901-1 and Tuc2009. Journal of bacteriology 188:55-63.
20. McGrath S, Fitzgerald GF, van Sinderen D. 2001. Improvement and optimization of two engineered phage resistance mechanisms in *Lactococcus lactis*. Applied and environmental microbiology 67:608-616.
21. Sambrook J, Russell DW. 2006. Purification of nucleic acids by extraction with phenol: chloroform. Cold spring harbor protocols 2006:pdb. prot4455.

22. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC bioinformatics 11:119.
23. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. Bioinformatics 16:944-5.
24. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A. 2010. The Pfam protein families database. Nucleic acids research 38:D211-22.
25. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA. 2003. The COG database: an updated version includes eukaryotes. BMC bioinformatics 4:41.
26. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic acids research 44:W16-W21.
27. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold spring harbor laboratory press.
28. Landete JM, Arqués JL, Peirotén Á, Langa S, Medina M. 2014. An improved method for the electrotransformation of lactic acid bacteria: a comparative survey. Journal of microbiological methods 105:130-133.
29. Holo H, Nes IF. 1989. High-frequency transformation, by electroporation, of *Lactococcus lactis* subsp. *cremoris* grown with glycine in osmotically stabilized media. Applied and environmental microbiology 55:3119-3123.
30. Garzon AE, Mahony J, Bottacini F, Kelleher P, van Sinderen D. 2019. Complete genome sequence of *Lactococcus lactis* subsp. *cremoris* 3107, host for the model lactococcal P335 bacteriophage TP901-1. Microbiology Resource Announcement 8:e01635-18.
31. Kelleher P, Bottacini F, Mahony J, Kilcawley KN, van Sinderen D. 2017. Comparative and functional genomics of the *Lactococcus lactis* taxon; insights into evolution and niche adaptation. BMC genomics 18:267.

32. Ainsworth S, Stockdale S, Bottacini F, Mahony J, van Sinderen D. 2014. The *Lactococcus lactis* plasmidome: much learnt, yet still lots to discover. FEMS microbiology reviews 38:1066-1088.
33. Kiewiet R, Bron S, de Jonge K, Venema G, Seegers JF. 1993. Theta replication of the lactococcal plasmid pWVO2. Molecular microbiology 10:319-327.
34. Seegers JF, Bron S, Franke CM, Venema G, Kiewiet R. 1994. The majority of lactococcal plasmids carry a highly related replicon. Microbiology 140:1291-1300.
35. Ostergaard Breum S, Neve H, Heller KJ, Vogensen FK. 2007. Temperate phages TP901-1 and phiLC3, belonging to the P335 species, apparently use different pathways for DNA injection in *Lactococcus lactis* subsp. *cremoris* 3107. FEMS microbiology letters 276:156-64.
36. Goh YJ, Goin C, O'Flaherty S, Altermann E, Hutkins R. Specialized adaptation of a lactic acid bacterium to the milk environment: the comparative genomics of *Streptococcus thermophilus* LMD-9, p S22. *In* (ed), BioMed Central,
37. Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N. 2006. Comparative genomics of the lactic acid bacteria. Proceedings of the national academy of sciences of the United States of America 103:15611-15616.
38. van Rooijen RJ, de Vos WM. 1990. Molecular cloning, transcriptional analysis, and nucleotide sequence of lacR, a gene encoding the repressor of the lactose phosphotransferase system of *Lactococcus lactis*. Journal of biological chemistry 265:18499-18503.
39. van Rooijen RJ, Gasson MJ, De Vos WM. 1992. Characterization of the *Lactococcus lactis l*actose operon promoter: contribution of flanking sequences and LacR repressor to promoter activity. Journal of bacteriology 174:2273-2280.
40. Drider D, Bekal S, Prévost H. 2004. Genetic organization and expression of citrate permease in lactic acid bacteria. Genetics and molecular research 3:271-281.
41. Cavanagh D, Fitzgerald GF, McAuliffe O. 2015. From field to fermentation: the origins of *Lactococcus lactis* and its domestication to the dairy environment. Food microbiology 47:45-61.

42. Cords B, McKay L, Guerry P. 1974. Extrachromosomal elements in group N streptococci. Journal of bacteriology 117:1149-1152.
43. Swaisgood H. 2003. Chemistry of the caseins, p 139-201, Advanced dairy chemistry—1 Proteins. Springer.
44. Yu W, Gillies K, Kondo JK, Broadbent JR, McKay LL. 1996. Loss of plasmid-mediated oligopeptide transport system in lactococci: another reason for slow milk coagulation. Plasmid 35:145-155.
45. Steele J, Broadbent J, Kok J. 2013. Perspectives on the contribution of lactic acid bacteria to cheese flavor development. Current opinion in biotechnology 24:135-141.
46. Kvint K, Nachin L, Diez A, Nyström T. 2003. The bacterial universal stress protein: function and regulation. Current opinion in microbiology 6:140-145.
47. Siezen RJ, Renckens B, van Swam I, Peters S, van Kranenburg R, Kleerebezem M, de Vos WM. 2005. Complete sequences of four plasmids of *Lactococcus lactis* subsp. *cremoris* SK11 reveal extensive adaptation to the dairy environment. Applied and environmental microbiology 71:8371-8382.
48. Keto-Timonen R, Hietala N, Palonen E, Hakakorpi A, Lindström M, Korkeala H. 2016. Cold shock proteins: a minireview with special emphasis on Csp-family of enteropathogenic *Yersinia*. Frontiers in microbiology 7:1151.
49. Duwat P, Cesselin B, Sourice S, Gruss A. 2000. *Lactococcus lactis*, a bacterial model for stress responses and survival. International journal of food microbiology 55:83-86.
50. Seegers JF, van Sinderen D, Fitzgerald GF. 2000. Molecular characterization of the lactococcal plasmid pCIS3: natural stacking of specificity subunits of a type I restriction/modification system in a single lactococcal strain. Microbiology 146:435-443.
51. Mills S, Coffey A, Hill C, Fitzgerald GF, McAuliffe O, Ross RP. 2005. Insertional inactivation of determinants for Mg^{2+} and Co^{2+} transport as a tool for screening recombinant *Lactococcus* species clones. Applied and environmental microbiology 71:4897-4901.
52. Mattheus W, Botteldoorn N, Heylen K, Pochet B, Dierick K. 2012. Trend analysis of antimicrobial resistance in *Campylobacter jejuni* and *Campylobacter coli* isolated from Belgian pork and poultry meat products

using surveillance data of 2004–2009. Foodborne pathogens and disease 9:465-472.

53. Amos G, Zhang L, Hawkey P, Gaze W, Wellington E. 2014. Functional metagenomic analysis reveals rivers are a reservoir for diverse antibiotic resistance genes. Veterinary microbiology 171:441-447.

54. Flórez AB, Ammor MS, Mayo B. 2008. Identification of tet (M) in two *Lactococcus lactis* strains isolated from a Spanish traditional starter-free cheese made of raw milk and conjugative transfer of tetracycline resistance to lactococci and enterococci. International journal of food microbiology 121:189-194.

55. Fallico V, McAuliffe O, Fitzgerald GF, Ross RP. 2011. Plasmids of raw milk cheese isolate *L. lactis* subsp. *lactis biovar*. diacetylactis DPC3901 suggest a plant-based origin for the strain. Applied and environmental microbiology:AEM. 00661-11.

56. Mills S, McAuliffe OE, Coffey A, Fitzgerald GF, Ross RP. 2005. Plasmids of lactococci–genetic accessories or genetic necessities? FEMS microbiology reviews 30:243-273.

57. O'Driscoll J, Glynn F, Fitzgerald GF, van Sinderen D. 2006. Sequence analysis of the lactococcal plasmid pNP40: a mobile replicon for coping with environmental hazards. Journal of bacteriology 188:6629-6639.

58. Grohmann E, Muth G, Espinosa M. 2003. Conjugative plasmid transfer in Gram-positive bacteria. Microbiology and molecular biology reviews 67:277-301.

59. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, Korlach J, Turner SW. 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. Nature methods 7:461.

Chapter III

Lactococcal bacteriophage TP901-1 requires the host-encoded GtfA protein, a putative glycosyltransferase, for DNA injection.

Note: Bioinformatic analysis performed by Dr. Francesca Bottacini.

Chapter III contents

3 Abstract

Three *L. lactis* subsp. *cremoris* 3107 derivatives, known as E119, E121 and E126, were previously shown to be resistant against phage TP901-1, while remaining sensitive to infection by another P335 group phage, LC3. It was suggested that the observed resistance to TP901-1 was due to an impairment in a step of the phage cycle beyond the initial recognition of and adsorption to the host cell. The genomes of the parent strain and the three TP901-1-resistant mutants were sequenced and compared, resulting in the identification of several mutations that may underpin the observed TP901-1 resistance. Subsequent complementation assays showed that introduction of a plasmid harbouring the *L3107_1442* gene (renamed here as *gtfA*), encoding a predicted glycosyltransferase, in mutant strains E119 and E126 restored phage sensitivity. Furthermore, the lysogenization frequency increased about 100-fold, indicating significantly improved DNA injection efficiency. Moreover, silencing of *gtfA* by an anti-sense mRNA interference strategy was shown to interfere with infection by phage TP901-1 in *L. lactis* 3107. It was furthermore experimentally validated that *gtfA* is involved in one of the initial steps of the infection process of the lactococcal P335 phage TP901-1.

3.1 Introduction

Lactic acid bacteria (LAB) represent a group of Gram-positive, anaerobic, non-spore forming bacteria. Some component species are used extensively in dairy fermentations as starter or adjunct cultures due to their acidification activity as well as their contribution to flavour and texture characteristics of the final product (1). Among the LAB, strains of the industrially significant species *Lactococcus lactis* are susceptible to infection by bacteriophages (phages), which are ubiquitous in the dairy environment. These phages may originate from the incoming raw milk and may not be fully inactivated by pasteurization treatments (2). Phage susceptibility of starter and adjunct strains is a major concern for the dairy fermentation industry as phage attacks may negatively impact on the final product quality and production regimes, with concomitant economic losses (3). Therefore, a thorough understanding of how these problematic phages interact with their hosts is necessary in order to devise rational strategies that prevent or at least limit their negative impact. Such strategies may, for example, involve the construction of robust starter cultures or the development of knowledge-based phage-incompatible strain rotations/blends (3-5).

The International Committee on Taxonomy of Viruses has classified lactococcal phages among the *Caudovirales* order of tailed phages, which includes over 95 % of all known phages (6). The *Caudovirales* order is comprised of three families: the *Myoviridae* (viruses with long, contractile tails), *Siphoviridae* (viruses with long, non-contractile tails) and *Podoviridae* (viruses with short tails). To date, morphological and genomic analyses have classified lactococcal phages into 10 genetically distinct groups, of which eight belong to the *Siphoviridae* family and two to the *Podoviridae* family (2). Members of the P335, 936 and c2 phage groups of lactococcal phages (all *Siphoviridae*) are most frequently and globally encountered in dairy fermentations (7). Currently, the complete genome sequences of twenty seven P335 phages are available (excluding predicted prophages identified on sequenced *L. lactis* genomes), and among these phages are the 'model' P335 phages Tuc2009 and TP901-1, as well as LC3, BK5-T, r1t, ul36, 4268, Q33 and BM13 (8-18). Tuc2009 and TP901-1 have received significant research attention over the past decade and have become prototypes for the analysis of phage-host interactions in Gram-positive bacteria (19-24). The P335 phage group has been further divided into five subgroups (subgroups I through to V) based on their distinct genetic lineages and morphological features (14, 25).

The first step in lactococcal phage infection is the recognition of, and adsorption to, a surface-exposed, host-encoded receptor followed by DNA release from the phage particle, resulting in the translocation of the phage genome into the bacterial cytoplasm (26, 27). The first phage-host interaction event is dependent on two factors: (i) a host-encoded receptor molecule(s), such as a cell wall polysaccharide (CWPS), a surface-exposed protein and/or teichoic acid (TA) (26), and (ii) the corresponding phage-encoded receptor binding protein(s) (RBPs). A membrane-bound proteinaceous receptor has been identified for certain Gram-positive infecting phages. For example, the model *Bacillus subtilis* siphophage SPP1 irreversibly binds to the membrane-associated YueB protein (28). Additionally, the purification of the extracellular region (ectodomain) of YueB was shown to trigger SPP1 DNA ejection *in vitro* (29). A membrane-associated protein, PIP (phage infection protein), is also required for bacteriophage c2 infection of *L. lactis* subsp. *lactis* C2 (30). Moreover, seven c2-like lactococcal phages, including c2, were able to reversibly bind to an *L. lactis* carbohydrate receptor moiety of unknown composition and then to irreversibly bind to PIP, causing the ejection of the phage genome (31). Furthermore, the primary receptors for members of the lactococcal P335 (32, 33), 936 (32, 34, 35), 1358 (36), 949 (37) and P087 (37) phage groups are saccharidic in nature. More specifically, the variable, so-called PSP component of the CWPS of *L. lactis* strains has been identified as the phage receptor in a number of cases (32-34, 36-40). In the case of *L. lactis* 3107 this PSP component was shown to act as the host cell surface receptor for the P335 group phages TP901-1 (sub-group II) and LC3 (sub-group III) (32).

Chemical mutagenesis of *L. lactis* subsp. *cremoris* 3107, which as mentioned above is sensitive to TP901-1 and LC3, culminated in the isolation of three distinct mutants, named E119, E121 and E126, which are resistant to phage TP901-1, while still susceptible to infection by phage LC3 (41). Subsequent analysis of these mutants showed that a step involving the TP901-1 DNA injection process into the host is impaired, suggesting that temperate phages TP901-1 and LC3 either utilize different pathways of injection or are differently triggered by their receptor (41). In the current study, the genomes of *L. lactis* 3107 and its derivatives were sequenced in order to identify gene(s) that underpin phage-host interactions between TP901-1 and its host. This resulted in the identification of *L3107_1442*, designated here as *gtfA*, which is predicted to encode a glycosyltransferase, and which is involved in one of the early steps of TP901-1 infection (42).

3.2 Materials and Methods

3.2.1 Bacterial strains, growth conditions, phage preparations, and plasmids.

Strains, phages and plasmids employed in this study are listed in Table 3.1. *L. lactis* strains were grown at 30 °C in M17 broth/agar (M17 (Oxoid, U.K.) supplemented with 0.5 % (v/v) glucose (GM17) and 5 μg ml^{-1} chloramphenicol (Cm) where necessary. Lactococcal phages were propagated in GM17 broth cultured with the host strain at an approximate OD_{600nm} of 0.2 and supplemented with 10 mM $CaCl_2$. TP901-1*erm* was induced from *L. lactis* subsp. *cremoris* NZ9000_TP901-1*erm* using 0.5 μg ml^{-1} mitomycin C when the growing culture reached an OD_{600nm} of approximately 0.2 (43). SM buffer (10 mM $CaCl_2$, 100 mM NaCl, 10 mM $MgSO_4$, 50 mM Tris-HCl at pH 7.5) was used as the diluent in all phage assays.

Table 3.1 Bacteria, phages, and plasmids used in this study.

	Description/purpose	Source
Bacterial Strains		
L. lactis subsp. *cremoris*		
3107	Host strain for the P335 phages TP901-1 and LC3	(44)
E119	Phage TP901-1-resistant derivative of 3107	(41)
E121	Phage TP901-1-resistant derivative of 3107	(41)
E126	Phage TP901-1-resistant derivative of 3107	(41)
NZ9000	MG1363 derivative in which the *nisRK* genes have been inserted in the chromosome	(45)
NZ9000Δ*8915*	NZ9000 harbouring a non-sense mutation in *LLNZ_08915*, homologue of $L3107\_1441_{3107}$	This study
NZ9000Δ*gtfA*	NZ9000 harbouring a non-sense mutation in *LLNZ_08920*, homologue of $gtfA_{3107}$	This study
NZ9000Δ*8915_ gtfA*	NZ9000 harbouring a non-sense mutation in *LLNZ_08915* and *LLNZ_08920*, homologues of $L3107\_1441_{3107}$ and $gtfA_{3107}$, respectively.	This study
NZ9000_TP901-1*erm*	NZ9000 harbouring TP901-1 containing an erythromycin (Em) marker induced from *L. lactis* subsp. *cremoris* 901-1	(43)
L. lactis subsp. *lactis* IL1403	Source of *yieH* which bears 84 % identity to $gtfA_{3107}$	(46)
Escherichia coli EC101	Cloning host	(47)
Bacteriophages		
TP901-1*erm*	Temperate P335 group phage infecting 3107, contains Em^r marker	(43)
TP901-1	Temperate P335 group phage induced from *L. lactis* subsp. *cremoris* 901-1	(48)
LC3	Temperate P335 group phage grown lytically on *L. lactis* subsp. *cremoris* 3107	(49)
Dub35A	P335 phage group	(25)
63301	P335 phage group	(25)
58601	P335 phage group	(25)

86501	P335 phage group	(25)
50101	P335 phage group	(25)
07501	P335 phage group	(50)
62605	936 phage group	(50)
62604	936 phage group	(50)
62601	936 phage group	(50)
66901	936 phage group	(50)
66902	936 phage group	(50)
66903	936 phage group	(50)
jj50	936 phage group	(51)
p2	936 phage group	(52)
sk1	936 phage group	(53)
712	936 phage group	(51)
c2	c2 phage group	(30)
P087	P087 phage group	(54)
949	949 phage group	(55)
WRP3	949 phage group	(37)
L47	949 phage group	(56)
LW81	949 phage group	(57)
Plasmids		
pJP005	Derivative of pNZ8048, carries *recT1* under control of nisin-inducible promoter P_{nisA}; Cm^r	(58)
pNZ44	*Escherichia coli–L. lactis* vector containing P44 promoter; Cm^r	(59)
pNZ*gtfA*	Complementing plasmid; pNZ44 vector harbouring *gtfA* from *L. lactis* 3107	This study
pNZS*gt41*	Silencing plasmid; pNZ44 vector harbouring *gtfA* from *L. lactis* 3107 in the anti-sense orientation	This study
pNZ*1441*	pNZ44 vector harbouring *L3107_1441* from *L. lactis* 3107	This study
pNZ*1441_ gtfA*	pNZ44 vector harbouring *L3107_1441* and *gtfA* from *L. lactis* 3107	This study
pNZ*yieH*	pNZ44 vector harbouring *yieH* from *L. lactis* IL1403	This study

3.2.2 Genome sequencing and general feature predictions of *L. lactis* 3107 and its derivatives E119, E121 and E126

The *L. lactis* subsp. *cremoris* 3107 genome was sequenced (18) using the sequencing technologies Single Molecule Real Time (SMRT) (60, 61) performed by GATC Biotech Ltd. (Germany) and Illumina (62) performed by Macrogen Inc. (South Korea) as described in Chapter II of this thesis (Sections 2.2.2). Following sequencing, the genomic features of 3107 were predicted as detailed in Chapter II of this thesis (Section 2.2.3). Similarly, the *L. lactis* 3107 derivatives E119, E121 and E126 were sequenced and their genomic features were predicted as described in Chapter II (Sections 2.2.2 and 2.2.3).

3.2.3 Bioinformatic analysis of *L. lactis* 3107 and its derivatives E119, E121 and E126

Single nucleotide polymorphism (SNP) analysis was performed between the genome sequence of *L. lactis* 3107 and the genome sequence of its derivatives to identify mutations that may have caused the observed phage-resistance. For this purpose, SNP analysis was performed using the *L. lactis* 3107 consensus sequence as reference for Bowtie2 alignment (63), followed by the use of SAMtools (64) to extract base variants. The sequence validity of randomly selected SNPs was confirmed by PCR amplification and Sanger sequencing of the relevant genomic regions.

3.2.4 Bioinformatic analysis of mutated *L. lactis* 3107 genes

In silico analysis of the potential *L. lactis* 3107 genes involved in TP901-1 phage resistance, as identified by SNP analysis, was performed using the online bioinformatics software TMHMM v. 2.0 (65) and HHpred (66) to identify predicted transmembrane regions in the encoded proteins, and to predict protein function, respectively. The identified *gtfA* gene was further analysed using Artemis software (67) to assess its genomic context. Moreover, the presence and conservation of the *gtfA* gene in other lactococcal strains was analysed using BlastP (68).

3.2.5 DNA manipulation and construction of plasmids

3.2.5.1 Construction of potential complementing plasmids

Primers (listed in Table S3.1) were designed and generated to amplify targeted *L. lactis* subsp. *cremoris* 3107 genes. The amplicons and the chloramphenicol-resistant (Cm^r) plasmid pNZ44 (3,396 bp) were restricted using the appropriate restriction enzymes (as indicated in Table S3.1). Restriction enzyme digestions, DNA ligations and agarose gel electrophoresis were performed following standard procedures (69). Ligated constructs were introduced into *E. coli* EC101 competent cells prepared using a previously described protocol (47). *E. coli* EC101 transformants expected to harbour the recombinant vector were isolated using LB agar plates (47) supplemented with 5 µg ml^{-1} chloramphenicol (Cm). Recombinant plasmids were then identified by the presence of a cloned DNA fragment of the expected size following PCR with the pNZ44_F (Table 3.2) primer and the corresponding reverse primer, which was specific for the cloned DNA fragment, e.g. *gtfA*_R (Table 3.2). Sequence integrity of the generated constructs was determined by Sanger sequencing the plasmid insert (performed by MWG Eurofins Genomics, Ebersberg, Germany) and subsequent analysis using the Seqman Pro program (DNAstar v7.0) (70). Verified recombinant plasmids were individually isolated using a Pure Link DNA Purification (Thermofisher Scientific) kit and introduced into competent cells of *L. lactis* mutants E119, E121 and E126. To confirm the presence of the desired constructs, PCR reactions using construct-specific primers (Tables S3.1 and 3.2) were performed on the Cm^r isolates.

3.2.5.2 Transduction assays

Transduction assays were performed in cases where particular plasmids could not be successfully introduced into *L. lactis* E119, E121 or E126 by electrotransformation. Such plasmids were first isolated from *E. coli* EC101 cells using a Pure Link DNA Purification (Thermofisher Scientific) kit and introduced into *L. lactis* 3107 by electrotransformation. Subsequently, LC3 was propagated on an *L. lactis* 3107 derivative harbouring such a plasmid. The lysate generated was then added to 50 mL of GM17 broth cultured with the desired host strain at an OD_{600nm} of 0.2 and supplemented with 10 mM $CaCl_2$. Every 30 minutes, 10 mL of this lysate-host mix was centrifuged at 1,900 × *g* for 10 minutes at 4 °C; the supernatant was discarded and the pellet was resuspended in 100 µl of GM17 and spread plated onto GM17 plates supplemented with 5 µg ml^{-1} Cm. The plates were then incubated at 30 °C for 48 h. PCR assessment using construct-specific primers (Tables S3.1 and 3.2) was employed

on resulting Cm^r colonies to confirm the presence of the desired construct. Sequence integrity of the introduced plasmid was verified by sequencing plasmid DNA using MWG Eurofins Genomics (Ebersberg, Germany) followed by sequence analysis employing the Seqman Pro program (DNAstar v7.0) (70).

3.2.5.3 Construction of the silencing plasmid pNZS*gtfA*

The *L3107_1442* (2,676 bp) gene, renamed here as *gtfA*, from *L. lactis* 3107 was cloned in the antisense orientation into the pNZ44 vector by first amplifying the gene using primers S*gtfA*_F and S*gtfA*_R (Table 3.2). The forward primer (S*gtfA*_F) incorporated the restriction site KpnI as well as the native ribosome binding site (RBS) AGGAGG, whereas the reverse primer (S*gtfA*_R) incorporated the restriction site PstI (Table 3.2). The amplicon was digested with KpnI and PstI, ligated with similarly digested pNZ44, and introduced into *E. coli* EC101 by electroporation. The desired plasmid, designated pNZS*gtfA*, was verified by sequencing and then introduced into *L. lactis* 3107 and *L. lactis* NZ9000.

3.2.5.4 Construction of pNZ*yieH*, pNZ*1441* and pNZ*1441_gtfA*

The *yieH* (2,676 bp) gene from *L. lactis* IL1403, the *L3107_1441* (786 bp) gene, and a DNA fragment encompassing the *L3107_1441* and *gtfA* (3,465 bp in total) genes from *L. lactis* 3107 were amplified using appropriate primers (listed in Table 3.2). The resulting amplicons were restricted with appropriate enzymes, ligated into similarly restricted pNZ44, and following this the ligation mixture was introduced into *E. coli* EC101 by electroporation. The desired recombinant plasmids, designated pNZ*yieH* (6,070 bp), pNZ*1441* (4,180 bp) and pNZ*1441_gtfA* (6,859 bp) were verified by sequencing. Following this, pNZ*yieH* was introduced into *L. lactis* E119 and E126, while pNZ*1441* and pNZ*1441_gtfA* were individually introduced into *L. lactis* NZ9000.

Table 3.2: Primers used in this study; introduced restriction sites are underlined.

Name	Sequence (5' – 3')	Target/purpose	Source
pNZ44_F	CTAATGTCACTAACCTGCCCCG	p44 lactococcal promoter and multiple cloning site (MCS) of pNZ44	Chapter II
pNZ44_R	GCTTTATCAACTGCTGCT		
*gtfA*_F	AGCAGCCTGCAGAGGAGGCACTCACTTGTATCTTTGGTATAATAG	*L. lactis* 3107 *gtfA* gene	This study
*gtfA*_R	AGCAGCGGTACCTTAGCTTTCATTTTGCTGAC		
S*gtfA*_F	AGCAGCGGTACCAGGAGGCACTCACTTGTATCTTTGGTATAATAG	*L. lactis* 3107 *gtfA* gene in the antisense orientation	This study
S*gtfA*_R	AGCAGCCTGCAGTTAGCTTTCATTTTGCTGAC		
*1441*_F	AGCAGCCTGCAGAGGAGGCACTCACATGATTAAACATAGTGAAAAT	*L. lactis* 3107 *L3107_1441* gene	This study
*1441*_R	AGCAGCTCTAGATCAGTATTTTTTATTTAATCG		
IL1403_*yieH_F*	AGCAGCCTGCAGAGGAGGCACTCACTTGTATCTTTGGTATAATAG	*L. lactis* IL1403 *yieH* gene	This study
IL1403_*yieH_R*	AGCAGCTCTAGATTAGCTTTCATTTT		
*1441*check_F	GACACCAATGACATCATAGCCTT	1053 bp of *L. lactis* 3107 *L3107_1441* gene	This study
*1441*check_R	AGATTATAAGAATCACAGCGACCAT		

*1441*mama_F	GTCCTTTTTGAATTCCCTAAT	Mutated *L. lactis* 3107 *L3107_1441* gene	This study
*gtfA*check_F	CAGGTTGTTGGCTTGGAA	984 bp of *L. lactis* 3107 *L3107_gtfA* gene	This study
*gtfA*check_R	TCATAGGTCATCCGCGTC		
*gtfA*mama_F	GATTTGAATTCCCTAACTTGG	Mutated *L. lactis* 3107 *L3107_gtfA* gene	This study
*1443*check_F	CGGCCCTCATTCTGGTGG	*L. lactis* 3107 *L3107_1442* gene	This study
*1443*checkR	AAATTGTACCGGACTTGC		
*1443*mamaF	TGTCTGAATTCCCTAACT	Mutated *L. lactis* 3107 *L3107_1441* gene	This study

3.2.6 Mutant generation by means of recombineering mutagenesis

Recombineering mutagenesis was performed as described previously (58, 71, 72) with some minor adjustments. Briefly, recombineering oligonucleotides designed to mutate *L. lactis* NZ9000 genes *LLNZ_08915*, *LLNZ_08920* (*gtfA*) and *LLNZ_08925* are listed in Table 3.3. These oligonucleotides are approximately 85-bases long and are partly identical to the lagging strand of the targeted gene with nucleotides introduced within the coding sequence of each targeted gene to incorporate an EcoRI (GAATTC) site immediately downstream of a "T" to create the stop codon TGA, and then followed by the incorporation of additional nucleotides to create a second TAA stop codon (see Table 3.3). Additionally, each mutational oligo contained 5' phosphorothioate modifications to prevent endonuclease and exonuclease degradation of the oligo following introduction into lactococcal cells. Cells were allowed to recover for 1 hour after electroporation prior to spread plating the relevant dilutions on GM17 agar. Screening for mutants was performed by PCR amplification using primers (Table 3.2) incorporating the overlap region spanning the mutated segment. Colonies harbouring the expected mutation were purified to ensure a pure genotype. PCR amplification was performed on the colonies using primers (Table 3.2) targeting the DNA region spanning the mutated segment. The expected mutated sequence of the amplified region of each mutant phage in the genome of *L. lactis* NZ9000 was confirmed by Sanger sequencing using MWG Eurofins Genomics (Ebersberg, Germany) and analysing the results using the Seqman Pro program (DNAstar v7.0) (70).

Table 3.3: Recombineering oligonucleotides designed to introduce mutations in specific *L. lactis* subsp. *cremoris* NZ9000 genes; introduced mutations are underlined.

Mutant strain	ssDNA Oligonucleotide sequence (5' – 3')	Mutation	Source
Lactococcus lactis subsp. *cremoris*			
NZ9000Δ*8915*	A*A*C*T*G*TTGTACCCAAAGAAAAAATTTGTCCTTTTTGAATTCCCTAATGTGATGTC GAACACCTGACCGGTATATTAAAAACTTCCG	*LLNZ_08915*$_{NZ9000}$	This study
NZ9000Δ*gtfA*	A*A*T*C*T*TAATGGTTATCGTTCTTGCAATGACTGGGATTTGAATTCCCTAACTTGGG GAAGTTCACGGTCAATTTTGGCTGGAGATGC	*LLNZ_08920*$_{NZ9000}$ / *gtfA*$_{NZ9000}$	This study
NZ9000Δ*8925*	C*T*T*G*G*AATTCTAGTAGCGGTAAGCATTCCTGTCTGAATTCCCTAATCTTATTCTCT AAAAACAAAACATTTTGAAGTCAATTTCAA	*LLNZ_08925*$_{NZ9000}$	This study

Nucleotides introduced within the coding sequence of each targeted gene are underlined, the incorporation of an EcoRI (GAATTC) site after a "T" created the stop codon TGA followed by the incorporation of additional mismatches to create the stop codon TAA. The asterisks represent 5' phosphorothioate modifications in the mutational primers.

3.2.7 Electrotransformation of *L. lactis* cells

L. lactis 3107 competent cells were prepared and used for electrotransformation as described in Chapter II (Section 2.2.6). Similarly, electrotransformation of *L. lactis* 3107 derivative strains E119, E121 and E126 was performed as previously described (Chapter II, Section 2.2.6) with the exception that these mutant strains were inoculated in 50 mL of GM17 broth supplemented with 0.6 – 0.8 % glycine. *L. lactis* NZ9000 competent cells were prepared and used for electrotransformation as previously described (73).

3.2.8 Phage assays

3.2.8.1 Efficiency of plaquing (E.O.P.)

To determine the E.O.P. of a particular phage on a specific host, solid and semi-solid agar was prepared using GM17 broth supplemented with 1.5 % (w/v) or 0.7 % (w/v) bacteriological agar, respectively, plus 10 mM $CaCl_2$, and 0.5 % glycine (w/v) where necessary (as E119, E121 and E126 strains are intolerant to glycine), and employed according to a previously described method (74). Lawns of the corresponding indicator strain were plated with progressive 10-fold phage dilutions that had been prepared in SM buffer. The E.O.P. was determined as the ratio of the titre obtained from the test strain to that of the control strain (*L. lactis* 3107).

3.2.8.2 Lysogenization assays

These assays were performed as previously described (41) with the following modifications. Briefly, 10 mL GM17 broth was inoculated with 2 % of the required strain until an approximate OD_{600nm} of 0.2 was reached. 500 µl of the culture was then infected with 500 µl of TP901-1*erm* at a multiplicity of infection (MOI) of 0.5 and the mixture was incubated at 30 °C for 1 hour to allow one round of infection to occur. Following incubation, the cells were diluted as necessary and plated on GM17 agar supplemented with 1 µg ml^{-1} Em. Lysogens were observed following anaerobic incubation for 48 hours at 30 °C. The frequency of lysogenization was determined by calculating the total number of obtained erythromycin resistant (Em^r) lysogens per millilitre, divided by the total colony forming units per millilitre (CFU ml^{-1}) obtained in the absence of Em.

3.2.8.3 Reversible and irreversible adsorption assays

Reversible and irreversible adsorption assays were performed as previously described (41) with the following modifications. Briefly, 10 mL GM17 broth was inoculated with 2 % of the appropriate strain until an OD_{600nm} of approximately 0.5 was reached. The culture was supplemented with 10 mM $CaCl_2$, following this 700 µl of the mixture was added to an equal volume of phage lysate (at a multiplicity of infection [MOI] of 0.01), and the mixture was incubated at 30 °C for 10 minutes. As a control 700 µl GM17 broth supplemented with 10 mM $CaCl_2$ was mixed with 700 µl of the phage lysate and incubated at 30 °C for 10 minutes. Following incubation, irreversible adsorption was determined by diluting samples 1:100 in ice-cold quarter-strength Ringer's solution supplemented with 1 M NaCl. Cell-free supernatants were prepared by centrifugation of the samples at 6,000 × *g* for 1 minute and phage numbers in the supernatants were determined by standard plaque assays using *L. lactis* 3107 as the indicator strain. Following incubation, reversible adsorption was also determined, samples were centrifuged at 6,000 × *g* for 1 minute to pellet cells and the supernatant was diluted 1:100 in quarter strength Ringer's solution supplemented with 1 M NaCl. Phage numbers in the supernatant were determined by standard plaque assays using *L. lactis* 3107 as the host strain. Adsorption was measured as percentage of total phages attached to the host when compared to control phage titre (which lacks bacterial cells) using the following formula: [(Control phage titre – Free phage titre in supernatant) / Control phage titre] x 100.

3.2.8.4 Lysis-in-broth assays

50 mL of GM17 broth was inoculated with the desired (host) strain and incubated at 30°C until an OD_{600nm} of approximately 0.2 was reached. The culture was then supplemented with 10 mM $CaCl_2$ and infected (by adding 10^8 pfu ml^{-1} phages) of the desired lactococcal phage and incubated at 30 °C. Following this, every 30 minutes, OD_{600nm} readings of the infected culture were performed for four and a half hours.

3.2.9 Generation of TP901-1*erm* escape mutants

In order to isolate phages that have gained the ability to infect the TP901-1-resistant *L. lactis* 3107-derived strains E119, E121 and E126 the following experiment was performed: 10 mL of GM17 broth was inoculated with 2 % fresh overnight culture

of one of the three afore-mentioned strains, i.e. E119, E121 or E126, and incubated at 30 °C until an OD_{600nm} of approximately 0.2 was reached. The culture was then supplemented with 10 mM $CaCl_2$ and 1 mL of a TP901-1*erm* lysate containing 10^8 pfu ml^{-1}, and incubated at room temperature overnight. The resulting mixture was then filtered using a MF-Millipore™ Membrane Filter of 0.45 µm pore size (Merck, Ireland) to remove bacterial cells and obtain a cell free supernatant (CFSN) that had the potential to contain TP901-1*erm* escape mutants. E.O.P. assays were then performed using this CFSN to allow the detection of TP901-1*erm* escape mutants (if present) that were capable of infecting the strain with which they were co-incubated with. In the absence of visible plaque formation, the generated CFSN was exposed repeatedly to the same TP901-1-resistant strain until TP901-1*erm* escape mutants were isolated by screening the CFSN produced in each round. This process was repeated twice using each TP901-1-resistant strain (i.e. E119, E121 or E126), and in total six TP901-1*erm* escape mutants were isolated.

3.2.10 Phage DNA extraction, genome sequencing, assembly and bioinformatic analysis

Phage genomic DNA was extracted from TP901-1*erm* escape mutants and TP901-1*erm* WT using a Phage DNA Isolation Kit (Norgen Biotek, Canada). Phage genomic DNA sequencing was performed by the commercial sequencing service provider Probiogenomics Lab (Parco Area de la Scienze, Parma, Italy) which uses Illumina sequencing based on a MiSeq platform. Subsequently, comparative genome and SNP analysis was performed between the phage genome sequence of TP901-1*erm* escape mutants and that of the sequenced TP901-1*erm* WT to identify potential mutations that may have caused these mutant phages to overcome the TP901-1 phage-resistance of 3107 mutant strains E119, E121 and E126. Identified SNPs were confirmed using PCR and Sanger sequencing. Furthermore, in order to verify the integrity of the mutations identified in the genomes on the TP901-1*erm* escape mutants, a comparative and SNP analysis was performed using the data of the newly sequenced TP901-1*erm* genome and the published TP901-1 genome (9), thus discarding any mutations observed as potential causes for overcoming phage resistance. *In silico* analysis of the potential TP901-1 genes involved in overcoming phage resistance was performed using the online bioinformatic software HHpred (66) to predict protein function.

3.3 Results

3.3.1 Genome sequencing of *L. lactis* 3107 and its derivatives

Chemical mutagenesis of *L. lactis* subsp. *cremoris* 3107, a strain sensitive to P335 group phages TP901-1 and LC3, had previously resulted in the isolation of three distinct TP901-1-resistant mutants (41). Adsorption of TP901-1 to these mutants was unaffected suggesting that a step beyond initial host recognition and attachment is impaired in these mutants (41). The genomes of *L. lactis* 3107 and its three derivatives, E119, E121 and E126, were sequenced in order to identify the gene(s) that is (are) responsible for the observed phage resistance phenotype. The genome of *L. lactis* 3107 (wild type) is 2.38 Mbp in length with 2,491 predicted protein-coding regions, while it also harbours six plasmids ranging in size from 2.2 to 60.2 kb (18). The genome of *L. lactis* 3107 has a 35.82 % G+C content, which is in keeping with those of other lactococcal strains (75). As expected, the genomes of the derivatives exhibit near identical characteristics to that of the parent strain and carry an identical plasmid content. Comparative genomic analysis did not reveal deletion or insertion events, and therefore a SNP analysis was undertaken to identify point mutations that may have caused the observed phage-resistance phenotype. This resulted in the identification of 36, 15 and 21 SNPs on the genomes of E119, E121 and E126, respectively, and the detailed results of this analysis are presented below.

3.3.2 Bioinformatic analysis and construction of potential complementing plasmids

Sequence differences were identified following comparative genome and SNP analyses of *L. lactis* 3107 and its three derivatives, E119, E121 and E126. In total, 72 SNPs (Table S3.1) were identified with seven of these SNPs occurring in more than one of the *L. lactis* 3107 mutant genomes, and with some genes containing more than one mutation in a given genome. Previous studies have demonstrated that Gram-positive-infecting phages utilize membrane-bound receptors for their infection process (28, 30, 31). Therefore, each of the proteins encoded by *L. lactis* 3107 genes that were carrying SNPs in E119, E121 and/or E126, and thus potentially involved in TP901-1 phage resistance, was analysed using the online bioinformatics software TMHMM v. 2.0 (65) to identify predicted transmembrane regions, and employing HHpred to investigate its homology and predicted structure (66) (Table S3.1). A total of 31, 15

and 17 genes were mutated in the genomes of strains E119, E121 and E126, respectively, with 5 of these genes mutated in both E119 and E126. In order to identify the mutated gene(s) responsible for the TP901-1-resistant phenotype of the three *L. lactis* 3107 derivatives, pNZ44 derivatives were constructed each carrying a gene, which had been amplified from *L. lactis* 3107 WT, and which had been shown to be mutated in E119, E121 and/or E126, and predicted to encode a membrane-associated protein (Table S3.1). Out of the 58 identified mutated genes, a total of 33 different recombinant plasmids (each carrying an *L. lactis* 3107 gene) were constructed prior to identifying a gene involved in TP901-1 phage infection. 14, 15 and 5 individual recombinant plasmids were introduced into strains E119, E121 and E126, respectively, using electroporation or transduction utilizing LC3 (Table S3.1). Following the verification of the presence of each construct in the corresponding mutant strain(s) by PCR, the resulting 34 strains were subjected to efficiency of plaquing and lysogenization assays to assess if the introduction of the recombinant plasmid (containing a gene that had been shown to be mutated in such a mutant) resulted in the restoration of a TP901-1-sensitive phenotype. One recombinant plasmid, pNZ*gtfA*, appeared to restore the phage sensitive phenotype for two of the three *L. lactis* 3107 mutants, E119 and E126. None of the generated recombinant plasmids restored TP901-1 sensitivity of mutant E121, which does not harbour a *gtfA* mutation, thereby indicating that this mutant is affected in a gene that is different from that in mutants E119 and E126, and showing the involvement of a secondary factor for successful TP901-1 infection.

3.3.3 Confirmation of restoration of phage-sensitivity phenotype through efficiency of plaquing (E.O.P.) assays

In order to verify if, and to what extent, phage sensitivity was restored in *L. lactis* E119 and E126 in the presence of pNZ*gtfA*, quantitative plaque assays were performed. The E.O.P. of TP901-1 was determined on *L. lactis* 3107*,* as well as the E119 and E126 mutants harbouring pNZ*gtfA* or the empty vector pNZ44 as a control. As expected, TP901-1 was not able to form visible plaques on E119 and E126, nor on the mutant strains harbouring pNZ44, yielding an E.O.P. of $\leq 10^{-8}$. However, in the presence of plasmid pNZ*gtfA*, TP901-1 was able to form plaques on these two mutants with the same E.O.P. as on *L. lactis* 3107 (Table 3.4). To further confirm that the presence of the pNZ*gtfA* construct caused the reversion of the TP901-1 phage

resistance phenotype, the mutants harbouring pNZ*gtfA* were passaged several times in the absence of Cm and then plated in order to cure the complemented mutant strains of pNZ*gtfA*. The absence of the pNZ*gtfA* plasmid was confirmed by the inability of the cured E119 and E126 mutants to grow in the presence of Cm (while also being negative for the presence of this plasmid as determined by PCR). As anticipated, TP901-1 was incapable of forming visible plaques on any of the obtained pNZ*gtfA*-free mutant strains, exhibiting an E.O.P. of $\leq 10^{-8}$. Phage LC3 formed visible plaques on all strains, including those harbouring the complementing plasmid, with the same efficiency as on wild type *L. lactis* 3107 (Table 3.4).

3.3.4 Lysogenization assays

Lysogenization assays were performed to support the results of the infection assays outlined in the previous paragraph. To this end TP901-1*erm*, a derivative of phage TP901-1 containing an erythromycin-resistant (Em^r) marker (43), was employed to determine the frequency of lysogenization by establishing the number of erythromycin-resistant colonies obtained per infecting phage (see Materials and Methods). As expected, the lysogenization frequencies of the two *L. lactis* 3107 mutant strains, E119 and E126, were lower than that of the WT strain (10^{-3} vs. 10^{-6}, Table 3.4). Interestingly, the approximately 100-fold increase in the frequency of lysogeny observed for E119 and E126 derivatives harbouring pNZ*gtfA* ($3.03 \times 10^{-4} \pm 4.93 \times 10^{-6}$ and $4.17 \times 10^{-4} \pm 3.58 \times 10^{-4}$, respectively) as compared to E119 and E126 ($5.54 \times 10^{-6} \pm 2.93 \times 10^{-7}$ and $5.39 \times 10^{-6} \pm 2.11 \times 10^{-7}$, respectively) represents a statistically significant increase (p-values of <0.05). This shows that the efficiency of infection and associated lysogenization frequency of TP901-1*erm* significantly improved in the presence of pNZ*gtfA*, though the latter was not fully restored to the level of the WT strain 3107. The cured E119 and E126 mutant strains and the mutants harbouring the empty pNZ44 vector exhibited the same frequency of lysogenization (10^{-6}, Table 3.4) as the E119 and E126 mutant strains indicating that pNZ44 has no impact on internalisation or lysogeny, and that in the absence of the complementing plasmid the ability of TP901-1*erm* to lysogenize E119 and E126 is substantially reduced.

3.3.5 Reduction of *gtfA*$_{3107}$ expression by the use of the silencing plasmid pNZS*gtfA*

To further validate that *gtfA* is involved in host interactions of phage TP901-1, a gene silencing plasmid (pNZS*gtfA*) was constructed by cloning the *gtfA* gene from *L. lactis* 3107 in the antisense orientation into pNZ44. The resulting plasmid, designated pNZS*gtfA*, was introduced into *L. lactis* 3107 in an effort to reduce *gtfA* expression by antisense-mediated silencing. The effects of *gtfA* silencing were assessed using TP901-1*erm* lysogenization frequencies and E.O.P. assays (Table 3.4). Interestingly, *L. lactis* 3107 harbouring the silencing plasmid exhibited a two log decrease in lysogenization frequency compared to the strain 3107 ($5.71 \times 10^{-5} \pm 1.79 \times 10^{-6}$ vs $5.53 \times 10^{-3} \pm 4.10 \times 10^{-4}$). This decrease is of statistical significance with a *p*-value of <0.05. Plasmids pNZ44 and pNZ*gtfA* were transformed into *L. lactis* 3107 as controls and the resulting two strains displayed frequencies of lysogeny that were comparable to *L. lactis* 3107 (Table 3.4). Interestingly, the introduction of pNZS*gtfA* caused a minor, but statistically significant reduction (*p*-value of <0.05) in the plaquing efficiency of TP901-1, and a noteworthy reduction in plaque size (Table 3.4). To confirm that the presence of pNZ44 did not affect TP901-1 plaquing efficiency or plaque size, plaque assays were also performed on *L. lactis* 3107 harbouring pNZ44 or pNZ*gtfA* separately. The E.O.P. and plaque morphology of TP901-1 on all of these strains appeared to be similar to that of *L. lactis* 3107. As expected, LC3 was able to infect all tested strains, including *L. lactis* 3107 harbouring the silencing plasmid, with the same efficiency as observed for the control *L. lactis* 3107 (Table 3.4). Furthermore, the silencing plasmid (pNZS*gtfA*) was transformed into *L. lactis* subsp. *cremoris* NZ9000, which harbours an almost identical (99 % identity) homologue of $gtfA_{3107}$ on its genome. Plaque assays were performed using the NZ9000-infecting 936 group phages jj50, p2, sk1 and 712 and the namesake of the c2 group (Table 3.1, data not shown) and no obvious changes in E.O.P. or plaque size were observed. Moreover, lysis-in-broth assays were performed to further analyse the effects of silencing *gtfA* on TP901-1 phage development in liquid cultures. Figure 3.1A highlights that *gtfA* silencing causes a delay in lysis allowing the strain to grow unaffected for a longer period of time as an extended exponential growth phase of the host was observed, similar to the uninfected control. Furthermore, the infection capability of TP901-1 does not appear to be affected by the presence of pNZ44 as the lysis rate is similar to that of *L. lactis* 3107 in the presence of TP901-1 which is shown by the similar growth/lysis pattern observed for *L. lactis* 3107 and *L. lactis* 3107 harbouring pNZ44 when each of these strains is infected by this phage. These results indicate that

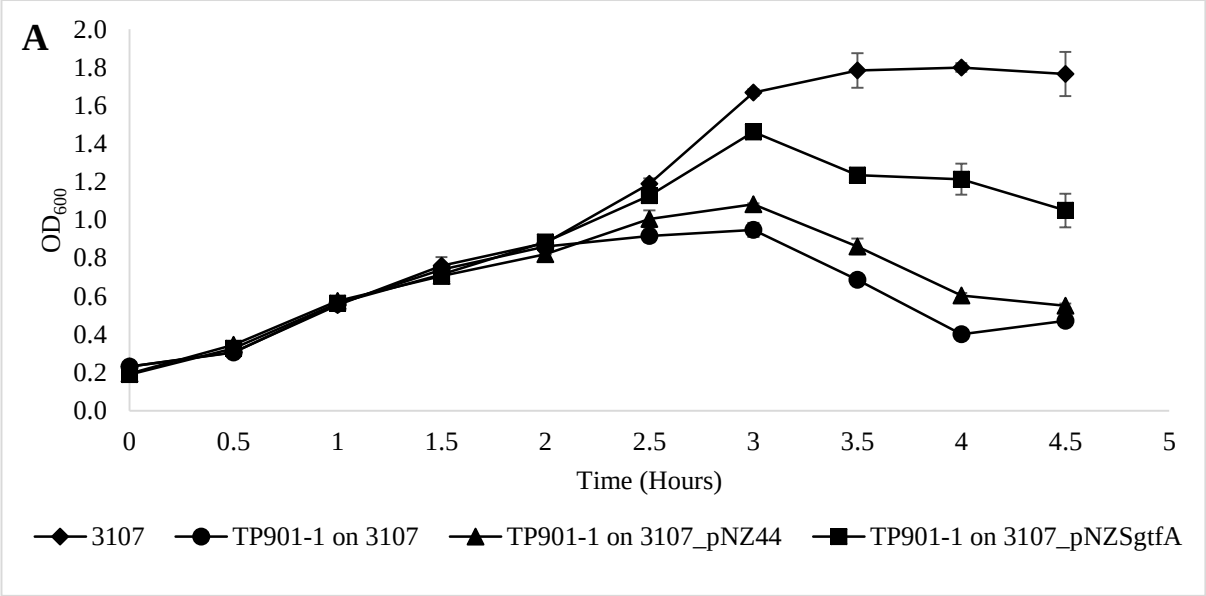
decreased expression of GtfA in *L. lactis* 3107 interferes with the infection ability of TP901-1. As a control lysis-in-broth assays were performed using LC3 as shown in Figure 3.1B, which demonstrates lysis of the control and test cultures in congruence with the plaque assay data (Table 3.4). Furthermore, adsorption assays were performed to analyse if the complementation of the mutant strains, as well as the mutation and silencing of *gtfA* had an effect on the reversible and/or irreversible binding of TP901-1 (Table 3.5). TP901-1 adsorbs reversibly and irreversibly to *L. lactis* 3107 and its derivatives with a high efficiency (Table 3.5). Mutation or silencing of *gtfA* has a negative impact on TP901-1 irreversible adsorption (*p*-values of <0.05). Furthermore, when *L. lactis* strains E119 and E126 harbour the complementing plasmid (pNZ*gtfA*) an 8.1 ± 0.2 % and 14.5 ± 1.1 % improvement in irreversible adsorption by TP901-1 is observed, respectively, showing a statistical significant (*p*-values of <0.05) increase. In all, these findings indicate that Gtf41 is involved in the first stages of infection, most likely irreversible adsorption and/or DNA injection of TP901-1.

Table 3.4: Restoration of phage sensitivity phenotype of *L. lactis* 3107 mutants E119 and E126 and silencing of *gtfA*.

Strain	E.O.P. of TP901-1	E.O.P. of LC3	TP901-1*erm* Lysogenization Frequency
L. lactis subsp. *cremoris*			
3107	1	1	$5.53 \times 10^{-3} \pm 4.10 \times 10^{-4}$
E119	$\leq 10^{-8}$	0.71 ± 0.10	$5.54 \times 10^{-6} \pm 2.93 \times 10^{-7}$
E119_pNZ44	$\leq 10^{-8}$	0.71 ± 0.08	$5.23 \times 10^{-6} \pm 2.52 \times 10^{-7}$
E119_pNZ*gtfA*	0.86 ± 0.01	0.76 ± 0.07	$3.03 \times 10^{-4} \pm 4.93 \times 10^{-6}$
E119_pNZ*gtfA* cured	$\leq 10^{-8}$	0.78 ± 0.08	$3.00 \times 10^{-6} \pm 5.00 \times 10^{-7}$
E126	$\leq 10^{-8}$	0.71 ± 0.09	$5.39 \times 10^{-6} \pm 2.11 \times 10^{-7}$
E126_pNZ44	$\leq 10^{-8}$	0.71 ± 0.06	$4.40 \times 10^{-6} \pm 4.00 \times 10^{-7}$
E126_pNZ*gtfA*	0.86 ± 0.01	0.77 ± 0.07	$4.17 \times 10^{-4} \pm 3.58 \times 10^{-4}$
E126_pNZ*gtfA* cured	$\leq 10^{-8}$	0.75 ± 0.04	$4.93 \times 10^{-6} \pm 3.06 \times 10^{-7}$
3107_pNZ44	0.76 ± 0.12	0.70 ± 0.07	$4.37 \times 10^{-3} \pm 3.29 \times 10^{-4}$
3107_pNZ*gtfA*	0.75 ± 0.12	0.66 ± 0.06	$4.44 \times 10^{-3} \pm 2.12 \times 10^{-4}$
3107_PNZS*gtfA*	$*0.10 \pm 0.01$	0.69 ± 0.05	$5.71 \times 10^{-5} \pm 1.79 \times 10^{-6}$

*hazy plaques and plaque size reduction was observed

TP901-1 and LC3 E.O.P., and TP901-1*erm* lysogenization frequencies for *L. lactis* 3107, its derivatives harbouring the complementing plasmid (pNZ*gtfA*) or the empty vector pNZ44, plasmid pNZ*gtfA*-cured derivatives, and *L. lactis* 3107 harbouring the silencing plasmid (pNZS*gtfA*), over-expressing plasmid (pNZ*gtfA*) or empty vector pNZ44.



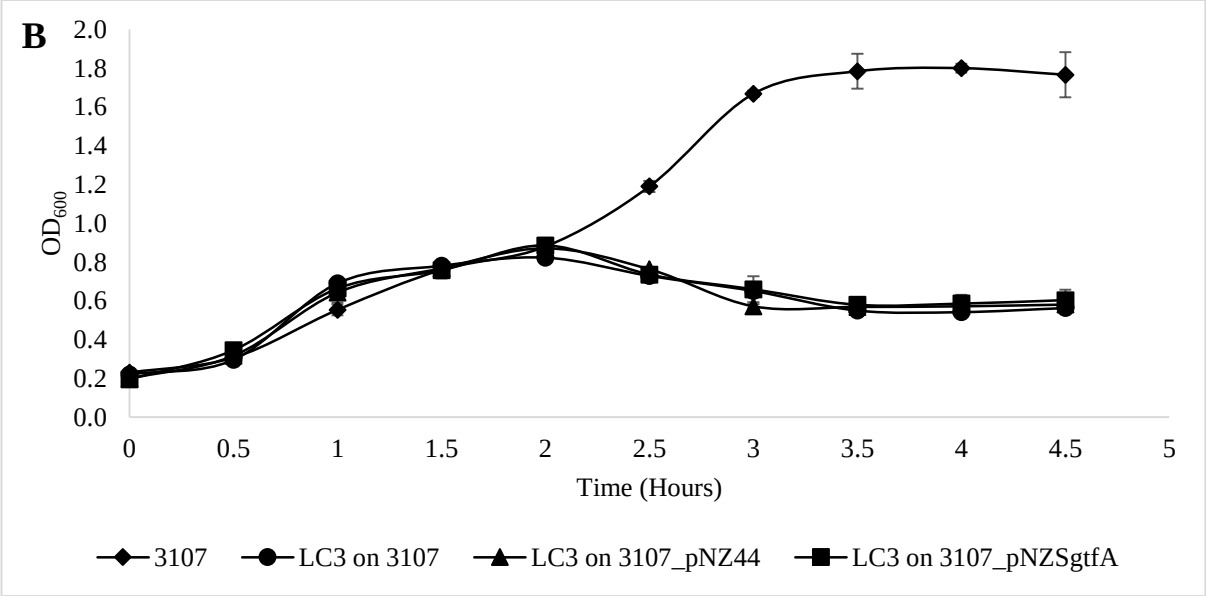


Figure 3.1: Effects of silencing ORF*gtfA*$_{3107}$ for TP901-1 phage propagation. TP901-1 (A) and LC3 (B) lysis-in-broth assays performed on *L. lactis* 3107 WT, *L. lactis* 3107 harbouring the silencing plasmid (pNZS*gtfA*) and pNZ44 individually.

Table 3.5: Adsorption (%) of phages TP901-1 to *L. lactis* subsp. *cremoris* 3107 and its derivatives.

	Adsorption of TP901-1 (%)	
Strain	**Reversible**	**Irreversible**
L. lactis		
3107	98.0 ± 0.2	97.5 ± 0.3
3107_pNZ44	97.7 ± 0.2	97.2 ± 0.1
3107_pNZ*gtfA*	94.0 ± 0.6	93.1 ± 0.9
3107_pNZS*gtfA*	92.4 ± 0.1	85.9 ± 1.5
E119	94.7 ± 0.4	75.2 ± 1.5
E119_pNZ44	91.3 ± 0.8	73.2 ± 0.8
E119_pNZ*gtfA*	94.1 ± 0.8	83.3 ± 0.2
E126	95.7 ± 0.7	68.2 ± 0.8
E126_pNZ44	90.3 ± 0.5	66.6 ± 1.7
E126_pNZ*gtfA*	95.7 ± 0.1	82.4 ± 1.1
NZ9000	18.9 ± 5.2	6.30 ± 5.2

3.3.6 Involvement of *gtfA* in infection of representatives of the P335, 936, P087 and 949 lactococcal phage groups

In order to determine if *gtfA* is also required for infection by other lactoccocal phages capable of infecting *L. lactis* 3107, a number of phages belonging to the P335 (Dub35A, 63301, 58601, 86501, 50101 and 07501), 936 (62605, 62604, 62601, 66901, 66902 and 66903), P087 (P087) and 949 (949, WRP3, L47 and LW81) groups (Table 3.1) were tested on the following strains: *L. lactis* 3107 WT; *L. lactis* 3107 harbouring pNZ*gtfA*, pNZS*gtfA* or pNZ44; *L. lactis* E119 and E126; strains E119 and E126 harbouring pNZ*gtfA* or pNZ44.

Phages 949 and WRP3 were shown to infect E119 and E126, although the E.O.P. of these phages was reduced by 4 orders of magnitude compared to the WT strain 3107 (Table 3.6). Furthermore, the observed statistically significant E.O.P. reduction (p-values of <0.05) in these mutants was reverted to WT levels when E119 and E126 harboured plasmid pNZ*gft41*. Silencing and overexpression of *gtfA* in the WT background appeared to have no impact on the E.O.P. or plaque size of 949 and WRP3, nor did the presence of the control plasmid pNZ44 cause any effect on these two parameters of infection. This indicates that these phages employ GtfA as a component of their infection pathway. No obvious changes in E.O.P. or plaque size were observed for the assessed P335, 936 and P087 phages, or for 949 phages L47 and LW81 (data not shown).

Table 3.6: Testing the 949 phages, 949 and WRP3, on *L. lactis* subsp. *cremoris* 3107 and its derivatives.

L. lactis* subsp. *cremoris	**E.O.P. of phages belonging to the 949 phage group***	
	949	**WRP3**
3107	1 (0.5 mm)	1 (0.6 mm)
3107_pNZ44	0.83 ± 0.14 (0.5 mm)	0.72 ± 0.29 (0.6 mm)
3107_pNZ*gtfA*	0.78 ± 1.0.02 (0.5mm)	0.69 ± 0.33 (1 mm)
3107_pNZS*gt41*	0.36 ± 0.08 (0.5 mm, slightly hazy plaques)	0.32 ± 0.07 (0.6 mm)
E119	$5.82 \times 10^{-4} \pm 9.06 \times 10^{-5}$ (1mm)	$9.11 \times 10^{-4} \pm 2.59 \times 10^{-4}$ (1.1mm)
E119_pNZ44	$6.06 \times 10^{-4} \pm 1.06 \times 10^{-4}$ (1.1 mm)	$1.04 \times 10^{-3} \pm 2.69 \times 10^{-4}$ (1.2 mm)
E119_pNZ*gtfA*	0.19 ± 0.01 (0.6mm)	0.36 ± 0.05 (0.9mm)
E126	$9.80 \times 10^{-4} \pm 1.15 \times 10^{-4}$ (1mm)	$1.19 \times 10^{-3} \pm 1.87 \times 10^{-4}$ (1mm)
E126_pNZ44	$8.13 \times 10^{-4} \pm 7.66 \times 10^{-5}$ (1.1 mm)	$9.33 \times 10^{-4} \pm 1.20 \times 10^{-4}$ (2 mm)
E126_pNZ*gtfA*	0.64 ± 0.05 (1mm)	0.89 ± 0.19 (1mm)

*The E.O.P. of phage applied and the plaque size in mm is shown in brackets.

3.3.7 *In silico* analysis of GtfA

L. lactis 3107 derivatives E119 and E126 both carry a mutation in *gtfA* that introduces a stop codon in the middle of this gene thus likely rendering *gtfA* defunct in these mutants (Fig. 3.2). HHpred modelling (66) of the deduced protein product of $gtfA_{3107}$ predicts that GtfA is an integral membrane protein (with fourteen predicted transmembrane regions), possessing a glycosyltransferase domain and a carbohydrate binding domain with high probabilities of 96.8 % and 61.2 %, respectively. GtfA characteristics are similar to members of the GT-C family of glycosyltransferases which are characterized by 8 – 13 transmembrane helices and an extracellular loop (76). GtfA may thus be a member of the GT-C family of glycosyltransferases, which are known to be involved in glycosylation of substrates at the periplasmic side of the cytoplasmic membrane (77). *In silico* analysis using BLASTp (68) was performed to analyse the presence and conservation of the *gtfA* gene in other lactococcal strains, revealing that it is present and conserved in all assessed *L. lactis* genomes. *gtfA* homologues in subsp. *cremoris* strains bear 99 % identity to that of $gtfA_{3107}$, while its equivalents in all assessed *L. lactis* subsp. *lactis* strains have a reduced sequence identity of 84 % indicating a subspecies-specific sequence for this gene. Inspection of the genomic region surrounding $gtfA_{3107}$ suggests that this gene is part of a tricistronic operon with a gene upstream and downstream of $gtfA_{3107}$, represented by *L3107_1441* and *L3107_1443*, respectively (Fig. 3.2). A previous transcriptome analysis performed on the genome of *L. lactis* MG1363 revealed that its homologue of $gtfA_{3107}$, which is designated *LLMG_RS08685*, is indeed part of a tricistronic operon with a promoter identified directly upstream of the coding sequence of the homologue of *L3107_1441*, *LLMG_RS08680* which is found upstream of the $gtfA_{3107}$ homologue (78). *L3107_1441* is predicted to encode a galactose-1-phosphate-uridylyltransferase (GalT) with a high probability (100 %) based on HHPred analysis, while *L3107_1443* is predicted to encode a protein exhibiting one transmembrane helix of unknown function. Similar to the *gtfA* gene, *L3107_1441* is also present and conserved in all *L. lactis* genomes sequenced to date and homologues of this gene in *L. lactis* subsp. *cremoris* strains bear 99 % identity to that of *L3107_1441.* In contrast, *L3107_1443* is not present in all *L. lactis* genomes sequenced to date, yet homologues of this gene are present in all assessed *L. lactis* subsp. *cremoris* and the majority of analysed *L. lactis* subsp. *lactis* strains bearing 94 - 100 % identity to that of *L3107_1443*. The GalT enzyme catalyses the reversible transfer of an uridine 5'-monophosphate (UMP)

moiety from uridine 5'-diphosphate (UDP)-glucose to the phosphate group of galactose 1-phosphate to form UDP-galactose. Following the binding of UDP-glucose to the GalT enzyme, an uridylylated enzyme intermediate is generated, and glucose 1-phosphate is released. Subsequently, galactose 1-phosphate binds to the active site and the UMP moiety is transferred to generate UDP-galactose (79). Based on the predicted function of GalT, this enzyme may be involved in the production of the substrate for GtfA.

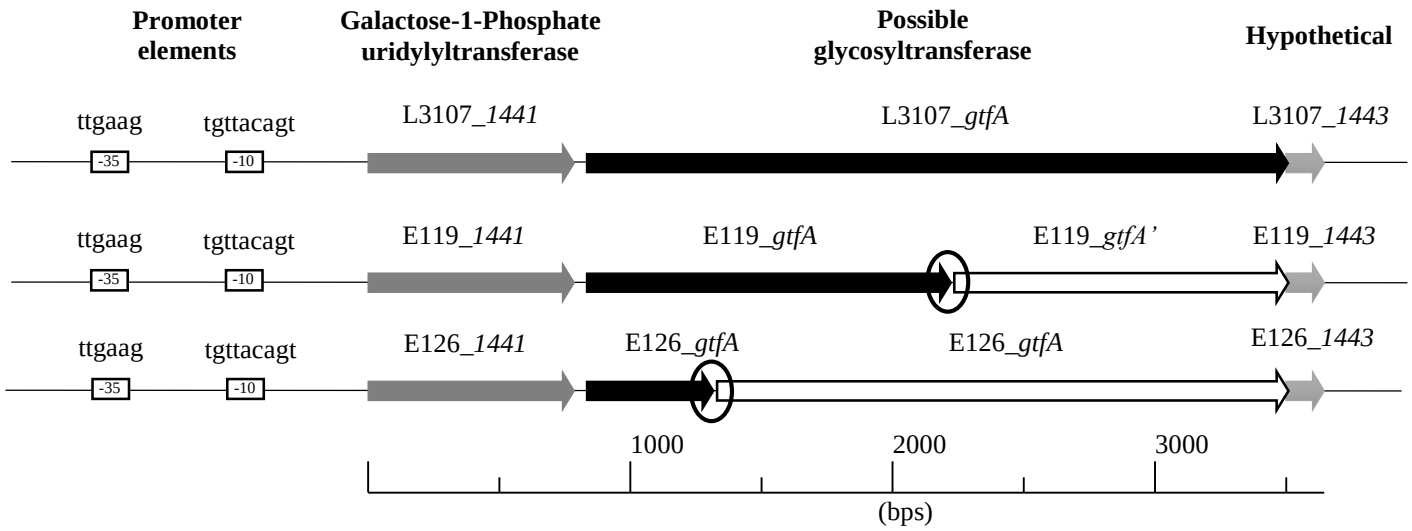


Figure 3.2: Schematic representation of the genomic region surrounding of *gtfA* in *L. lactis* 3107 and derived mutants E119 and E126. Predicted promoter elements -10 and -35 are located upstream *L3107_1441*; both mutations lead to the incorporation of a stop codon in *gtfA*, the position of which is indicated by a circle.

3.3.8 Restoration of the phage-sensitivity phenotype using a homologue of *gtfA* from *L. lactis* subsp. *lactis* IL1403

The identified homologue of *gtfA* on the *L. lactis* subsp. *lactis* IL1403 genome, designated as *yieH*, was cloned into pNZ44 (resulting in plamid pNZ*yieH*) and introduced into *L. lactis* 3107 mutants E119 and E126. Plaque assays were performed in order to determine if these mutants harbouring pNZ*yieH* restored their phage-sensitivity phenotype (Table 3.7). The results revealed that TP901-1 was indeed able to form plaques on E119 and E126 mutants harbouring pNZ*yieH* with the same E.O.P. as on the *L. lactis* 3107 and *L. lactis* 3107 harbouring the pNZ44 empty vector. Furthermore, as expected, following the loss of pNZ*yieH* from strains E119 and E126, TP901-1 was incapable of forming visible plaques on cured mutant strains, exhibiting an E.O.P. of $\leq 10^{-8}$, confirming that the presence of the pNZ*yieH* construct caused the TP901-1 phage sensitivity phenotype to be restored in both mutants. These results demonstrate that the expression of *yieH* (which bears 84 % sequence identity to *gtfA*) in E119 and E126 restores their TP901-1-sensitivity phenotype. As anticipated, phage LC3 formed visible plaques on all strains with the same efficiency as on the *L. lactis* 3107 WT.

Table 3.7: Effects of expressing *yieH* on infectivity TP901-1 and LC3.

Strain	E.O.P. of TP901-1	E.O.P. of LC3
L. lactis subsp. *cremoris*		
3107	1	1
E119	$\leq 10^{-8}$	0.71 ± 0.10
E119_pNZ44	$\leq 10^{-8}$	0.71 ± 0.08
E119_ pNZ*yieH*	0.85 ± 0.01	0.75 ± 0.06
E119_ pNZ*yieH* cured	$\leq 10^{-8}$	0.76 ± 0.02
E126	$\leq 10^{-8}$	0.71 ± 0.09
E126_pNZ44	$\leq 10^{-8}$	0.71 ± 0.06
E126_ pNZ*yieH*	0.85 ± 0.01	0.75 ± 0.04
E126_ pNZ*yieH* cured	$\leq 10^{-8}$	0.78 ± 0.10

TP901-1 and LC3 E.O.P. on (i) *L. lactis* 3107, (ii) *L. lactis* 3107 derivatives harbouring the complementing plasmid (pNZ*yieH*), or the empty vector pNZ44, or its pNZ*yieH*-cured derivative.

3.3.9 Construction of *L. lactis* subsp. *cremoris* NZ9000 mutants

As previously indicated (Section 3.3.7), *L. lactis* subsp. *cremoris* NZ9000 harbours an almost identical (99 % identity) homologue of $gtfA_{3107}$ on its genome named *LLNZ_08920*, designated here as $gtfA_{NZ9000}$. Furthermore, the gene upstream (*LLNZ_08915*) and downstream (*LLNZ_08925*) of $gtfA_{NZ9000}$ represent near identical (99 % identity) homologues of *L3107_1441* and *L3107_1443*, while the gene order was also conserved in the genomes of these strains. We had already shown that GtfA is not required for infection by *L. lactis* NZ9000-infecting 936 group phages as expressing the silencing plasmid pNZS*gtfA* in *L. lactis* NZ9000 had no negative effect on the sensitivity profile of the strain (Section 3.3.5). The targeted genes (*LLNZ_08915*, $gtfA_{NZ9000}$ and *LLNZ_08925*) and overall genome of *L. lactis* NZ9000 is 99 % identical to that of *L. lactis* 3107. Therefore, by means of targeted mutagenesis using recombineering technology (58, 72) the homologues of $gtfA_{3107}$ and/or *L3107_1441* in the chromosome of *L. lactis* NZ9000 were mutated which resulted in the construction of three *L. lactis* NZ9000 derivatives, i.e. NZ9000*Δ8915*, NZ9000*ΔgtfA*, and NZ9000*Δ8915_gtfA* (Table 3.3). These derivatives were generated in such a manner so as to introduce stop codons (Table 3.3) into the coding sequence of *LLNZ_08915*, $gtfA_{NZ9000}$, or in both *LLNZ_08915* and $gtfA_{NZ9000}$, respectively. Thus, these NZ9000 mutant strains are not expected to produce the predicted galactose-1-phosphate-uridylyltransferase (LLNZ_08915), the GtfA glycosyltransferase (LLNZ_08920), or either of these two proteins, respectively. These NZ9000 derivatives were then tested to assess if either or both of these mutated genes are involved in the infection process of NZ9000-infecting phages.

3.3.10 Involvement of $gtfA_{NZ9000}$ and/or *LLNZ_08915* in infection by representatives of the 936 and c2 lactococcal phage groups

Quantitative plaque assays were performed in order to determine if, and to what extent, $gtfA_{NZ9000}$ and/or *LLNZ_08915* are required for infection by lactoccocal phages capable of infecting *L. lactis* NZ9000 (Table 3.8). Phages belonging to the 936 (jj50, 712, sk1 and p2) and c2 (c2) groups were tested against *L. lactis* NZ9000 and its derivatives NZ9000*Δ8915*, NZ9000*ΔgtfA*, and NZ9000*Δ8915_gtfA*. Furthermore, in order to determine if the overexpression of L3107_1441, GtfA, or a combination of GtfA and L3107_1441 has an effect on the infection characteristics of these 936 and c2 phages, these lactococcal phages were tested on *L. lactis* NZ9000 derivatives

harbouring pNZ*1441*, pNZ*gtfA* and pNZ*1441_gtfA*, respectively. Plaque assays were also performed on *L. lactis* NZ9000 harbouring the empty pNZ44 vector as a control. The mutation and overexpression of *gtfA* and *LLNZ_08915* in the *L. lactis* NZ9000 background appeared to have no impact on the E.O.P. or plaque size of the assessed NZ9000-infecting phages, nor did the presence of the control plasmid pNZ44 cause any effect on these two parameters of infection. This suggests that these phages do not employ GtfA and/or LLNZ_08915 as a component of their infection pathway or perhaps an additional required component which is present in *L. lactis* 3107 is being utilized in combination with GtfA which is not present in *L. lactis* NZ9000.

Table 3.8: Testing lactococcal 936 and c2 phage groups on *L. lactis* subsp. *cremoris* NZ9000 and its derivatives; the E.O.P of phage applied and plaque size in mm is shown.

Strain	E.O.P of 936 group (plaque size)				E.O.P of c2 group (plaque size)
	jj50	**712**	**sk1**	**p2**	**c2**
L. lactis subsp. *cremoris*					
NZ9000	1.0 ± 0.00 (2 mm)	1.0 ± 0.00 (1.5 mm)	1.00 ± 0.00 (1.5 mm)	1.00 ± 0.00 (2 mm)	1.00 ± 0.00 (2.5 mm)
NZ9000Δ*8915*	0.98 ± 0.03 (2 mm)	0.81 ± 0.48 (1.5 mm)	0.80 ± 0.02 (1.5 mm)	0.86 ± 0.48 (2 mm)	0.91 ± 0.02 (2.5 mm)
NZ9000Δ*gtfA*	0.92 ± 0.05 (2 mm)	0.88 ± 0.55 (1.5 mm)	1.33 ± 0.07 (2 mm)	0.95 ± 0.14 (2 mm)	0.65 ± 0.03 (2.5 mm)
NZ9000Δ*8915_ gtfA*	0.93 ± 0.05 (2 mm)	1.18 ± 0.70 (1.5 mm)	0.82 ± 0.02 (2 mm)	0.94 ± 0.50 (2 mm)	0.86 ± 0.05 (2.5 mm)
NZ9000_pNZ44	0.41 ± 0.04 (2 mm)	0.74 ± 0.43 (1.5 mm)	0.58 ± 0.04 (1.5 mm)	0.39 ± 0.20 (2 mm)	0.75 ± 0.02 (2.5 mm)
NZ9000_pNZ*1441*	0.53 ± 0.03 (2 mm)	0.65 ± 0.37 (1.5 mm)	0.67 ± 0.02 (1.5 mm)	0.88 ± 0.44 (2 mm)	0.47 ± 0.03 (2.5mm)
NZ9000_pNZ*gtfA*	0.76 ± 0.03 (2 mm)	0.59 ± 0.34 (1.5 mm)	0.76 ± 0.02 (1.5 mm)	0.77 ± 0.37 (2 mm)	0.75 ± 0.01 (2.5 mm)
NZ9000_pNZ*1441_gtfA*	0.41 ± 0.05 (2 mm)	0.69 ± 0.39 (1.5 mm)	0.49 ± 0.03 (1.5 mm)	0.52 ± 0.28 (2 mm)	0.72 ± 0.05 (2.5 mm)

3.3.11 E.O.P. and lysogenization assays of TP901-1*erm* escape mutants

In an attempt to generate TP901-1 (so-called escape) mutants capable of infecting the TP901-1-resistant *L. lactis* E119 strain, TP901-1*erm* was repeatedly exposed to strain E119 (see Materials and Methods). This resulted in the isolation of two TP901-1*erm* escape mutants named TP901-1*erm*-E119A and TP901-1*erm*-E119B. Similarly, exposure of TP901-1*erm* to strains E121 and E126 allowed the isolation of four additional TP901-1*erm* escape mutants named TP901-1*erm*-E121A and TP901-1*erm*-E121B, and TP901-1*erm*-E126A and TP901-1*erm*-E126B, respectively. The E.O.P. of each TP901-1*erm* escape mutant and the TP901-1*erm* WT was determined on *L. lactis* E119, E121 and E126, as well as the 3107 WT strain (Table 3.9). As expected, TP901-1*erm* was not able to form visible plaques on E119, E121 and E126, yielding an E.O.P. of $\leq 10^{-8}$. Interestingly, all TP901-1*erm* escape mutants, irrespective of the mutant strain used to isolate them, were able to form plaques on all three TP901-1-resistant strains, E119, E121 or E126, although with a minor yet significantly reduced E.O.P. (p-values <0.05) as compared to the *L. lactis* 3107 WT (Table 3.9). The infection of both strains, E119 and E126 by TP901-1*erm*-E119A/B and TP901-1*erm*-E126A/B was perhaps unsurprising as both strains carry a mutation in *gtfA*, but the infection of the E121 strain by these escape mutants and the infection of E119 and E126 by TP901-1*erm*-E121A/B was somewhat unexpected as E121 does not carry a mutation in *gtfA*. This suggests that the three TP901-1-resistant strains (E119, E121 and E126) harbour mutations involved in the same pathway of infection by TP901-1, though a different step of the pathway is mutated for E121 (unknown) than for both E119 and E126 (GtfA), and that the escape mutants somehow bypass the requirement for this pathway for infection.

To further support the E.O.P. results which suggest that all of the TP901-1*erm* escape mutants have overcome the phage resistant phenotype of E119, E121 and E126, lysogenization assays were performed. For this purpose, TP901-1*erm* was used instead of the TP901-1 WT when attempting to isolate TP901-1*erm* escape mutants as the Em^r marker would allow us to determine the frequency of lysogenization achieved by the TP901-1*erm* escape mutants based on the number of Em^r colonies obtained per infecting phage (see Materials and Methods). As expected, TP901-1*erm* lysogenization frequencies of the three *L. lactis* 3107 derivative strains, E119, E121 and E126, were lower than that of the 3107 WT strain (10^{-3} vs. 10^{-6}, Table 3.9). Interestingly, when the TP901-1*erm* escape mutants were applied, an increase by three

orders of magnitude in the frequency of lysogeny was observed for the E119, E121 and E126 derivatives compared to the TP901-1*erm* WT (10^{-3} vs. 10^{-6}, Table 3.9), showing a statistically significant increase (*p*-values <0.05). This suggests that the efficiency of infection and lysogeny frequency of the TP901-1*erm* escape mutants was restored to the level of the WT TP901-1*erm* phage.

Table 3.9: Testing TP901-1*erm* escape mutants on *L. lactis* subsp. *cremoris* 3107 and its derivatives; the E.O.P. and lysogenization frequency are indicated for each tested phage.

	***E.O.P.**				***Lysogenization Frequency**			
L. lactis subsp. *cremoris*	TP901-1*erm* _E119A	TP901-1*erm* _E126A	TP901-1*erm* _E121A	TP901-1*erm*	TP901-1*erm* _E119A	TP901-1*erm* _E126A	TP901-1*erm* _E121A	TP901-1*erm*
3107	1.65	1.65	1.65	1.65	$2.44 \times 10^{-3} \pm 2.19 \times 10^{-4}$	$2.41 \times 10^{-3} \pm 1.20 \times 10^{-4}$	$2.47 \times 10^{-3} \pm 5.86 \times 10^{-5}$	$5.53 \times 10^{-3} \pm 4.10 \times 10^{-4}$
E119	1.00 ± 0.01	0.85 ± 0.01	0.85 ± 0.03	$\leq 10^{-8}$	$5.49 \times 10^{-3} \pm 1.27 \times 10^{-4}$	$2.43 \times 10^{-3} \pm 1.57 \times 10^{-4}$	$2.47 \times 10^{-3} \pm 1.78 \times 10^{-4}$	$5.54 \times 10^{-6} \pm 2.93 \times 10^{-7}$
E121	0.85 ± 0.01	0.85 ± 0.01	1.00 ± 0.03	$\leq 10^{-8}$	$2.43 \times 10^{-3} \pm 2.53 \times 10^{-4}$	$2.45 \times 10^{-3} \pm 1.29 \times 10^{-4}$	$5.45 \times 10^{-3} \pm 1.08 \times 10^{-4}$	$5.31 \times 10^{-6} \pm 2.54 \times 10^{-7}$
E126	0.85 ± 0.01	1.00 ± 0.01	0.85 ± 0.03	$\leq 10^{-8}$	$2.45 \times 10^{-3} \pm 1.51 \times 10^{-4}$	$5.43 \times 10^{-3} \pm 2.10 \times 10^{-4}$	$2.46 \times 10^{-3} \pm 1.87 \times 10^{-4}$	$5.39 \times 10^{-6} \pm 2.11 \times 10^{-7}$

*Similar results were observed for the three remaining TP901-1*erm* escape mutants, TP901-1*erm*_E119B, TP901-1*erm*_E121B, and TP901-1*erm*_E126B.

3.3.12 Genome sequencing and bioinformatic analysis of TP901-1*erm* escape mutants

Analysis of the six escape mutant phages, i.e. TP901-1*erm*-E119_A/B, TP901-1*erm*-E121_A/B and TP901-1*erm*-E126_A/B, showed that they have overcome the TP901-1 phage-resistant phenotype of the E119, E121 and E126 strains (Table 3.9). The genomes of these six TP901-1*erm* escape mutants as well as the genome of TP901-1*erm* WT, were sequenced in order to identify the mutated gene(s) that is (are) responsible for allowing the phage to overcome the phage-resistance phenotype. The genome of TP901-1 WT is 37.6 Kb in length with 56 predicted protein-coding regions and possesses 35.38 % G+C content (9). The genomes of the TP901-1*erm* derivatives as well as the WT exhibit similar characteristics to that of the parent phage. Following comparative genome analysis between the Illumina sequenced WT and derivatives, deletion or insertion events were not identified, and therefore a SNP analysis was performed to identify point mutations that may have caused the TP901-1*erm* escape mutants to overcome phage-resistance. Moreover, a SNP analysis was performed using the published TP901-1 data and the Illumina TP901-1*erm* WT data to rule out sequencing errors, thus confirming the integrity of the SNPs identified in the genomes on the TP901-1*erm* escape mutants.

This resulted in the identification of 96, 178, 128, 45, 111 and 5 SNPs on the genomes of TP901-1*erm*_E119A, E119B, E121A, E121B, E126A and E126B, respectively. In total, 563 SNPs (Table S3.2) were identified with 113 of these SNPs occurring in more than one of the TP901-1*erm* escape mutant genomes, and with some genes containing more than one mutation in a given genome. Interestingly, the only gene mutated in all of the six TP901-1 mutant genomes was *tal* (tail-associated lysin/*orf47*), with each genome exhibiting one or more SNPs in the coding sequence of this 2,757 bp *tal* gene. Moreover, the genome of the TP901-1*erm*_E119B escape mutant contains 5 SNPs, all of which are located within the *tal* gene, thus clearly implicating Tal in the phage DNA injection process and suggesting that TP901-1 can undergo specific *tal* modification(s) to overcome the phage-insensitivity of the E119, E121 and E126 mutants.

Interestingly, five of the TP901-1*erm* escape mutants, E119A/B, E121A/B, and E126A contain a single mutation in *tal*, with four of these mutants exhibiting the same SNP (T to C), while TP901-1*erm*_E126B contains five mutations in *tal* as detailed in Table 3.10. A comparative amino acid sequence analysis was performed between the

predicted protein sequence of each of the genetically altered Tal proteins and $Tal_{TP901-1}$ which revealed the specific mutations caused in the encoded amino acids as shown in Table 3.10. HHpred modelling of $Tal_{TP901-1}$ predicts that this protein possesses a structural and lytic domain. Based on this analysis the observed mutations in the encoded amino acids of the altered Tal proteins appear to be located in the structural or lytic domains as shown in Table 3.10 and Figure 3.3.

These findings indicate that a single nucleotide modification in the structural domain of *tal* or a small number of nucleotide modifications in the gene sequence encoding the lytic domain of *tal* is sufficient to modify the encoded $Tal_{TP901-1}$ in order to infect the E119, E121 and E126 mutants.

Table 3.10: Analysis of confirmed SNPs in $tal_{TP901\text{-}1}$ demonstrating SNP position and nucleotide change, as well as the resulting mutation in the encoded amino acid and its domain location.

Mutant	**SNP position in *tal***	**Nucleotide change**			**Amino acid change**	**Location / domain**
E119B	653	g	to	c	G218A	structural domain
E119A/E121A/E121B/E126A	1,141	t	to	c	W381R	structural domain
E126B	2,454	a	to	c	G818G	lytic domain
E126B	2,481	t	to	a	S827S	lytic domain
E126B	2,482	g	to	a	G828N	lytic domain
E126B	2,483	g	to	a	G828N	lytic domain
E126B	2,521	g	to	a	D841N	lytic domain

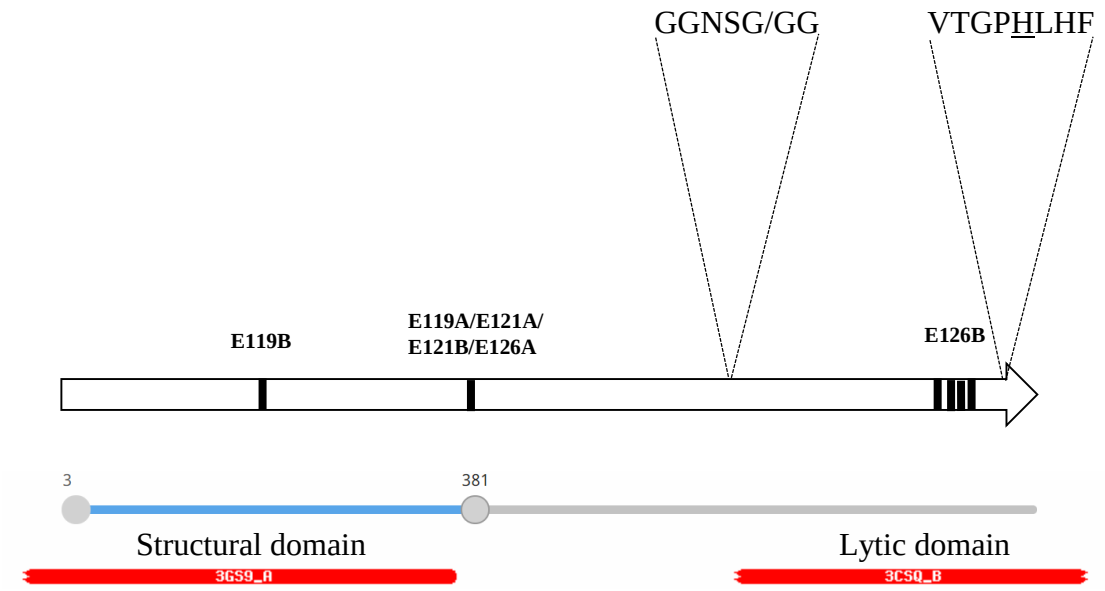


Figure 3.3: HHpred modelling of $Tal_{TP901\text{-}1}$ exhibiting predicted (99.9 % probability) structural and lytic domains. Arrow represents encoded Tal protein with vertical lines representing the amino acid position of mutations in the encoded protein. The autoproteolytic site GGNSG/GG at amino acid positions 598-604 and the catalytic centre VTGPHLHF at amino acid positions 888-895 are shown.

3.4 Discussion

Previous studies have demonstrated that phage-host interactions may require the involvement of more than one bacterial receptor. For example, the model siphophage SPP1, initially binds reversibly to a carbohydrate moiety on the cell surface of *B. subtilis*, followed by irreversible binding to the ectodomain of the YueB protein (28), which then triggers DNA ejection (29). In a similar manner, it has been shown that phage c2 irreversibly binds to the *L. lactis* subsp. *lactis* C2 membrane protein, PIP (phage infection protein), following an initial reversible adsorption to a cell wall carbohydrate (30, 31). Furthermore, the PtsG (a membrane glucose transporter protein) and FkpA (a periplasmic chaperone) proteins in *E. coli* are known to act as phage receptors for phage HK97 (80).

It has previously been suggested that the P335 temperate phages TP901-1 and LC3 either attach to different receptor(s) or utilize distinct pathways of injection (13, 41). The delineation of five genetically and morphologically distinct P335 lactococcal phage sub-groups further strengthens this notion since TP901-1 and LC3 belong to distinct sub-groups (sub-groups II and III, respectively) (13), where TP901-1 possesses a double-disc baseplate structure (81), while LC3 exhibits a "stubby" tail tip region (49, 82). Therefore, the distinct baseplate type of LC3 may represent a different receptor recognition platform that follows a specific host adsorption and DNA injection pathway (13). The results in the present study also support the notion that TP901-1 and LC3, either attach to different receptor(s) or utilize distinct pathways of DNA injection: expression of *gtfA* in the E119 and E126 mutant strains restored infectivity by phage TP901-1, yet had no effect on LC3 phage infection, while silencing of *gtfA* interfered with infection by phage TP901-1, yet had no impact on the infection ability of LC3.

This current study shows that the resistance of strains E119 and E126 to infection by phage TP901-1 was caused by mutations in the predicted glycosyltransferase-encoding gene, *gtfA*. This conclusion is supported by complementation assays in which phage sensitivity to TP901-1 was restored in both mutants. Additionally, silencing of *gtfA* interfered with infection by phage TP901-1, indicating that *gtfA* is involved in the initial interactions between TP901-1 and its host *L. lactis* 3107.

This study indicates that multiple genes are involved in phage-host interactions of TP901-1. The finding that the phage-resistant phenotype of only two out of the three

mutants was restored by the introduction of *gtfA* indicates that the third strain, E121, contains a mutation in a gene other than *gtfA*. It seems likely that *gtfA* is also involved in phage-host interactions of the rare lactococcal 949 phages (949 and WRP3) since E119 and E126 were shown to be less sensitive to these phages, while complementation assays fully restored infectivity by these phages. Interestingly, *gtfA* is ubiquitous among *Lactococcus* genomes. However, silencing and introducing a mutation (stop codon) in $gtfA_{NZ9000}$ had no effect on NZ9000-infecting 936 group phages.

Interestingly, a homologue of *gtfA* identified in the genome of the lactococcal strain IL1403 with 84 % nucleotide identity was able to fully restore the phage-sensitivity phenotype of E119 and E126, thus showing that a *gtfA* subspecies-specific sequence with reduced sequence identity is capable of acting as a target for TP901-1. It is noteworthy that P335 phages are not available for testing in this strain and it would be an interesting avenue of research to explore the effect of suppression of *gtfA* expression in lactococcal strains that are sensitive to (sub-group II) P335 phages.

The generation of TP901-1 escape mutants capable of infecting phage-resistant derivatives of the host strain 3107 revealed that TP901-1 is capable of evolving over time in order to adapt to cells which have acquired phage resistance. Interestingly, despite the fact that strains E119 and E126 do not share the same mutation (*gtfA*) responsible for their TP901-1 resistant phenotype as E121, all of the TP901-1 escape mutants generated in this study infect each of the three TP901-1-resistant strains. This indicates that even though the mutations in these strains are involved in different steps of (producing the molecule required for) the TP901-1 infection pathway, escape mutants are capable of bypassing the requirement for this pathway of infection.

Previous studies have demonstrated that the deletion of specific TP901-1 phage genes prevents or reduces infection of the host (83). Previously, it has been shown that the tail-associated lysin-encoding gene (*tal*) encodes the tail fiber protein of the TP901-1 distal tail structure which is believed to be involved in the opening of the DNA ejection conduit through a conformational change during infection (84). Here, we demonstrate that regardless of the mutated gene responsible for TP901-1-resistance, TP901-1 is capable of introducing specific mutation(s) in the coding sequence of *tal* allowing it to infect phage-resistant host cells. Interestingly, in order to overcome phage resistant strains TP901-1 appears to modify either the N-terminal or C-terminal part of Tal. The N-terminal part of $Tal_{TP901-1}$ may have a structural role,

whereas the C-terminal part of $Tal_{TP901-1}$ has been shown to encode a lytic domain necessary for the cleavage of peptidoglycan to allow DNA transfer into the cytoplasm of the host (19). This current study has demonstrated that alterations in the structural or lytic domain of Tal leads to successful infection of TP901-1 resistant strains. We hypothesize that the modified Tal is capable of recognizing a different host injection trigger or attaching to a modified host injection trigger, thereby generating the desired DNA release signal which is believed to trigger a cascade of conserved conformational changes in the tail tube to ultimately allow DNA release (84). This demonstrates that in a natural environment phages will adapt to host populations which have acquired phage resistance mechanism(s), and that phage evolution is continually driven by changes in the bacterial host population.

In conclusion, we identified *gtfA* as a genetic component required for infection of the model phage TP901-1 to its host and have demonstrated that subsequent adaptation of TP901-1 to such phage-resistant host cells is achieved through the modification of *tal*. Currently it is not clear how GtfA is involved in triggering DNA injection of TP901-1; GtfA may act directly as a proteinaceous injection trigger or the saccharidic product of its predicted transferase activity may be a target for TP901-1. Chapter IV of this thesis will provide information on efforts to identify the possible cell envelope-associated substrates for GtfA.

3.5 References

1. Teuber M. 1993. Lactic acid bacteria. Biotechnology set second edition:325-366.
2. Deveau H, Labrie SJ, Chopin M-C, Moineau S. 2006. Biodiversity and classification of lactococcal phages. Applied and environmental microbiology 72:4338-4346.
3. Coffey A, Ross RP. 2002. Bacteriophage-resistance systems in dairy starter strains: molecular analysis to application, p 303-321, Lactic acid bacteria: genetics, metabolism and applications. Springer.
4. Garneau JE, Moineau S. 2011. Bacteriophages of lactic acid bacteria and their impact on milk fermentations. Microbial cell factories 10:S20.
5. Emond E, Moineau S. 2007. Bacteriophages and food fermentations. Bacteriophage: genetics and molecular biology:93-124.
6. Maniloff J, Ackermann H-W. 1998. Taxonomy of bacterial viruses: establishment of tailed virus genera and the other *Caudovirales*. Archives of virology 143:2051-2063.
7. Solovyev V, Salamov A. 2011. Automatic annotation of microbial genomes and metagenomic sequences. Metagenomics and its applications in agriculture, biomedicine and environmental studies:61-78.
8. Blatny JM, Godager L, Lunde M, Nes IF. 2004. Complete genome sequence of the *Lactococcus lactis* temperate phage φLC3: comparative analysis of φLC3 and its relatives in lactococci and streptococci. Virology 318:231-244.
9. Brøndsted L, Østergaard S, Pedersen M, Hammer K, Vogensen FK. 2001. Analysis of the complete DNA sequence of the temperate bacteriophage TP901-1: evolution, structure, and genome organization of lactococcal bacteriophages. Virology 283:93-109.
10. Labrie S, Moineau S. 2002. Complete genomic sequence of bacteriophage ul36: demonstration of phage heterogeneity within the P335 quasi-species of lactococcal phages. Virology 296:308-320.
11. Labrie SJ, Josephsen J, Neve H, Vogensen FK, Moineau S. 2008. Morphology, genome sequence, and structural proteome of type phage P335 from *Lactococcus lactis*. Applied and environmental microbiology 74:4636-4644.
12. Mahanivong C, Boyce JD, Davidson BE, Hillier AJ. 2001. Sequence analysis and molecular characterization of the *Lactococcus lactis* temperate

bacteriophage BK5-T. Applied and environmental microbiology 67:3564-3576.

13. Mahony J, Martel B, Tremblay DM, Neve H, Heller KJ, Moineau S, Sinderen D. 2013. Identification of a new P335 subgroup through molecular analysis of lactococcal phages Q33 and BM13. Applied and environmental microbiology 79(14), 4401-4409.
14. Mahony J, Oliveira J, Collins B, Hanemaaijer L, Lugli GA, Neve H, Ventura M, Kouwen TR, Cambillau C, van Sinderen D. 2017. Genetic and functional characterisation of the lactococcal P335 phage-host interactions. BMC Genomics 18:146.
15. Seegers JF, Mc Grath S, O'Connell-Motherway M, Arendt EK, van de Guchte M, Creaven M, Fitzgerald GF, van Sinderen D. 2004. Molecular and transcriptional analysis of the temperate lactococcal bacteriophage Tuc2009. Virology 329:40-52.
16. Trotter M, McAuliffe O, Callanan M, Edwards R, Fitzgerald GF, Coffey A, Ross RP. 2006. Genome analysis of the obligately lytic bacteriophage 4268 of *Lactococcus lactis* provides insight into its adaptable nature. Gene 366:189-199.
17. Venema G, Nauta A. 1996. Sequence analysis and molecular characterization of the temperate lactococcal bacteriophage r1t. Molecular microbiology 19:1343-1355.
18. Garzon AE, Mahony J, Bottacini F, Kelleher P, van Sinderen D. 2019. Complete genome sequence of *Lactococcus lactis* subsp. *cremoris* 3107, host for the model lactococcal P335 bacteriophage TP901-1. Microbiology Resource Announcement 8:e01635-18.
19. Bebeacua C, Bron P, Lai L, Vegge CS, Brondsted L, Spinelli S, Campanacci V, Veesler D, Heel M, Cambillau C. 2010. Structure and molecular assignment of lactococcal phage TP901-1 baseplate. Journal of biological chemistry 285: 39079 –39086.
20. Bebeacua C, Lai L, Vegge CS, Brondsted L, Heel M, Veesler D, Cambillau C. 2013. Visualizing a complete *Siphoviridae* member by single-particle electron microscopy: the structure of lactococcal phage TP901-1. Journal of virology 87:1061-8.

21. Collins B, Bebeacua C, Mahony J, Blangy S, Douillard FP, Veesler D, Cambillau C, Sinderen D. 2013. Structure and functional analysis of the host recognition device of lactococcal phage Tuc2009. Journal of virology 87:8429-40.
22. Mc Grath S, Neve H, Seegers JF, Eijlander R, Vegge CS, Brondsted L, Heller KJ, Fitzgerald GF, Vogensen FK, Sinderen D. 2006. Anatomy of a lactococcal phage tail. Journal of bacteriology 188:3972-82.
23. Stockdale SR, Collins B, Spinelli S, Douillard FP, Mahony J, Cambillau C, Sinderen D. 2015. Structure and assembly of TP901-1 virion unveiled by mutagenesis. PLOS One 10:e0131676.
24. Veesler D, Spinelli S, Mahony J, Lichiere J, Blangy S, Bricogne G, Legrand P, Ortiz-Lombardia M, Campanacci V, Sinderen D. 2012. Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. Proceedings of the national academy of sciences of the United States of America 109:8954-8.
25. Oliveira J, Mahony J, Lugli GA, Hanemaaijer L, Kouwen T, Ventura M, van Sinderen D. 2016. Genome sequences of eight prophages isolated from *Lactococcus lactis* dairy strains. Genome announcements 4:e00906-16.
26. Rakhuba D, Kolomiets E, Dey ES, Novik G. 2010. Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. Polish Journal of Microbiology 59:145-155.
27. Mahony J, Ainsworth S, Stockdale S, van Sinderen D. 2012. Phages of lactic acid bacteria: the role of genetics in understanding phage-host interactions and their co-evolutionary processes. Virology 434:143-150.
28. São-José C, Baptista C, Santos MA. 2004. *Bacillus subtilis* operon encoding a membrane receptor for bacteriophage SPP1. Journal of bacteriology 186:8337-8346.
29. São-José C, Lhuillier S, Lurz R, Melki R, Lepault J, Santos MA, Tavares P. 2006. The ectodomain of the viral receptor YueB forms a fiber that triggers ejection of bacteriophage SPP1 DNA. Journal of biological chemistry 281:11464-11470.
30. Valyasevi R, Sandine W, Geller B. 1991. A membrane protein is required for bacteriophage c2 infection of *Lactococcus lactis* subsp. *lactis* C2. Journal of bacteriology 173:6095-6100.

31. Monteville MR, Ardestani B, Geller BL. 1994. Lactococcal bacteriophages require a host cell wall carbohydrate and a plasma membrane protein for adsorption and ejection of DNA. Applied and environmental microbiology 60:3204-3211.
32. Ainsworth S, Sadovskaya I, Vinogradov E, Courtin P, Guerardel Y, Mahony J, Grard T, Cambillau C, Chapot-Chartier MP, Sinderen D. 2014. Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. MBio 5:e00880-14.
33. Legrand P, Collins B, Blangy S, Murphy J, Spinelli S, Gutierrez C, Richet N, Kellenberger C, Desmyter A, Mahony J. 2016. The atomic structure of the phage Tuc2009 baseplate tripod suggests that host recognition involves two different carbohydrate binding modules. MBio 7:e01781-15.
34. Chapot-Chartier MP, Vinogradov E, Sadovskaya I, Andre G, Mistou MY, Trieu-Cuot P, Furlan S, Bidnenko E, Courtin P, Pechoux C. 2010. Cell surface of *Lactococcus lactis* is covered by a protective polysaccharide pellicle. Journal of Biological Chemestry 285:10464-71.
35. Dupont K, Janzen T, Vogensen FK, Josephsen J, Stuer-Lauridsen B. 2004. Identification of *Lactococcus lactis* genes required for bacteriophage adsorption. Applied and environmental microbiology 70:5825-32.
36. Farenc C, Spinelli S, Vinogradov E, Tremblay D, Blangy S, Sadovskaya I, Moineau S, Cambillau C. 2014. Molecular insights on the recognition of a *Lactococcus lactis* cell wall pellicle by the phage 1358 receptor binding protein. Journal of virology 88:7005-15.
37. Mahony J, Randazzo W, Neve H, Settanni L, Sinderen D. 2015. Lactococcal 949 group phages recognize a carbohydrate receptor on the host cell surface. Applied and environmental microbiology 81:3299-305.
38. Andre G, Kulakauskas S, Chapot-Chartier M-P, Navet B, Deghorain M, Bernard E, Hols P, Dufrêne YF. 2010. Imaging the nanoscale organization of peptidoglycan in living *Lactococcus lactis* cells. Nature communications 1:27.
39. Vinogradov E, Sadovskaya I, Courtin P, Kulakauskas S, Grard T, Mahony J, van Sinderen D, Chapot-Chartier M-P. 2018. Determination of the cell wall polysaccharide and teichoic acid structures from *Lactococcus lactis* IL1403. Carbohydrate research 462:39-44.

40. Vinogradov E, Sadovskaya I, Grard T, Murphy J, Mahony J, Chapot-Chartier M-P, van Sinderen D. 2018. Structural studies of the cell wall polysaccharide from *Lactococcus lactis* UC509.9. Carbohydrate research 461:25-31.
41. Breum SØ, Neve H, Heller KJ, Vogensen FK. 2007. Temperate phages TP901-1 and φLC3, belonging to the P335 species, apparently use different pathways for DNA injection in *Lactococcus lactis* subsp. *cremoris* 3107. FEMS microbiology letters 276:156-164.
42. Spinelli S, Campanacci V, Blangy S, Moineau S, Tegoni M, Cambillau C. 2006. Modular structure of the receptor binding proteins of *Lactococcus lactis* phages: the RBP structure of the temperate phage TP901-1. Journal of Biological Chemistry 281:14256-14262.
43. Koch B, Christiansen B, Evison T, Vogensen FK, Hammer K. 1997. Construction of specific erythromycin resistance mutations in the temperate lactococcal bacteriophage TP901-1 and their use in studies of phage biology. Applied and environmental microbiology 63:2439-2441.
44. Braun Jr V, Hertwig S, Neve H, Geis A, Teuber M. 1989. Taxonomic differentiation of bacteriophages of *Lactococcus lactis* by electron microscopy, DNA-DNA hybridization, and protein profiles. Microbiology 135:2551-2560.
45. Kuipers OP, de Ruyter PG, Kleerebezem M, de Vos WM. 1998. Quorum sensing-controlled gene expression in lactic acid bacteria. Journal of biotechnology 64:15-21.
46. Bolotin A, Wincker P, Mauger S, Jaillon O, Malarme K, Weissenbach J, Ehrlich SD, Sorokin A. 2001. The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. Genome research 11:731-53.
47. Sambrook J, Russell DW. 2006. Transformation of *E. coli* by electroporation. Cold spring harbor protocols 2006:pdb. prot3933.
48. Christiansen B, Johnsen M, Stenby E, Vogensen F, Hammer K. 1994. Characterization of the lactococcal temperate phage TP901-1 and its site-specific integration. Journal of bacteriology 176:1069-1076.
49. Lillehaug D, Lindqvist B, Birkeland N. 1991. Characterization of phiLC3, a *Lactococcus lactis* subsp. *cremoris* temperature bacteriophage with cohesive single-stranded DNA ends. Applied and environmental microbiology 57:3206-3211.

50. Oliveira J. 2017. Biodiversity of lactococcal bacteriophages and associated phage-host interactions: PhD thesis, University College Cork.
51. Mahony J, Deveau H, Mc Grath S, Ventura M, Canchaya C, Moineau S, Fitzgerald GF, van Sinderen D. 2006. Sequence and comparative genomic analysis of lactococcal bacteriophages jj50, 712 and P008: evolutionary insights into the 936 phage species. FEMS Microbiology letters 261:253-61.
52. Tremblay DM, Tegoni M, Spinelli S, Campanacci V, Blangy S, Huyghe C, Desmyter A, Labrie S, Moineau S, Cambillau C. 2006. Receptor-binding protein of *Lactococcus lactis* phages: identification and characterization of the saccharide receptor-binding site. Journal of bacteriology 188:2400-10.
53. Valyasevi R, Sandine WE, Geller BL. 1994. *Lactococcus lactis* ssp. *lactis* C2 bacteriophage sk1 receptor involving rhamnose and glucose moieties in the cell wall. Journal of dairy science 77:1-6.
54. Villion M, Chopin M-C, Deveau H, Ehrlich SD, Moineau S, Chopin A. 2009. P087, a lactococcal phage with a morphogenesis module similar to an *Enterococcus faecalis* prophage. Virology 388:49-56.
55. Samson JE, Moineau S. 2010. Characterization of *Lactococcus lactis* phage 949 and comparison with other lactococcal phages. Applied and environmental microbiology 76:6843-6852.
56. Cavanagh D, Guinane CM, Neve H, Coffey A, Ross RP, Fitzgerald GF, McAuliffe O. 2013. Phages of non-dairy lactococci: isolation and characterization of ΦL47, a phage infecting the grass isolate *Lactococcus lactis* ssp. *cremoris* DPC6860. Frontiers in microbiology 4 p.417.
57. Mahony J, Moscarelli A, Kelleher P, Lugli GA, Ventura M, Settanni L, van Sinderen D. 2017. Phage biodiversity in artisanal cheese wheys reflects the complexity of the fermentation process. Viruses 9:45.
58. van Pijkeren J-P, Britton RA. 2012. High efficiency recombineering in lactic acid bacteria. Nucleic acids research 40:e76-e76.
59. McGrath S, Fitzgerald GF, van Sinderen D. 2001. Improvement and optimization of two engineered phage resistance mechanisms in *Lactococcus lactis*. Applied and environmental microbiology 67:608-616.
60. Gupta PK. 2008. Single-molecule DNA sequencing technologies for future genomics research. Trends in biotechnology 26:602-611.

61. McCarthy A. 2010. Third generation DNA sequencing: pacific biosciences' single molecule real time technology. Chemistry & biology 17:675-676.
62. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and environmental microbiology: 79 (17) 5112-5120.
63. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10:R25.
64. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079.
65. Sonnhammer EL, Von Heijne G, Krogh A. A hidden Markov model for predicting transmembrane helices in protein sequences, p 175-182. *In* (ed),
66. Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic acids research 33:W244-W248.
67. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream M-A, Barrell B. 2000. Artemis: sequence visualization and annotation. Bioinformatics 16:944-945.
68. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. Journal of molecular biology 215:403-410.
69. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold spring harbor laboratory press.
70. Swindell SR, Plasterer TN. 1997. SeqMan: Contig assembly. Sequence data analysis guidebook:75-89.
71. Stockdale SR, Mahony J, Courtin P, Chapot-Chartier MP, van Pijkeren JP, Britton RA, Neve H, Heller KJ, Aideh B, Vogensen FK, van Sinderen D. 2013. The lactococcal phages Tuc2009 and TP901-1 incorporate two alternate forms of their tail fiber into their virions for infection specialization. Journal of biological chemistry 288:5581-90.
72. van Pijkeren J-P, Neoh KM, Sirias D, Findley AS, Britton RA. 2012. Exploring optimization parameters to increase ssDNA recombineering in *Lactococcus lactis* and *Lactobacillus reuteri*. Bioengineered 3:209-217.

73. Landete JM, Arqués JL, Peirotén Á, Langa S, Medina M. 2014. An improved method for the electrotransformation of lactic acid bacteria: a comparative survey. Journal of microbiological methods 105:130-133.
74. Lillehaug D. 1997. An improved plaque assay for poor plaque-producing temperate lactococcal bacteriophages. Journal of applied microbiology 83:85-90.
75. Bolotin A, Wincker P, Mauger S, Jaillon O, Malarme K, Weissenbach J, Ehrlich SD, Sorokin A. 2001. The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. Genome research 11:731-753.
76. Liu J, Mushegian A. 2003. Three monophyletic superfamilies account for the majority of the known glycosyltransferases. Protein science 12:1418-1431.
77. Rismondo J, Percy MG, Gründling A. 2018. Discovery of genes required for lipoteichoic acid glycosylation predicts two distinct mechanisms for wall teichoic acid glycosylation. Journal of biological chemistry 293:3293-3306.
78. van der Meulen SB, de Jong A, Kok J. 2016. Transcriptome landscape of *Lactococcus lactis* reveals many novel RNAs including a small regulatory RNA involved in carbon uptake and metabolism. RNA biology 13:353-366.
79. Wedekind JE, Frey PA, Rayment I. 1995. Three-dimensional structure of galactose-1-phosphate uridylyltransferase from *Escherichia coli* at 1.8. ANG. resolution. Biochemistry 34:11049-11061.
80. Cumby N, Reimer K, Mengin-Lecreulx D, Davidson AR, Maxwell KL. 2015. The phage tail tape measure protein, an inner membrane protein and a periplasmic chaperone play connected roles in the genome injection process of *E. coli* phage HK97. Molecular microbiology 96:437-447.
81. Vegge CS, Vogensen FK, Mc Grath S, Neve H, van Sinderen D, Brøndsted L. 2006. Identification of the lower baseplate protein as the antireceptor of the temperate lactococcal bacteriophages TP901-1 and Tuc2009. Journal of bacteriology 188:55-63.
82. Rousseau GM, Moineau S. 2009. Evolution of *Lactococcus lactis* phages within a cheese factory. Applied and environmental microbiology 75:5336-5344.
83. Vegge CS, Brøndsted L, Neve H, Mc Grath S, van Sinderen D, Vogensen FK. 2005. Structural characterization and assembly of the distal tail structure of the

temperate lactococcal bacteriophage TP901-1. Journal of bacteriology 187:4187-4197.

84. Veesler D, Spinelli S, Mahony J, Lichiere J, Blangy S, Bricogne G, Legrand P, Ortiz-Lombardia M, Campanacci V, van Sinderen D, Cambillau C. 2012. Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. Proceedings of the national academy of sciences of the United States of America 109:8954-8.

Table S3.1: Confirmed SNPs with proposed (or confirmed) involvement in TP901-1 phage insensitivity in derived bacteriophage insensitive mutants (BIMs).

SNP	BIM	SNP coordinates	Base modification	Location *(orf)*	Predicted function of product (BLAST)	Primers used for cloning (5' to 3')
1	E126	40809	c to a	Intergenic	n/a	
2	E119	47852	g to a	L3107_0037	acetoin dehydrogenase complex E1 component alpha subunit	
3	E119	78801	g to a	L3107_0070	ATPase component of ABC transporter with duplicated ATPase domains	
4	E121	112463	g to a	L3107_0104	argininosuccinate synthase	104_F AGCAGCGGTACCAGGAGGCACTCACATGATGGGAAACAAAAAA 104_R AGCAGCTCTAGATTATTTGAGATTTACTTG
5	E119	117593	c to t	L3107_0109	Preprotein translocase subunit YidC	109_F AGCAGCCTGCAGAGGAGGCACTCACTTGAAAAAGAAATTTAGT 109_R AGCAGCTCTAGATTATTTTTTTCTTGCCTTC
6	E119	139073	a to t	Intergenic	n/a	
7	E121	181584	t to a	L3107_0174	hypothetical protein	174_F AGCAGCCTGCAGAGGAGGCACTCACATGAAAATGAGTGTTATT 174_R AGCAGCTCTAGACTAATCCTCAACCACCTTC
8	E119	270693	g to a	L3107_0263	hypothetical protein	263_F AGCAGCCTGCAGAGGAGGCACTCACATGAAAATAATAAAAAAC 263_R AGCAGCTCTAGACTATTCGAGGTTTTTGCT
9	E119	291322	g to a	L3107_0286	hypothetical protein	286_F AGCAGCCTGCAGAGGAGGCACTCACATGGCAAAAAGACAGAACC 286_R AGCAGCTCTAGATTAGCCGAAGTTAATGGTG
10	E119	317385	c to t	Intergenic	n/a	
11	E119	393093	c to t	L3107_0378	C4-dicarboxylate transporter / malic acid transporter	378_F F AGCAGCCTGCAGAGGAGGCACTCACATGATCTATTCAAAATTT 378_R AGCAGCTCTAGATTAGATGCTTTCAGCAGA
12	E119	475194	g to a	L3107_0457	hypothetical protein	457_F AGCAGCCTGCAGAGGAGGCACTCACATGAAAATAATTGGAAAT 457_R AGCAGCTCTAGATTAGAGCGTCTTTTTATA
13	E119	480099	c to t	L3107_0464	Predicted Rossmann fold nucleotide-binding protein	464_F F AGCAGCCTGCAGAGGAGGCACTCACATGAATGTTACTGTCTTC 464_R AGCAGCTCTAGATTATTTTGAATAACAGAA

14	E119	480967	g to a	Intergenic	n/a	
15	E126	484073	g to a	L3107_0468	DNA polymerase III catalytic subunit DnaE type	
16	E119	527927	c to t	L3107_0505	ATP-binding subunit of Clp protease and DnaK / DnaJ chaperones	
17	E119	593696	g to a	Intergenic	n/a	
18	E119	636929	c to t	L3107_0616	Branched-chain amino acid permease	
19	E119	649840	c to t	L3107_0630	glycogen synthase (ADP-glucose)	630_F AGCAGCCTGCAGAGGAGGCACTCACATGAAAGTTTTATTTGCG 630_R AGCAGCTCTAGATTACTTTAATATTGTTTG
20	E119	652014	g to a	L3107_0631	Glucan phosphorylase	
21	E126	741347	c to t	L3107_0715	3-oxoacyl-[acyl-carrier-protein] synthase III	715_F AGCAGCCTGCAGAGGAGGCACTCACATGACTTTTGCGAAAATC 715_R AGCAGCTCTAGATTATAAATTAATAATTGC
22	E121	869757	c to t	L3107_0866	hypothetical protein	866_F AGCAGCGGTACCAGGAGGCACTCACATGTCAAAACAAAAAACA 866_R AGCAGCTCTAGATTATACACCCGTAAATGT
23	E119	884291	t to c	L3107_0885	hypothetical protein	885_F AGCAGCCTGCAGAGGAGGCACTCACATGAGTTTAGATAATTTC 885_R AGCAGCTCTAGATTATGCAGTCCGTTGCCA
24	E119 / E126	962789	c to t	Intergenic	n/a	
25	E121	973425	c to t	L3107_0976	Multidrug resistance protein B MF superfamily	0976_F AGCAGCCTGCAGAGGAGGCACTCACATGGTTCCTATTTCAGCT 0976_R AGCAGCTCTAGATTATTTATTTTCTCCTTTC
26	E121	993617	a to t	Intergenic	Phosphopentomutase	SNP26_F AGCAGCCTGCAGAGGAGGCACTCACAAAGATAAAAATCATTGGTC 1001_R AGCAGCTCTAGATTAAACTAACGCATCCAAGA
27	E119 / E126	1132674	t to c	L3107_1143	Alpha-acetolactate decarboxylase	
28	E119 / E126	1133202	t to g	Intergenic	n/a	
29	E126	1158710	g to t	L3107_1163	DNA-3-methyladenine glycosylase I	1163_F AGCAGCCTGCAGAGGAGGCACTCACATGACTGAAAAAACAAGA 1163_R AGCAGCTCTAGATCAATCCATTAATTCTAC
30	E126	1169995	c to t	L3107_1173	FAD synthase	
31	E121	1193427	c to t	L3107_1195	Xanthine phosphoribosyltransferase	1195_F AGCAGCCTGCAGAGGAGGCACTCACATGAAGTTATTAGAAGAC 1195_R AGCAGCTCTAGATTAATCGAAAGCGATATC
32	E119	1203781	g to a	L3107_1208	Deoxynucleoside kinase	1208_F AGCAGCCTGCAGAGGAGGCACTCACTTGATCGTATTAGCAGGT 1208_R AGCAGCTCTAGATTAAAGTAAATCAAGTTT

33	E119	1262123	c to t	L3107_1265	3-isopropylmalate dehydrogenase	
34	E126	1272700	g to a	L3107_1276	threonine dehydratase	
35	E126	1310588	a to t	Intergenic	n/a	
36	E121	1311194	g to a	L3107_1311	hypothetical protein	1311_F AGCAGCCTGCAGAGGAGGCACTCACATGATACTTTTCATTATT 1311_R AGCAGCTCTAGATTAATCCTTGATTCCAAG
37	E121	1392004	c to t	L3107_1382	N-acetylglucosamine 6-phosphate deacetylase	1382_F AGCAGCCTGCAGAGGAGGCACTCACATGACTTACTATATTAAAGC 1382_R AGCAGCTCTAGATTAAGCTTCGTAACGTTTTTCAC
38	E119	1425572	g to a	L3107_1417	hypothetical protein	
39	E126	1452747	c to t	*L3107_1442	Glycosyltransferase	1442_F AGCAGCCTGCAGAGGAGGCACTCACTTGTATCTTTGGTATAATAG
40	E119	1453653				1442_R AGCAGCGGTACCTTAGCTTTCATTTTGCTGAC
41	E126	1559076	g to t	L3107_1546	Surface antigen	
42	E119	1680258	g to a	L3107_1670	Phosphotransferase system mannose/fructose-specific component IIA	
43	E119	1703015	a to t	L3107_1694	Predicted O-methyltransferase	
44	E119	1734506	c to t	L3107_1726	3-dehydroquinate synthase	
45	E119	1735828	c to t	L3107_1727	shikimate dehydrogenase	1727_F AGCAGCCTGCAGAGGAGGCACTCACATGAAAATTGACGGTTAC 1727_R AGCAGCGGTACCTCATACTTCTCTCTCCAT
46	E119	1777955	g to a	L3107_1773	Polar amino acid ABC transporter permease protein	1773_F AGCAGCCTGCAGAGGAGGCACTCACATGAATTTTGATTTTTTACC 1773_R AGCAGCTCTAGATTATTTCGCATATCCTTTCC
47	E126	1785474	c to a	L3107_1776	DNA-directed RNA polymerase subunit beta	
48	E126	1789091	g to a	L3107_1779	Oligoendopeptidase O	
49	E121	1833247	g to a	L3107_1821	ABC-type dipeptide/oligopeptide/nickel transport system ATPase component	1821_F AGCAGCCTGCAGAGGAGGCACTCACATGTACGAAAAAAGATACG 1821_R AGCAGCTCTAGATCACCTCTCACCTCCTTTA
50	E119 / E126	1900069	t to c	L3107_1891	DNA polymerase IV	
51	E119 / E126	1900084	a to g	L3107_1891	DNA polymerase IV	
52	E119 / E126	1900726	a to t	L3107_1891	DNA polymerase IV	
53	E119	1946818	c to t	L3107_1940	hypothetical protein	1940_F AGCAGCCTGCAGAGGAGGCACTCACATGGAAAAATTGATTTCT 1940_R AGCAGCTCTAGACTATCTTTTTCTCATCAT
54	E119	1957802	c to t	L3107_1955	head-tail joining protein	

55	E126	1996872	g to a	L3107_1994	Predicted membrane protein	1994_F AGCAGCCTGCAGAGGAGGCACTCACATGGAAGTGATAAGCAGT 1994_R AGCAGCTCTAGATTAATATTCTATTCTTGG
56	E119	2017636	c to t	L3107_2014	phenylalanyl-tRNA synthetase beta subunit	
57	E121	2073132	c to t	L3107_2071	Phage DNA polymerase (ATPase domain)	2071_F AGCAGCCTGCAGAGGAGGCACTCACATGACAGACCAATTAGAT 2071_R AGCAGCTCTAGATCATTTTCGCTTTTTCTT
58	E119 / E126	2081538	t to c	L3107_2086	hypothetical protein	2086_F AGCAGCCTGCAGAGGAGGCACTCACATGCCTACACAGGAAATA 2086_R AGCAGCTCTAGACTATTTTCGCCATAGTCC
59	E126	2176455	g to a	L3107_2192	30S ribosomal protein S5	
60	E121	2197193	g to a	Intergenic	Ribosomal large subunit pseudourine synthase	SNP60_F AGCAGCCTGCAGAGGAGGCACTCACTATCCGTTCATCATTGGTAAG 2223_R AGCAGCTCTAGATTATAATAGTTGCTCTACTTC
61	E121	2205536	c to a	L3107_2235	Competence protein ComGB	2235_F AGCAGCCTGCAGAGGAGGCACTCACATGGGGAATCTTTTGAAT 2235_R AGCAGCTCTAGATTAAATTCCAGAACCCAT
62	E121	2228548	a to g	L3107_2253	DNA polymerase I	2253_F AGCAGCCTGCAGAGGAGGCACTCACATGGAAGATAAGAATAGA 2253_R AGCAGCGGTACCTCATTTGGCTTCATACCA
63	E121	2274649	c to t	L3107_2293	recombination regulator RecX	2293_F AGCAGCCTGCAGAGGAGGCACTCACATGGGAAAAATCACAGCA 2293_R AGCAGCGGTACCTTATTCTGGGAATGTGTA
64	E126	2325916	c to t	L3107_2342	Activator of 2-hydroxyglutaryl-CoA dehydratase and two uncharacterized domains	
65	E121	2340790	g to a	Intergenic	ATP-dependent DNA helicase RecG.	SNP65_F AGCAGCCTGCAGAGGAGGCACTCACAAAGTAATCAGAAAATTTTCT 2352_R AGCAGCTCTAGATCAGTCAAATCCACCCTCAAC

*The cloning of the *L3107_1442* gene (renamed *gtfA* in this study) complemented the E119 and E126 mutants.

Table S3.2: Confirmed SNPs with proposed involvement in TP901-1 phage infection of bacteriophage insensitive mutants (BIMs).

SNP	TP901-1*erm* mutant	*SNP coordinates	Base modification			Location (*orf*)	Predicted function of product (Blast)
1	E119B/E121A/E126A	325	a	to	g	ORF31	large terminase subunit (terL)
2	E119B/E121A/E126A	326	t	to	c	ORF31	large terminase subunit (terL)
3	E119B/E121A/E126A	343	g	to	a	ORF31	large terminase subunit (terL)
4	E119B/E121A/E126A	364	c	to	t	ORF31	large terminase subunit (terL)
5	E119B/E121A/E126A	415	t	to	g	ORF31	large terminase subunit (terL)
6	E119B/E121A/E126A	466	c	to	t	ORF31	large terminase subunit (terL)
7	E119B/E121A/E126A	469	a	to	g	ORF31	large terminase subunit (terL)
8	E119B/E121A/E126A	496	a	to	g	ORF31	large terminase subunit (terL)
9	E119B/E121A/E126A	514	t	to	c	ORF31	large terminase subunit (terL)
10	E119B/E121A/E126A	524	c	to	a	ORF31	large terminase subunit (terL)
11	E119B/E121A/E126A	526	g	to	a	ORF31	large terminase subunit (terL)
12	E119B/E121A/E126A	538	a	to	t	ORF31	large terminase subunit (terL)
13	E119A/E119B/E121A/E126A	1771	g	to	a	ORF32	portal protein
14	E119A/E119B/E121A/E126A	1777	a	to	g	ORF32	portal protein
15	E119A/E119B/E121A/E126A	1780	c	to	t	ORF32	portal protein
16	E119A/E119B/E121A/E126A	1781	t	to	g	ORF32	portal protein

17	E119A/E119B/E121A/E126A	1810	a	to	t	ORF32	portal protein
18	E119A/E119B/E121A/E126A	1834	g	to	a	ORF32	portal protein
19	E119A/E119B/E121A/E126A	1837	a	to	g	ORF32	portal protein
20	E119A/E119B/E121A/E126A	1843	a	to	c	ORF32	portal protein
21	E119A/E119B/E121A/E126A	1858	c	to	t	ORF32	portal protein
22	E119A/E119B/E121A/E126A	1870	c	to	t	ORF32	portal protein
23	E119A/E119B/E121A/E126A	1885	t	to	a	ORF32	portal protein
24	E119A/E119B/E121A/E126A	1891	a	to	g	ORF32	portal protein
25	E119A/E119B/E121A/E126A	1909	c	to	t	ORF32	portal protein
26	E119A/E119B/E121A/E126A	1918	t	to	c	ORF32	portal protein
27	E119A/E119B/E121A/E126A	1973	a	to	c	ORF32	portal protein
28	E119A/E119B/E121A/E126A	2189	t	to	g	ORF33	minor capsid protein 1 (MCP1)
29	E119A/E119B/E121A/E126A	2284	t	to	a	ORF33	minor capsid protein 1 (MCP1)
30	E119A/E119B/E121A/E126A	2285	c	to	g	ORF33	minor capsid protein 1 (MCP1)
31	E119A/E119B/E121A/E126A	2303	a	to	c	ORF33	minor capsid protein 1 (MCP1)
32	E119A/E119B/E121A/E126A	2377	t	to	c	ORF33	minor capsid protein 1 (MCP1)
33	E119A/E119B/E121A/E126A	2384	c	to	t	ORF33	minor capsid protein 1 (MCP1)
34	E119A/E119B/E121A/E126A	2386	g	to	a	ORF33	minor capsid protein 1 (MCP1)

35	E119A/E119B/E121A/E126A	2389	g	to	a	ORF33	minor capsid protein 1 (MCP1)
36	E119A/E119B/E121A/E126A	2390	c	to	t	ORF33	minor capsid protein 1 (MCP1)
37	E119A/E119B/E121A/E126A	2392	t	to	g	ORF33	minor capsid protein 1 (MCP1)
38	E119A/E119B/E121A/E126A	2396	t	to	g	ORF33	minor capsid protein 1 (MCP1)
39	E119A/E119B/E121A/E126A	2407	t	to	c	ORF33	minor capsid protein 1 (MCP1)
40	E119A/E119B/E121A/E126A	2410	c	to	t	ORF33	minor capsid protein 1 (MCP1)
41	E119A/E119B/E121A/E126A	2431	t	to	g	ORF33	minor capsid protein 1 (MCP1)
42	E119A/E119B/E121A/E126A	2437	a	to	g	ORF33	minor capsid protein 1 (MCP1)
43	E119A/E119B/E121A/E126A	2441	t	to	g	ORF33	minor capsid protein 1 (MCP1)
44	E119A/E119B/E121A/E126A	2442	c	to	g	ORF33	minor capsid protein 1 (MCP1)
45	E119A/E119B/E121A/E126A	2453	a	to	g	ORF33	minor capsid protein 1 (MCP1)
46	E119A/E119B/E121A/E126A	2476	a	to	c	ORF33	minor capsid protein 1 (MCP1)
47	E119A/E119B/E121A/E126A	2479	t	to	a	ORF33	minor capsid protein 1 (MCP1)
48	E119A/E119B/E121A/E126A	2485	t	to	c	ORF33	minor capsid protein 1 (MCP1)
49	E119A/E119B/E121A/E126A	2486	t	to	c	ORF33	minor capsid protein 1 (MCP1)
50	E119A/E119B/E121A/E126A	2491	c	to	t	ORF33	minor capsid protein 1 (MCP1)
51	E119A/E119B/E121A/E126A	2497	a	to	c	ORF33	minor capsid protein 1 (MCP1)
52	E119A/E119B/E121A/E126A	2501	g	to	a	ORF33	minor capsid protein 1 (MCP1)

53	E119A/E119B/E121A/E126A	2503	a	to	c	ORF33	minor capsid protein 1 (MCP1)
54	E119A/E119B/E121A/E126A	2506	a	to	t	ORF33	minor capsid protein 1 (MCP1)
55	E119A/E119B/E121A/E126A	2530	g	to	t	ORF33	minor capsid protein 1 (MCP1)
56	E119A/E119B/E121A/E126A	2533	a	to	t	ORF33	minor capsid protein 1 (MCP1)
57	E119A/E119B/E121A/E126A	2578	a	to	g	ORF33	minor capsid protein 1 (MCP1)
58	E119A/E119B/E121A/E121B/E126A	2660	a	to	g	ORF33	minor capsid protein 1 (MCP1)
59	E119A/E119B/E121A/E121B/E126A	2698	g	to	a	ORF33	minor capsid protein 1 (MCP1)
60	E119A/E119B/E121A/E126A	2716	a	to	t	ORF33	minor capsid protein 1 (MCP1)
61	E119A/E119B/E121A/E126A	2722	c	to	t	ORF33	minor capsid protein 1 (MCP1)
62	E119A/E119B/E121A/E126A	2725	a	to	t	ORF33	minor capsid protein 1 (MCP1)
63	E119A/E119B/E121A/E121B/E126A	2737	t	to	a	ORF33	minor capsid protein 1 (MCP1)
64	E119A/E119B/E121A/E121B/E126A	2740	g	to	t	ORF33	minor capsid protein 1 (MCP1)
65	E119A/E119B/E121A/E121B/E126A	2743	g	to	a	ORF33	minor capsid protein 1 (MCP1)
66	E119A/E119B/E121A/E121B/E126A	2920	t	to	c	ORF33	minor capsid protein 1 (MCP1)
67	E119A/E119B/E121A/E121B/E126A	2929	t	to	c	ORF33	minor capsid protein 1 (MCP1)
68	E119A/E119B/E121A/E121B/E126A	2938	g	to	a	ORF33	minor capsid protein 1 (MCP1)
69	E119A/E119B/E121A/E121B/E126A	2958	c	to	g	ORF33	minor capsid protein 1 (MCP1)
70	E119A	2994	a	to	c	ORF33	minor capsid protein 1 (MCP1)

71	E119A	2995	a	to	g	ORF33	minor capsid protein 1 (MCP1)
72	E119A/E119B/E121A/E121B/E126A	3955	g	to	c	ORF34	hypothetical protein
73	E119A/E119B/E121A/E121B/E126A	3956	a	to	c	ORF34	hypothetical protein
74	E119A/E119B/E121A/E121B/E126A	3958	a	to	g	ORF34	hypothetical protein
75	E119A/E119B/E121A/E121B/E126A	3959	a	to	g	ORF34	hypothetical protein
76	E119A/E119B/E121A/E121B/E126A	3961	t	to	a	ORF34	hypothetical protein
77	E119A/E119B/E121A/E121B/E126A	3962	a	to	c	ORF34	hypothetical protein
78	E119A/E119B/E121A/E121B/E126A	3966	t	to	a	ORF34	hypothetical protein
79	E119A/E119B/E121A/E121B/E126A	3969	a	to	c	ORF34	hypothetical protein
80	E119A/E119B/E121A/E121B/E126A	3972	t	to	c	ORF34	hypothetical protein
81	E119A/E119B/E121A/E121B/E126A	3988	t	to	g	Intergenic	n/a
82	E119A/E119B/E121A/E121B/E126A	4146	a	to	c	ORF35	scaffolding protein (Sfp)
83	E119A/E119B/E121A/E121B/E126A	4163	c	to	t	ORF35	scaffolding protein (Sfp)
84	E119A/E119B/E121A/E121B/E126A	4170	g	to	a	ORF35	scaffolding protein (Sfp)
85	E119A/E119B/E121A/E121B/E126A	4200	g	to	t	ORF35	scaffolding protein (Sfp)
86	E119A/E119B/E121A/E121B/E126A	4312	t	to	g	ORF35	scaffolding protein (Sfp)
87	E119A/E119B/E121A/E121B/E126A	4464	t	to	c	ORF35	scaffolding protein (Sfp)
88	E119A/E119B/E121A/E121B/E126A	4470	g	to	a	ORF35	scaffolding protein (Sfp)

89	E119A/E119B/E121A/E121B/E126A	4494	g	to	a	ORF35	scaffolding protein (Sfp)
90	E119A/E119B/E121A/E121B/E126A	4548	c	to	t	ORF35	scaffolding protein (Sfp)
91	E119A/E119B/E121A/E121B/E126A	4617	a	to	g	ORF35	scaffolding protein (Sfp)
92	E119A/E119B/E121A/E121B/E126A	4778	g	to	a	ORF36	major head protein (Mhp)
93	E119A/E119B/E121A/E121B/E126A	4792	t	to	c	ORF36	major head protein (Mhp)
94	E119A/E119B/E121A/E121B/E126A	4841	a	to	c	ORF36	major head protein (Mhp)
95	E119A/E119B/E121A/E121B/E126A	4843	g	to	a	ORF36	major head protein (Mhp)
96	E119A/E119B/E121A/E121B/E126A	5020	g	to	t	ORF36	major head protein (Mhp)
97	E119A/E119B/E121A/E121B/E126A	5042	g	to	t	ORF36	major head protein (Mhp)
98	E119A/E119B/E121A/E121B/E126A	5043	c	to	t	ORF36	major head protein (Mhp)
99	E119A/E119B/E121A/E121B/E126A	5130	g	to	a	ORF36	major head protein (Mhp)
100	E119A/E119B/E121A/E121B/E126A	5131	c	to	a	ORF36	major head protein (Mhp)
101	E121A/E121B	5227	c	to	t	ORF36	major head protein (Mhp)
102	E119A/E119B	5290	c	to	t	ORF36	major head protein (Mhp)
103	E121A/E121B	5290	c	to	t	ORF36	major head protein (Mhp)
104	E119A/E119B	5384	g	to	a	ORF36	major head protein (Mhp)
105	E121A/E121B	5384	g	to	a	ORF36	major head protein (Mhp)
106	E119A/E119B	5385	g	to	a	ORF36	major head protein (Mhp)

107	E121A/E121B	5385	g	to	a	ORF36	major head protein (Mhp)
108	E119A/E119B	5396	a	to	t	ORF36	major head protein (Mhp)
109	E121A/E121B	5396	a	to	t	ORF36	major head protein (Mhp)
110	E119A/E119B	5608	a	to	g	ORF37	hypothetical protein
111	E121A/E121B	5608	a	to	g	ORF37	hypothetical protein
112	E119A/E119B/E121A	5725	t	to	g	ORF37	hypothetical protein
113	E119A/E119B/E121A	5726	c	to	a	ORF37	hypothetical protein
114	E119B/E121A	5768	g	to	t	ORF37	hypothetical protein
115	E119B/E121A	5773	t	to	a	ORF38	DNA packaging protein
116	E119B/E121A	5774	t	to	a	ORF38	DNA packaging protein
117	E119B/E121A	5781	t	to	a	ORF38	DNA packaging protein
118	E119B/E121A	5799	t	to	c	ORF38	DNA packaging protein
119	E119B/E121A	5816	g	to	a	ORF38	DNA packaging protein
120	E119B/E121A	5839	c	to	a	ORF38	DNA packaging protein
121	E121A	5867	a	to	c	ORF38	DNA packaging protein
122	E119B	12698	g	to	c	ORF47	tail-associated lysin (Tal)
123	E119A/E121A/E121B/E126A	13186	t	to	c	ORF47	tail-associated lysin (Tal)
124	E126B	14499	a	to	c	ORF47	tail-associated lysin (Tal)

125	E126B	14526	t	to	a	ORF47	tail-associated lysin (Tal)
126	E126B	14527	g	to	a	ORF47	tail-associated lysin (Tal)
127	E126B	14528	g	to	a	ORF47	tail-associated lysin (Tal)
128	E126B	14566	g	to	a	ORF47	tail-associated lysin (Tal)
129	E119B	34861	c	to	a	ORF22	hypothetical protein
130	E119B	34878	a	to	c	ORF22	hypothetical protein
131	E119B	34982	a	to	g	ORF22	hypothetical protein
132	E119B	35755	a	to	c	ORF22	hypothetical protein
133	E119B	35793	g	to	a	ORF22	hypothetical protein
134	E119B	35808	a	to	g	ORF22	hypothetical protein
135	E119B	35811	a	to	g	ORF22	hypothetical protein
136	E119B	35819	a	to	t	ORF22	hypothetical protein
137	E119B	35918	t	to	c	Intergenic	n/a
138	E119B	35969	c	to	t	Intergenic	n/a
139	E119B	36013	g	to	t	Intergenic	n/a
140	E119B	36027	a	to	g	Intergenic	n/a
141	E119B	36065	t	to	g	Intergenic	n/a
142	E119B	36068	t	to	c	Intergenic	n/a

143	E119B	36071	t	to	a	Intergenic	n/a
144	E119B	36687	t	to	c	Intergenic	n/a
145	E119B	36722	a	to	c	Intergenic	n/a
146	E119B	36727	g	to	a	Intergenic	n/a
147	E119B	36729	c	to	a	Intergenic	n/a
148	E119B	36742	t	to	g	Intergenic	n/a
149	E119B	36821	c	to	t	ORF29	activator of late transcription
150	E119B	36857	c	to	t	ORF29	activator of late transcription
151	E119B	36866	a	to	g	ORF29	activator of late transcription
152	E119B	36887	g	to	a	ORF29	activator of late transcription
153	E119B	36917	g	to	a	ORF29	activator of late transcription
154	E119B	36937	a	to	t	ORF29	activator of late transcription
155	E119B	36938	c	to	a	ORF29	activator of late transcription
156	E119B	36943	a	to	g	ORF29	activator of late transcription
157	E119B	36945	c	to	t	ORF29	activator of late transcription
158	E119B	36947	c	to	a	ORF29	activator of late transcription
159	E119B	36949	a	to	g	ORF29	activator of late transcription
160	E119B	36953	c	to	a	ORF29	activator of late transcription

161	E119B	36954	c	to	t	ORF29	activator of late transcription
162	E119B	36959	g	to	a	ORF29	activator of late transcription
163	E119B	36960	t	to	a	ORF29	activator of late transcription
164	E119B	36961	a	to	g	ORF29	activator of late transcription
165	E119B	36962	c	to	a	ORF29	activator of late transcription
166	E119B	36966	t	to	a	ORF29	activator of late transcription
167	E119B	36969	a	to	g	ORF29	activator of late transcription
168	E119B	36974	a	to	c	ORF29	activator of late transcription
169	E119B	38237	g	to	c	ORF31	large terminase subunit (terL)
170	E119B	38238	t	to	a	ORF31	large terminase subunit (terL)
171	E119B	38241	t	to	c	ORF31	large terminase subunit (terL)
172	E119B	38244	g	to	a	ORF31	large terminase subunit (terL)
173	E119B	38259	t	to	g	ORF31	large terminase subunit (terL)
174	E119B	38271	g	to	a	ORF31	large terminase subunit (terL)
175	E119B	38293	a	to	g	ORF31	large terminase subunit (terL)
176	E119B	38298	g	to	a	ORF31	large terminase subunit (terL)
177	E119B	38299	c	to	t	ORF31	large terminase subunit (terL)
178	E119B	38301	t	to	c	ORF31	large terminase subunit (terL)

179	E119B/E121A/E126A	39183	a	to	g	ORF31	large terminase subunit (terL)
180	E119B/E121A/E126A	39184	t	to	c	ORF31	large terminase subunit (terL)
181	E119B/E121A/E126A	39201	g	to	a	ORF31	large terminase subunit (terL)
182	E119B/E121A/E126A	39222	c	to	t	ORF31	large terminase subunit (terL)
183	E119B/E121A/E126A	39273	t	to	g	ORF31	large terminase subunit (terL)
184	E119B/E121A/E126A	39324	c	to	t	ORF31	large terminase subunit (terL)
185	E119B/E121A/E126A	39327	a	to	g	ORF31	large terminase subunit (terL)
186	E119B/E121A/E126A	39354	a	to	g	ORF31	large terminase subunit (terL)
187	E119B/E121A/E126A	39372	t	to	c	ORF31	large terminase subunit (terL)
188	E119B/E121A/E126A	39382	c	to	a	ORF31	large terminase subunit (terL)
189	E119B/E121A/E126A	39384	g	to	a	ORF31	large terminase subunit (terL)
190	E119B/E121A/E126A	39396	a	to	t	ORF31	large terminase subunit (terL)

*These SNP coordinates exhibit the position of the mutation on the Illumina sequenced TP901-1*erm* WT.

Chapter IV

Biochemical analysis of cell wall polysaccharides and teichoic acids in *Lactococcus lactis* subsp. *cremoris* strains 3107 and NZ9000, and derivatives.

Note: Cell wall polysaccharide (CWPS) analysis was performed by Dr. Marie-Pierre Chapot-Chartier, Dr. Simon Palussière and Dr. Pascal Courtin of Micalis Institute, INRA, Jouy-en-Josas, France. Teichoic acid analysis was performed by Dr. Irina Sadovskaya of the Université du Littoral Côte d'Opale, Boulogne-sur-Mer, France and Dr. Evgeny Vinogradov of the Institute of Biological Sciences, Ottawa, ON, Canada.

Chapter IV contents

4 Abstract

The predicted membrane-associated glycosyltransferase (GTF), $GtfA_{3107}$, is required for one of the first host-interaction steps of certain members of the lactococcal P335 and 949 phage groups (see Chapter III). GtfA is presumed to glycosylate a substrate, which is likely to be a cell envelope-associated glycopolymer, such as cell wall polysaccharides (CWPS) or teichoic acids (TAs). In an attempt to identify the GtfA substrate, the biochemical composition and structure of CWPS components (rhamnan and polysaccharide pellicle (PSP)) and lipoteichoic acids (LTAs) of *L. lactis* strains 3107 and NZ9000, and a number of derivative strains were analysed. Results show that derivatives possessing a non-functional *gtfA* do not possess any structural modifications in either the rhamnan or PSP components of the CWPS as compared to their parent strains, while LTA analysis revealed that *gtfA* is not responsible for terminal galactose substitution of LTA.

4.1 Introduction

Lactococcal phage-encoded receptor binding protein(s) (RBPs) recognize and attach to host-encoded receptor molecule(s), such as the cell wall polysaccharide (CWPS) or a surface-exposed protein and/or TA (1), an event that acts as a prelude to phage genome injection into the host cell cytoplasm (1, 2). Structural details of the CWPS have previously been investigated in certain lactococcal strains (3-9). The lactococcal CWPS consists of two covalently-linked constituents: the surface-exposed polysaccharide pellicle (PSP) and the recently described rhamnan polysaccharide, which is embedded within the peptidoglycan layer (3). The first PSP structure to be resolved was that of *L. lactis* subsp. *cremoris* strain MG1363 (4, 5), followed by that of *L. lactis* subsp. *cremoris* (*L. lactis*) 3107 (6) and SMQ-388 (7). These PSP structures exhibit subtle differences in their biochemical make-up and, interestingly, act as receptors to determine host specificity of many lactococcal phages (6). The receptor for the industrially significant P335 (6, 10) and 936 (5, 6, 11) phage groups, as well as the 1358 (7), 949 (12) and P087 (12) phage groups is saccharidic in nature, and the variable PSP has been identified as the principal phage receptor in a number of cases (6, 7, 10, 12). Interestingly, the c2 group of lactococcal phages appears to utilize both a carbohydrate receptor as well as a proteinaceous receptor moiety to achieve successful attachment and subsequent infection (13, 14).

Phage-host interaction studies in Gram-positive bacteria have demonstrated that also teichoic acids (TAs) may act as receptors for specific phages (15-19). In the cell wall of Gram-positive bacteria, such as *Bacillus subtilis*, *Listeria monocytogenes*, and *Staphylococcus aureus*, TAs are present as wall teichoic acids (WTA) (covalently attached to peptidoglycan) and/or lipoteichoic acids (LTA) (anchored in the cytoplasmic membrane by means of a lipid attachment) (20-22). LTA is a polyglycerol-phosphate (GroP) chain decorated with D-alanine or/and sugar residues (23-28). The enzymes involved in the D-alanylation and glycosylation process have been characterized for certain bacterial species (21, 29, 30). The LTA glycosylation process in *B. subtilis* and *L. monocytogenes* is instigated by a cytoplasmic glycosyltransferase (GTF) named CsbB and GtlA, which are presumed to transfer N-acetylglucosamine (GlcNAc) and galactose (Gal) residues, respectively, onto the undecaprenyl lipid carrier. The sugar-lipid intermediate is then re-oriented to the other side of the membrane by an as yet to be identified flippase enzyme. The GlcNAc/Gal residues are then transferred onto the GroP chain of the LTA by a second membrane-associated

GTF identified as YfhO and GtlB in *B. subtilis* and *L. monocytogenes*, respectively. Glycosylation of the GroP chain of LTA thus occurs at the periplasmic side of the membrane (31). WTA is also a GroP chain decorated with molecules such as glycosyl or D-alanyl, and unlike LTA, the GroP chain of WTA is attached to a "linkage unit" composed of *N*-acetylmannosaminy1β-(1-4)glucosamine which allows covalent attachment to the peptidoglycan (32). WTA glycosylation was initially thought to occur on the cytoplasmic side of the cell membrane as WTA subunit biosynthesis takes place at this location in *S. aureus* (33-36) and *B. subtilis* (37, 38). However, WTA glycosylation was recently proposed to be an extracellular process in *L. monocytogenes* and other Gram-positive bacteria (31). In *L. monocytogenes*, the glycosylation process of WTA has been suggested to be similar to that of LTA where Lmo2550 represents a GTF required for undecaprenyl glycosylation, while Lmo1079 is the GTF involved in periplasmic glycosylation of WTA (31).

Currently it is not known if *L. lactis* strains synthesize WTA and to what extent this structure is involved in phage infection. LTA extraction related to *L. lactis* subsp. *cremoris* SK110 was shown to contain galactosyl-containing LTA which was linked to phage infection (18). Subsequent to this study LTA isolated from *Lactococcus lactis* CRL 526 was chemically analysed, revealing (not surprisingly) a GroP chain (39). The first complete LTA structure from an *L. lactis* strain to be characterized was that of *L. lactis* subsp. *lactis* IL1403 as shown in Figure 4.1 (8). Biochemical analysis of this particular LTA showed that following LTA polymerization, the GroP chain of LTA is substituted with single α-galactose (α-Gal) or alanine (Ala) residues (8, 40).

In Chapter III of this thesis, we demonstrated that $gtfA_{3107}$, which encodes a predicted membrane-associated GTF protein, is required for injection of *L. lactis* strain 3107 by lactococcal phage TP901-1 (a member of the P335 phage group). Our current hypothesis is that $GtfA_{3107}$ catalyses the glycosidic modification of a cell envelope-associated polysaccharide component, which acts as an injection trigger for TP901-1 following its initial adsorption to a CWPS component. The presumed GtfA substrates may include the PSP, rhamnan, LTA, WTA (if present), peptidoglycan or a glycoprotein. The aim of this study is to determine if the substrate for GtfA is a CWPS component (PSP and/or rhamnan) or LTA, this was analysed by determining if the expression of $gtfA_{3107}$ evokes modification of LTA or CWPS components (PSP and/or rhamnan). This was assessed through the comparative biochemical analysis of

extracted LTA and/or CWPS from *L. lactis* strains 3107 and NZ9000, and their *gtfA*-mutated derivatives.

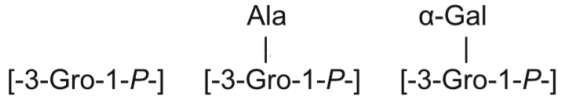


Figure 4.1 Composition of lipoteichoic acid structure of *L. lactis* IL1403; figure taken from (8). The LTA tructure is composed of a glycerol-phosphate polymer [-3-Gro-1-P] partially glycosylated with α-galactose (α-Gal) or acylated with alanine (Ala). The LTA mean [-3-Gro-1-P] chain length is 16, following polymerization 11.8 ± 1.6 % of the chain is substituted with galactose and 28.5 ± 6.6 % is substituted with D-Alanine, while 59.8 ± 5.2 % of the chain remains non-substituted (8, 40).

4.2 Materials and Methods

4.2.1 Bacterial strains and growth conditions

Strains employed in this study are listed in Table 4.1. *L. lactis* strains were grown at 30 °C in M17 broth/agar (Oxoid, U.K.) supplemented with 0.5 % (v/v) glucose (GM17) and 5 μg ml^{-1} chloramphenicol (Cm), where necessary.

Table 4.1 Bacterial strains used in this study.

Bacterial Strains	Description	Source
Lactococcus lactis subsp. *cremoris*		
3107	Host strain for model P335 phage TP901-1	(41)
E119	TP901-1-resistant derivative of 3107	(42)
E121	TP901-1-resistant derivative of 3107	(42)
E126	TP901-1-resistant derivative of 3107	(42)
E126_pNZ*gtfA*	E126 harbouring complementing plasmid pNZ*gtfA*; pNZ44 vector harbouring *gtfA* from *L. lactis* 3107	Chapter III
3107_pNZ44	3107 harbouring the empty vector pNZ44	Chapter III
3107_pNZS*gtfA*	3107 harbouring the silencing plasmid pNZS*gtfA*; pNZ44 vector harbouring *gtfA* from *L. lactis* 3107 in the anti-sense orientation	Chapter III
3107_pNZ*gtfA*	3107 harbouring the over-expressing plasmid pNZ*gtfA*	Chapter III
3107_pNZ*1441-1443*	3107 harbouring the over-expressing plasmid pNZ*1441-1443*; pNZ44 containing *L3107_1441*, *gtfA* and *L3107_1443* from *L. lactis* 3107.	Chapter III
NZ9000	MG1363 derivative in which the *nisRK* genes have been inserted in the chromosome	(43)
NZ9000Δ*8915*	NZ9000 harbouring a non-sense mutation in *LLNZ_08915*, homologue of *L3107_1441*	Chapter III
NZ9000Δ*gtfA*	NZ9000 harbouring a non-sense mutation in *LLNZ_08920*, homologue of $gtfA_{3107}$	Chapter III
NZ9000Δ*8915_ gtfA*	NZ9000 harbouring a non-sense mutation in *LLNZ_08915* as well as *LLNZ_08920*, homologues of *L3107_1441* and $gtfA_{3107}$, respectively.	Chapter III

4.2.2 Cell wall polysaccharide (CWPS) extraction, purification and analysis

CWPS was extracted as described previously (3). Briefly, cells from an exponentially growing culture (corresponding to an OD_{600nm} between 0.6 and 0.8) were harvested by centrifugation at 1,900 × *g* for 15 minutes at 4 °C and washed twice with deionized ice-cold water. CWPS was then extracted by 48 % hydrofluoric acid (HF) treatment for 48 h at 4 °C. After drying, rhamnan and PSP oligosaccharides released by HF treatment were separated by size exclusion chromatography (SEC) with a high-performance liquid chromatography (HPLC) system (SEC-HPLC) as described previously (3). Elution was performed with Milli-Q H_2O and detection of eluted compounds was performed with a refractometer (2414 Refractive Index Detector, Waters) and/or UV detector at 206 nm. Fractions corresponding to peaks containing rhamnan and PSP oligosaccharides were collected and dried under vacuum and further analyzed for composition and mass as described previously (3). Monosaccharide composition was determined after trifluoroacetic acid (TFA) hydrolysis by high performance anion exchange chromatography coupled with pulse-amperometric detection (HPAEC-PAD) (ICS5000 system, Thermo Fisher scientific).

4.2.3 Lipoteichoic acid (LTA) extraction and analysis

Strains *L. lactis* 3107 and its mutant derivative E126 were grown overnight in 4 L of GM17, after which cells were harvested by centrifugation at 1900 × *g* for 15 minutes at 4 °C and washed twice with deionized ice-cold water. LTA (anchored in the cytoplasmic membrane) was extracted from cells with hot aqueous phenol as previously described (18). Briefly, the cell pellet (~5 g) was suspended in Chloroform-Methanol-Water (30:15:2.5, v/v) and stirred overnight. Following centrifugation at 3,000 x *g* for 30 minutes, the defatted cells were suspended in 20 ml of 0.1 M sodium acetate buffer (pH 5) and 20 ml of 40 % hot aqueous phenol was added. The suspension was stirred at 65 °C for 1 h, cooled and transferred into teflon centrifuge tubes. The suspension was further cooled by placing the tubes in ice for the separation phase. The cooled suspension was then centrifuged to separate the aqueous and phenol phases; the aqueous phase (which is expected to contain LTA) was collected and re-extracted with chloroform. Crude aqueous extracts were further de-proteinated by the addition of trichloroacetic acid (TCA) to a final concentration of 5 %. Following centrifugation, the clear supernatant was dialyzed and lyophilized to obtain a crude LTA preparation. Crude extracts (5 mg) were subjected to composition and

methylation analyses by gas chromatography (GC) and gas chromatography-mass spectrometry (GC-MS) after treatment with 48 % HF (4 °C, 24 h), acid hydrolysis and conversion into alditol acetates, as described previously (44). For the detection of glycerol, hydrolysed samples were acetylated.

4.3 Results

4.3.1 Bioinformatic analysis and predicted function of GtfA

The *gtfA*$_{3107}$ gene encodes a putative membrane-associated GTF protein, GtfA, which is predicted to possess fourteen transmembrane helices and a large extracellular loop containing the catalytic site of the protein (Chapter III). These topological characteristics are similar to those of YfhO and Lmo1079 which are members of the GT-C family of GTFs, and which exhibit 19 % amino acid sequence identity to GtfA. The GT-C family of GTFs are characterized by 8 – 13 transmembrane helices and an extracellular loop (45). GtfA is thus predicted to be a member of the GT-C family of GTFs, which are known to be involved in glycosylation of substrates at the periplasmic side of the cytoplasmic membrane, e.g. YfhO and Lmo1079 of *B. subtilis* and *L. monocytogenes* glycosylate the GroP chain of LTA and WTA, respectively.

Based on the above, we predicted that GtfA catalyses a glycosylation event in *L. lactis* 3107 by transferring a mono/oligo-saccharide from its undecaprenyl carrier to its extracytoplasmic substrate. This proposed glycosylation process is schematically depicted in Figure 4.2. The presumed substrate for GtfA is TA (WTA or LTA), CWPS components (PSP or rhamnan), peptidoglycan or a glycoprotein of *L. lactis* 3107. Here, to identify the substrate of GtfA, candidates of which are LTA, PSP or rhamnan, biochemical analysis of LTA structures in the 3107 strain and a *gtfA* mutant (stop codon) derivative (E126) of 3107 was performed as described in more detail below. Furthermore, as described below, biochemical analysis of PSP and rhamnan structures in 3107 and NZ9000, and their derivatives was also performed.

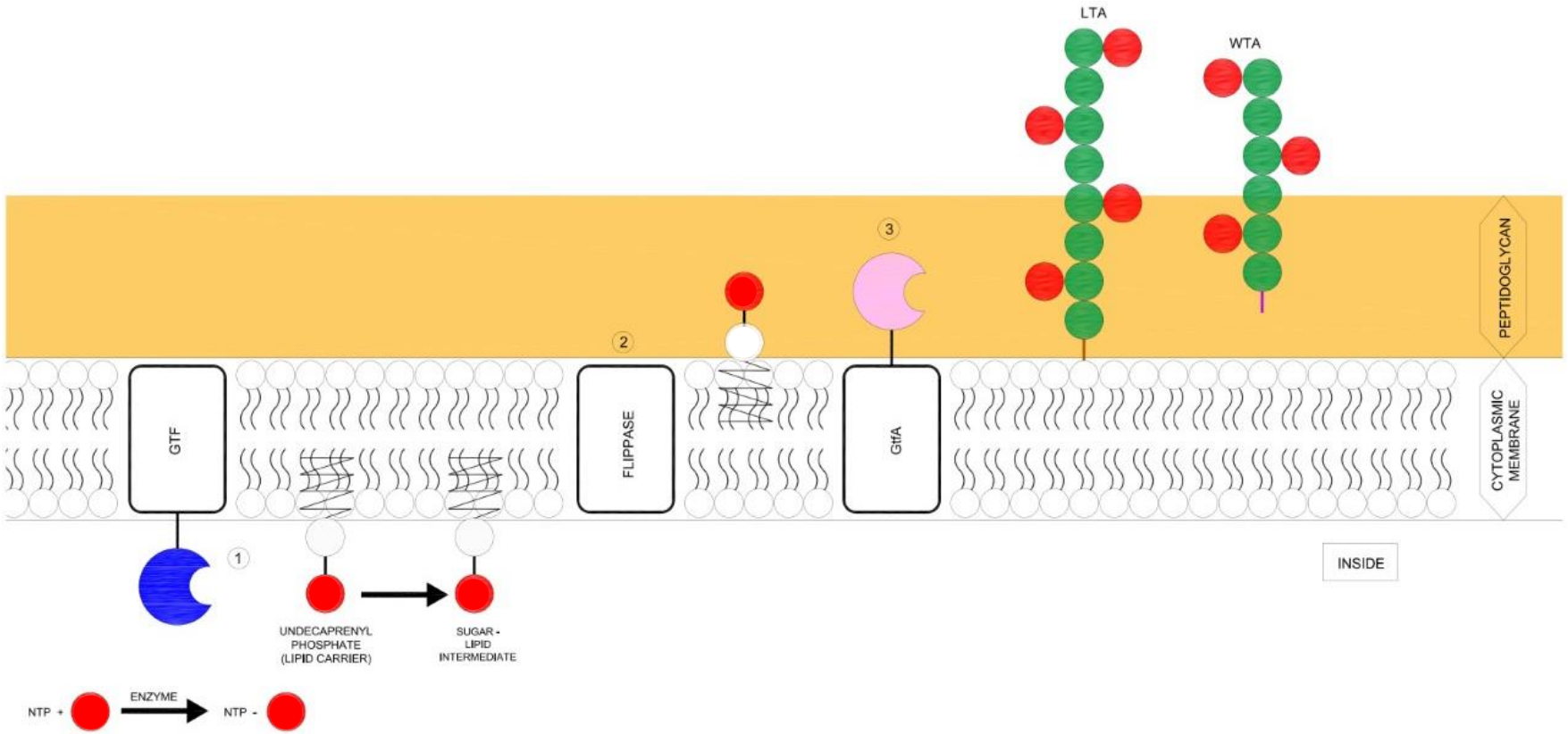


Figure 4.2: Model for TA (LTA or WTA) glycosylation in *L. lactis*. The LTA or WTA glycosylation process in *L. lactis* initially requires an as yet unidentified cytoplasmic GTF (blue) which transfers a sugar residue (red) onto an undecaprenyl-phosphate which acts as a lipid carrier (Step 1). The sugar-lipid intermediate is then transported across the membrane by an unknown flippase enzyme (Step 2). The sugar residue is then transferred by a second extracellular GTF proposed to be GtfA (pink) onto the GroP chain (green) of the LTA which is anchored in the cytoplasmic membrane by means of a lipid attachment (brown) or WTA which is covalently attached to peptidoglycan (purple) (Step 3).

4.3.2 Analysis of CWPS from *L. lactis* 3107 and its derivatives

To assess the predicted role of *gtfA*$_{3107}$ in glycosylation of a particular cell wall-associated glycan, the two components of CWPS (rhamnan and PSP) from *L. lactis* 3107 and its derivatives were purified and analysed. Rhamnan and PSP oligosaccharides were extracted from cell walls by HF treatment and purified by SEC-HPLC (Figs. 4.3, 4.4 and 4.5). Following this, the monosaccharide composition of the extracted rhamnan and PSP oligosaccharides was determined and analyzed by Maldi-Tof MS according to previously described protocols (3). No significant differences were detected at the level of rhamnan and PSP structures, between the *L. lactis* 3107 strain and its phage-resistant derivatives E119 and E126 (*gtfA* mutants), or the complemented versions of these strains. Furthermore, *L. lactis* 3107 containing the following plasmids were also analysed; plasmid pNZ*gtfA*, which causes constitutive (over)expression of *gtfA*$_{3107}$, plasmid pNZS*gtfA* in which expression of *gtfA*$_{3107}$ was silenced, plasmid pNZ*gtfA_1443* causing constitutive (over)expression of *L3107_1441*, *gtfA*$_{3107}$ and *L3107_1443*, as well as the empty pNZ44 vector as a control. All these *L. lactis* 3107 derivatives also displayed identical PSP and rhamnan structures to the parent strain (Figs. 4.6 and 4.7). Figures 4.6 and 4.7 provides data on the monosaccharide composition analysis of extracted rhamnan and PSP, respectively, from 3107 and some of its analysed derivatives relative to the amount of Glucosamine ($GlcNH_2$) present. Deacetylation of N-acetylglucosamine (GlcNAc) leads to the presence of $GlcNH_2$, thus an alteration on the level of $GlcNH_2$ present would be indicative of rhamnan or PSP modification. Furthermore, the composition of extracted rhamnan was also analysed relative to the amount of rhamnose, galactose and glucose (data not shown). Although variation was observed between the individual extraction attempts (Figs. 4.3 to 4.5) in PSP and rhamnan, the compositional analysis of these extracted structures exhibited no variation. These results therefore demonstrate that neither the PSP nor the rhamnan were structurally modified, and that *gtfA*$_{3107}$ is not involved in PSP or rhamnan modification.

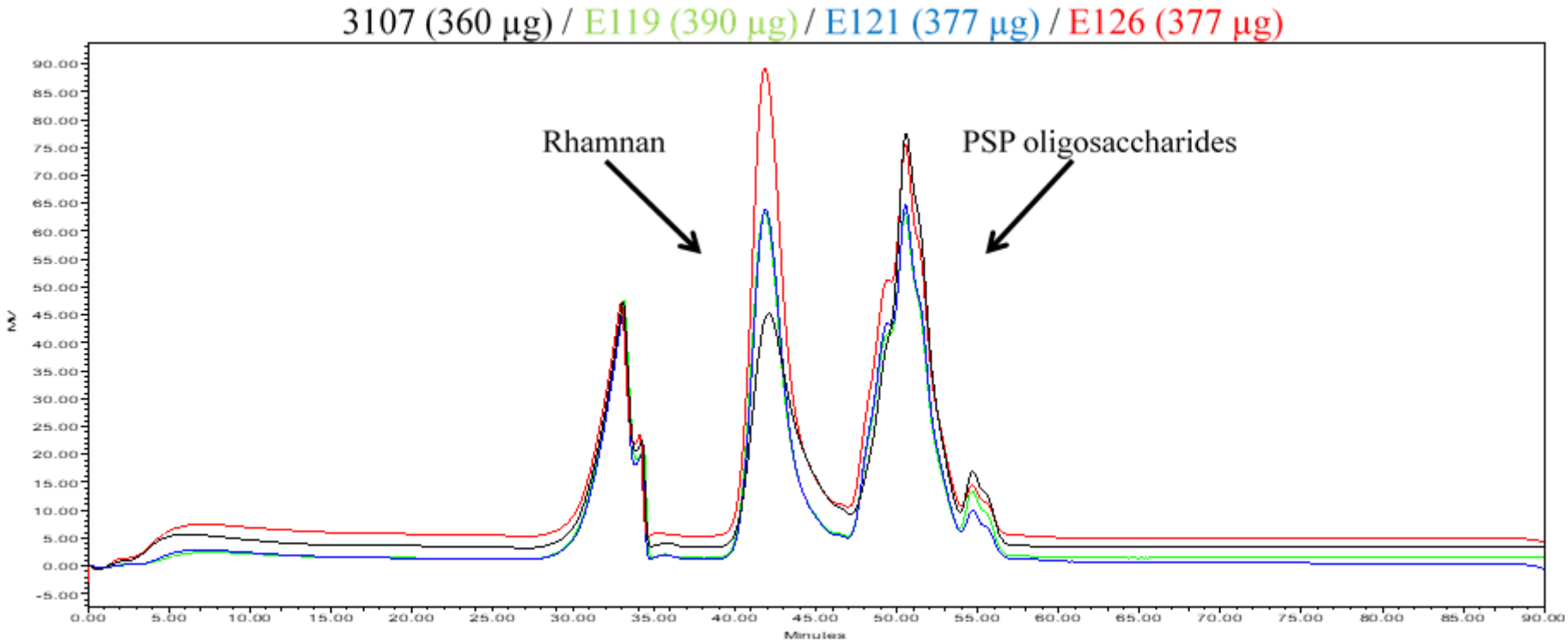


Figure 4.3: Size exclusion chromatography with high-performance liquid chromatography (SEC-HPLC) purification of rhamnan and PSP oligosaccharides extracted from cell walls of *L. lactis* 3107 and its derivatives, E119, E121 and E126 by hydrofluoric acid (HF) treatment.

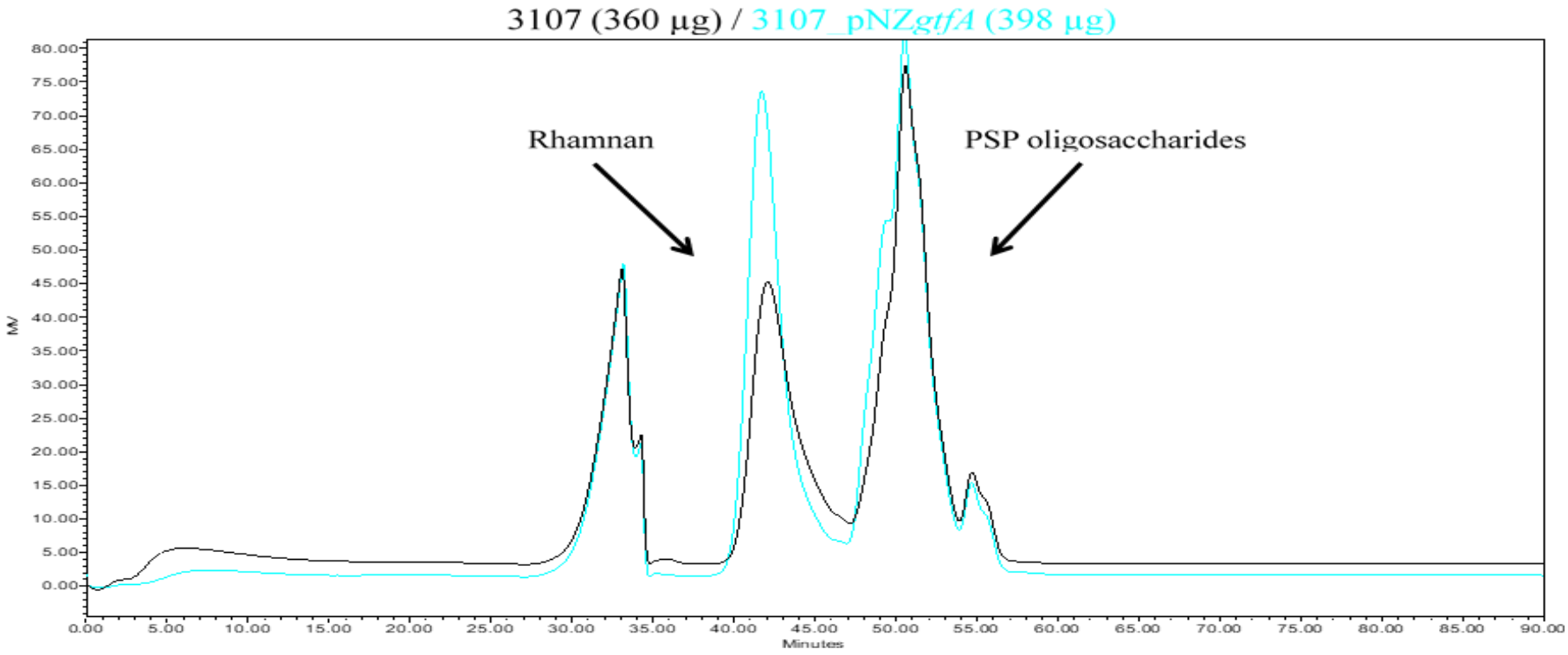


Figure 4.4: Size exclusion chromatography with high-performance liquid chromatography (SEC-HPLC) purification of rhamnan and PSP oligosaccharides extracted from cell walls of *L. lactis* 3107 and the 3107 overexpressing *gtfA* by hydrofluoric acid (HF) treatment.

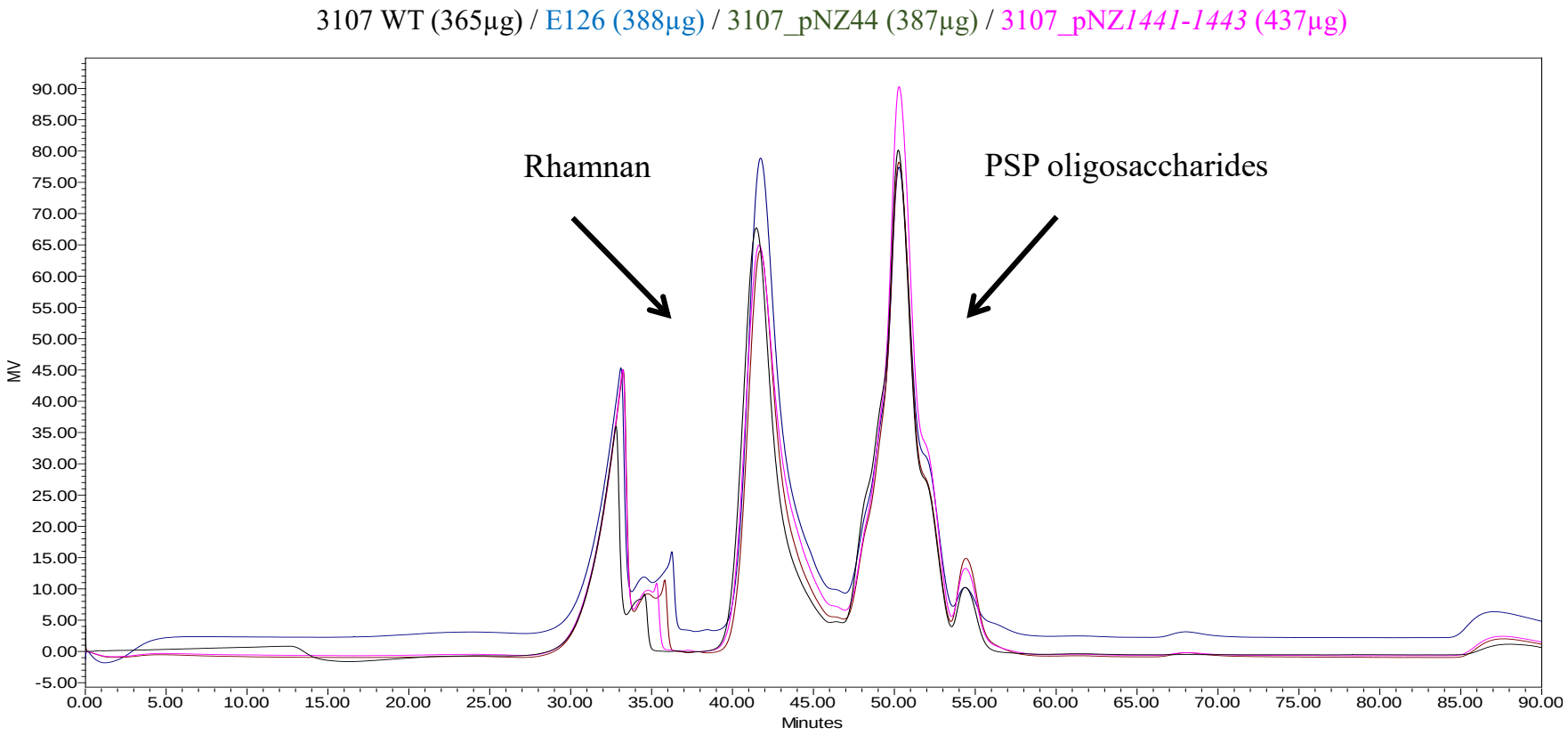


Figure 4.5: Size exclusion chromatography with high-performance liquid chromatography (SEC-HPLC) purification of rhamnan and PSP oligosaccharides extracted from cell walls of *L. lactis* 3107 and its derivatives, E126, 3107_pNZ44 and 3107_pNZ*1441-1443* by hydrofluoric acid (HF) treatment.

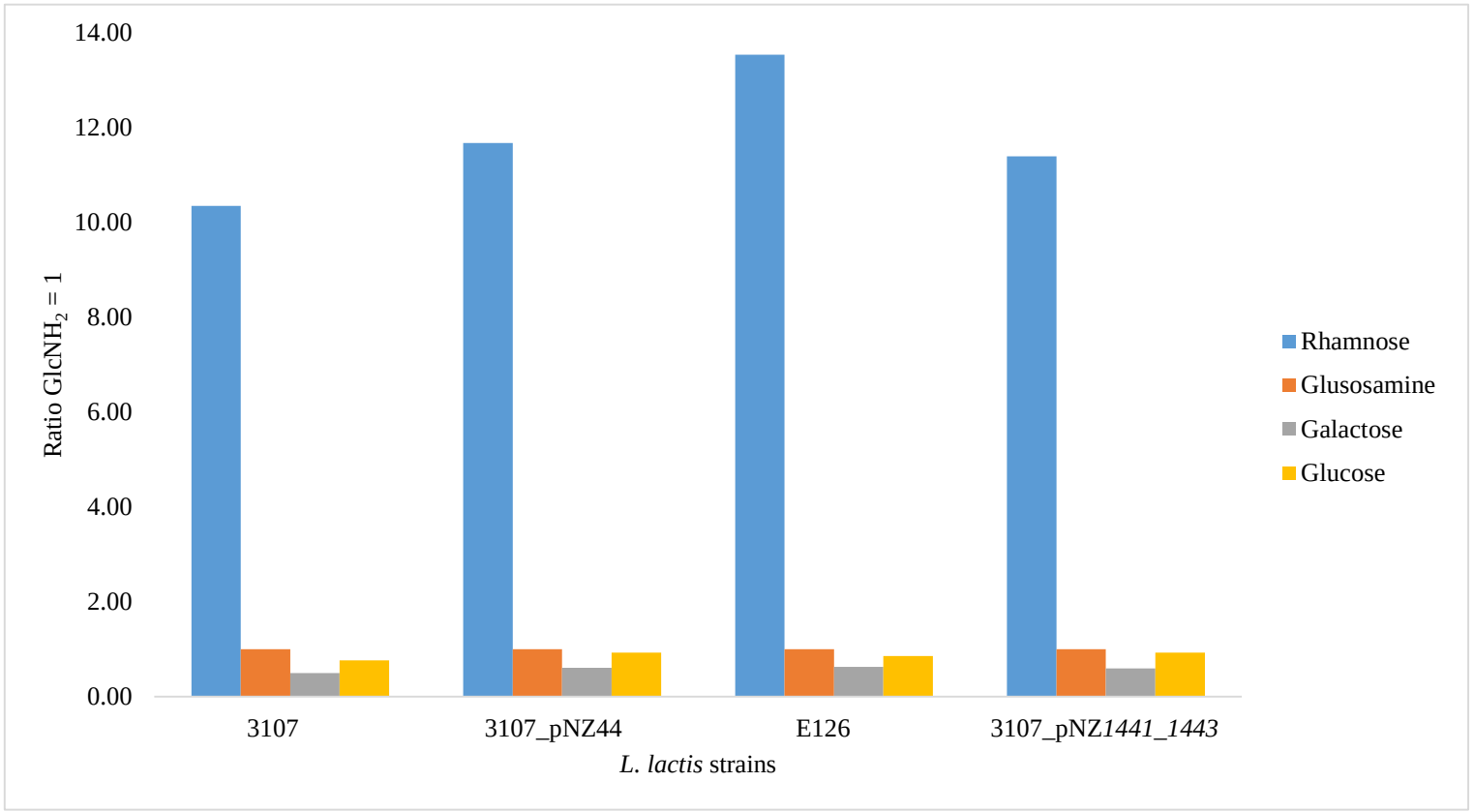


Figure 4.6: Monosaccharide composition of rhamnan peak relative to Glucosamine ($GlcNH_2$ = 1) following SEC-HPLC purification of rhamnan oligosaccharides extracted from cell walls of *L. lactis* 3107 and its derivatives, 3107_pNZ44, E126 and 3107_pNZ*1441-1443*.

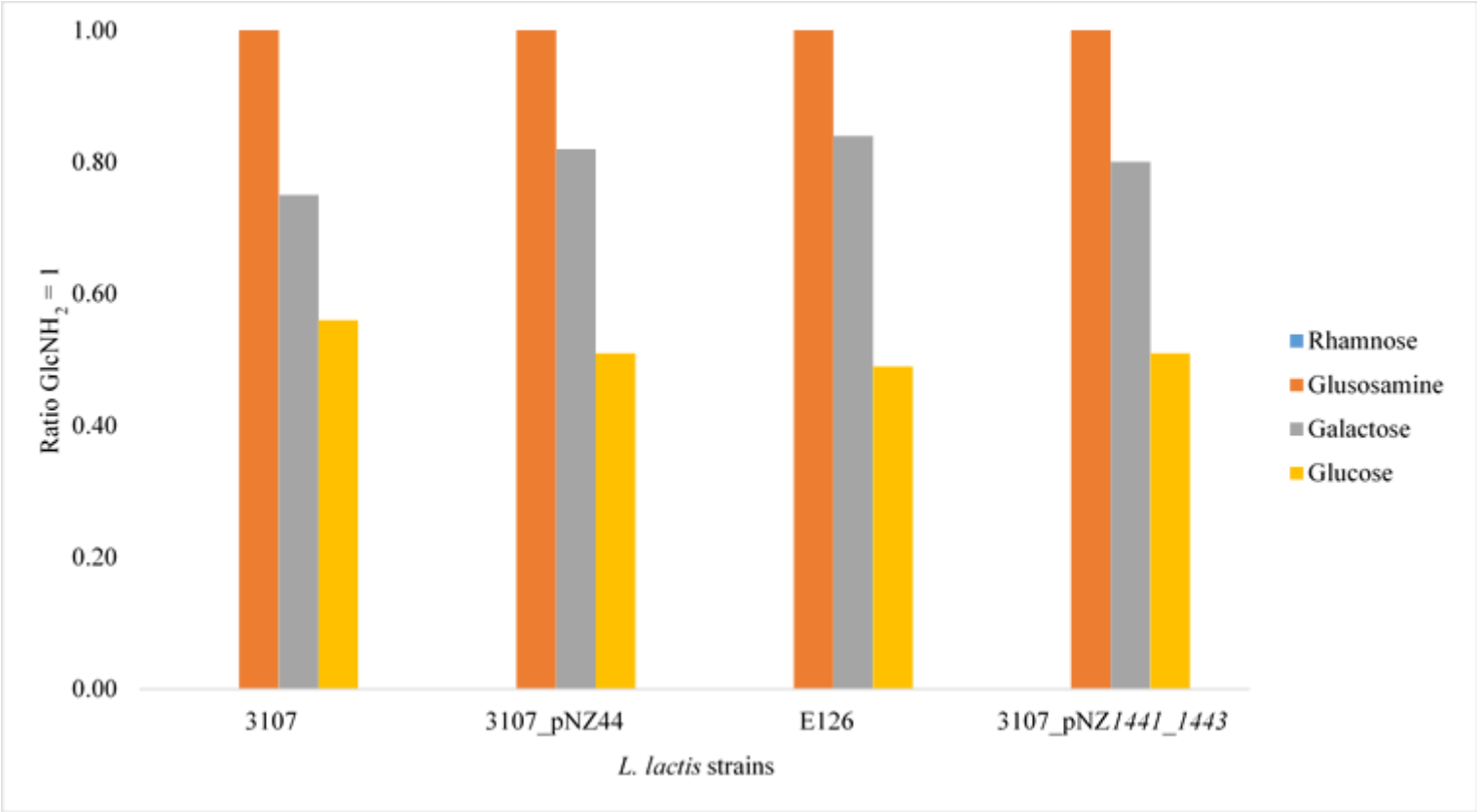


Figure 4.7: Monosaccharide composition of polysaccharide pellicle (PSP) peak relative to Glucosamine ($GlcNH_2$ = 1) following SEC-HPLC purification of PSP oligosaccharides extracted from cell walls of *L. lactis* 3107 and its derivatives, 3107_pNZ44, E126 and 3107_pNZ*1441-1443*.

4.3.3 CWPS analysis of *L. lactis* NZ9000 and its derivatives.

The results presented above indicate that GtfA is not involved in PSP or rhamnan modification in *L. lactis* 3107 (Section 4.3.2). *L. lactis* subsp. *cremoris* NZ9000 harbours almost identical (99 % identity) homologues of the *L3107_1441* and $gtfA_{3107}$ genes on its genome, referred to here as *LLNZ_08915* and $gtfA_{NZ9000}$ (corresponding to locus tag *LLNZ_08920*), respectively. Thus if these genes are involved in the modification of CWP components, the inactivation of these genes should generate mutant strains deficient in the glycosylation of PSP or rhamnan. Stop codons were introduced into the coding sequence of *LLNZ_08915* and $gtfA_{NZ9000}$ by targeted mutagenesis [32, 35] either individually or in combination, thereby generating NZ9000*Δ8915*, NZ9000*ΔgtfA* and NZ9000*Δ8915_gtfA*, respectively (Chapter III). Analysis of the biochemical composition and architecture of rhamnan and PSP structures in NZ9000 and its derivatives (data not shown) demonstrated results consistent with 3107 and its derivatives, thus corroborating the earlier finding for *L. lactis* 3107 that GtfA is not involved in CWPS modification.

4.3.4 Analysis of LTA from *L. lactis* 3107 and its derivatives

Since GtfA is not involved in modification of PSP nor rhamnan in 3107 or NZ9000, it may catalyse the extracellular glycosylation of LTA (or WTA if present). LTA analysis of *L. lactis* IL1403 has previously shown that the GroP chain of LTA is substituted with single α-Gal residues (8, 40) which can be identified after LTA depolymerization with HF and methylation analysis. Thus, assuming that the glycosylation process of LTA in strain 3107 is identical to that of *L. lactis* IL1403, the level of glycosylation can be analysed by the assessment of terminal Galactose (t-Gal) residues. The PSP and rhamnan components of the CWPS of *L. lactis* 3107 do not contain t-Gal, therefore, the presence of minor amounts of these polymers in TA preparations would not affect the analysis.

LTA extracts from *L. lactis* 3107, its derived mutant E126 (which does not produce $GtfA_{3107}$) and *L. lactis* IL1403 (as a positive control) were subjected to composition and methylation analyses. As expected, the positive control strain *L. lactis* IL1403 contained t-Gal in all preparations. In contrast, t-Gal was absent in LTA preparations from both 3107 and its E126 mutant (Fig. 4.8). This result indicates an absence of galactosylation of the retrieved LTAs of 3107. Thus, on the basis of current data we conclude that GtfA is not involved in LTA galactosylation in *L. lactis* 3107. Further analysis is ongoing to determine if GtfA is involved in LTA glycosylation, but involving a different terminal sugar, while glycosylation of WTA (if present) will also be assessed.

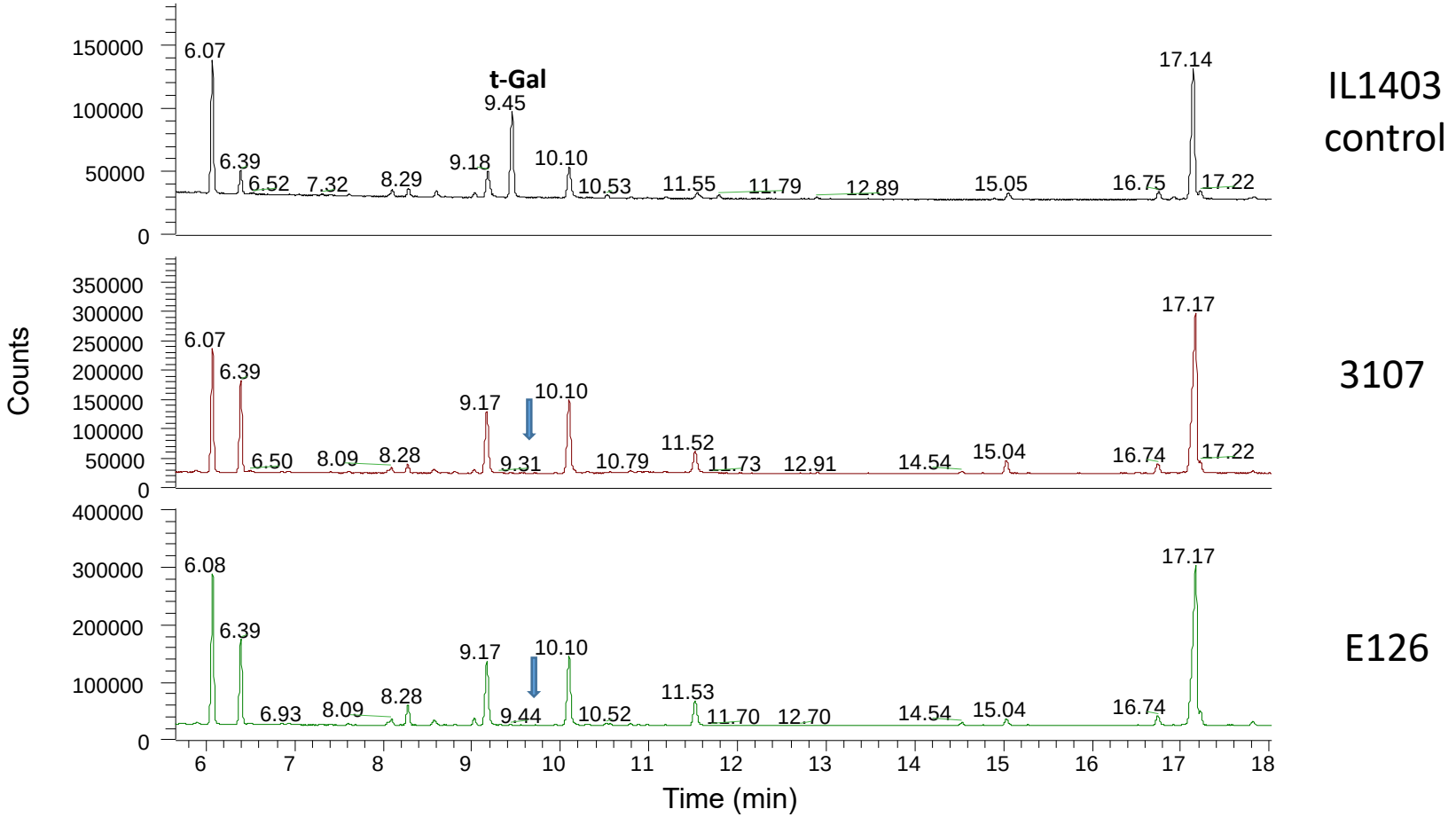


Figure 4.8: Gas chromatography-(GC) profiles corresponding to methylation analysis of LTA preparations of *L. lactis* IL1403 (as a control), *L. lactis* 3107 and its mutant E126 following phenol extraction. Presence of terminal-Galactose (t-Gal) shown with "t-Gal", and absence of t-Gal shown with a blue arrow.

4.4 Discussion

$GtfA_{3107}$ has been shown to be specifically involved in phage-host interactions of TP901-1, 949 and WRP3 (Chapter III). Previous studies have shown that differences in the PSP component of the CWPS play a crucial role in determining the phage adsorption and host range (6). Based on these findings, a possible substrate for GtfA was the CWPS, represented by its constituent components rhamnan and PSP. Thus, the structures of rhamnan and PSP, of *L. lactis* strains 3107 and NZ9000, and various derivatives were biochemically analysed. The obtained results clearly demonstrate that GtfA is not involved in the modification of rhamnan or PSP in these *L. lactis* strains. This suggests that the substrate for the GtfA encoded by $gtfA_{3107}$ is another cell wall glycopolymer, possibly LTA (or WTA if present).

Here, GtfA is predicted to exhibit the same GT-C fold as YfhO and Lmo1079, which catalyse the substitution of LTA (with GlcNAc) and WTA (with Gal) in *B. subtilis* and *L. monocytogenes*, respectively. Thus, in *L. lactis* 3107, GtfA is proposed to transfer a sugar from undecaprenyl-phosphate to its extracytoplasmic TA substrate. Assuming that TP901-1, 949 and WRP3 phages attach to such a glycosylated TA (WTA or LTA) on the surface of *L. lactis* 3107, the presence of a non-functional GtfA protein would result in 3107 mutant strains resistant to these lactococcal phages due to their deficiency in TA glycosylation.

Phage-host interaction studies have indeed demonstrated that bacterial WTA or LTA of *B. subtilis, L. lactis* and *Lactobacillus delbrueckii* can act as host-receptors for specific phages (15-19). Studies have also shown that modification of the WTA or LTA structures can result in phage-resistance (18, 46, 47). For example, it has been shown previously that the LTA composition of an *L. lactis* phage-resistant strain differed from its phage-sensitive variant by the presence of a galactosyl-containing component (18). Certain *B. subtilis*-infecting phages require glucosylated polyglycerol TA for adsorption, and thus *B. subtilis* phage-resistant mutant strains have been shown to be deficient in glucosylation of their TAs (46). Furthermore, a phage-resistant *Listeria* strain was shown to lack GlcNac in its TA (47). Additionally, using WTA biochemical and structural analysis, phage was previously shown to directly adsorb to galactosylated WTA in *L. monocytogenes* (48).

Our analysis shows that the LTAs present in 3107 or its phage-resistant derivative E126 do not appear to contain a Gal substitution. Therefore, no definite conclusion can be drawn on the involvement of GtfA in TA glycosylation in *L. lactis*

3107 as yet. Further biochemical analysis of TA (both LTA and, if present, WTA) composition in *L. lactis* 3107 has to be performed in order to identify the GtfA substrate and its relevance to phage infection.

4.6 References

1. Rakhuba D, Kolomiets E, Dey ES, Novik G. 2010. Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. Polish Journal of Microbiology 59:145-155.
2. Mahony J, Ainsworth S, Stockdale S, van Sinderen D. 2012. Phages of lactic acid bacteria: the role of genetics in understanding phage-host interactions and their co-evolutionary processes. Virology 434:143-150.
3. Sadovskaya I, Vinogradov E, Courtin P, Armalyte J, Meyrand M, Giaouris E, Palussière S, Furlan S, Péchoux C, Ainsworth S. 2017. Another brick in the wall: a rhamnan polysaccharide trapped inside peptidoglycan of *Lactococcus lactis*. MBio 8:e01303-17.
4. Andre G, Kulakauskas S, Chapot-Chartier M-P, Navet B, Deghorain M, Bernard E, Hols P, Dufrêne YF. 2010. Imaging the nanoscale organization of peptidoglycan in living *Lactococcus lactis* cells. Nature communications 1:27.
5. Chapot-Chartier MP, Vinogradov E, Sadovskaya I, Andre G, Mistou MY, Trieu-Cuot P, Furlan S, Bidnenko E, Courtin P, Pechoux C. 2010. Cell surface of *Lactococcus lactis* is covered by a protective polysaccharide pellicle. Journal of biological chemistry 285:10464-71.
6. Ainsworth S, Sadovskaya I, Vinogradov E, Courtin P, Guerardel Y, Mahony J, Grard T, Cambillau C, Chapot-Chartier MP, Sinderen D. 2014. Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. MBio 5:e00880-14.
7. Farenc C, Spinelli S, Vinogradov E, Tremblay D, Blangy S, Sadovskaya I, Moineau S, Cambillau C. 2014. Molecular insights on the recognition of a *Lactococcus lactis* cell wall pellicle by the phage 1358 receptor binding protein. Journal of virol 88:7005-15.
8. Vinogradov E, Sadovskaya I, Courtin P, Kulakauskas S, Grard T, Mahony J, van Sinderen D, Chapot-Chartier M-P. 2018. Determination of the cell wall polysaccharide and teichoic acid structures from *Lactococcus lactis* IL1403. Carbohydrate research 462:39-44.
9. Vinogradov E, Sadovskaya I, Grard T, Murphy J, Mahony J, Chapot-Chartier M-P, van Sinderen D. 2018. Structural studies of the cell wall polysaccharide from *Lactococcus lactis* UC509.9. Carbohydrate research 461:25-31.

10. Legrand P, Collins B, Blangy S, Murphy J, Spinelli S, Gutierrez C, Richet N, Kellenberger C, Desmyter A, Mahony J. 2016. The atomic structure of the phage Tuc2009 baseplate tripod suggests that host recognition involves two different carbohydrate binding modules. MBio 7:e01781-15.
11. Dupont K, Janzen T, Vogensen FK, Josephsen J, Stuer-Lauridsen B. 2004. Identification of *Lactococcus lactis* genes required for bacteriophage adsorption. Applied and environmental microbiology 70:5825-32.
12. Mahony J, Randazzo W, Neve H, Settanni L, Sinderen D. 2015. Lactococcal 949 group phages recognize a carbohydrate receptor on the host cell surface. Applied and environmental microbiology 81:3299-305.
13. Valyasevi R, Sandine W, Geller B. 1991. A membrane protein is required for bacteriophage c2 infection of *Lactococcus lactis* subsp. *lactis* C2. Journal of bacteriology 173:6095-6100.
14. Monteville MR, Ardestani B, Geller BL. 1994. Lactococcal bacteriophages require a host cell wall carbohydrate and a plasma membrane protein for adsorption and ejection of DNA. Applied and environmental microbiology 60:3204-3211.
15. Xia G, Corrigan RM, Winstel V, Goerke C, Gründling A, Peschel A. 2011. Wall teichoic acid-dependent adsorption of *Staphylococcal siphovirus* and *myovirus*. Journal of bacteriology 193:4006-4009.
16. Young F. 1967. Requirement of glucosylated teichoic acid for adsorption of phage in *Bacillus subtilis* 168. Proceedings of the national academy of sciences of the United States of America 58:2377-2384.
17. Baptista C, Santos MA, São-José C. 2008. Phage SPP1 reversible adsorption to *Bacillus subtilis* cell wall teichoic acids accelerates virus recognition of membrane receptor YueB. Journal of bacteriology 190:4989-4996.
18. Sijtsma L, Wouters J, Hellingwerf K. 1990. Isolation and characterization of lipoteichoic acid, a cell envelope component involved in preventing phage adsorption, from *Lactococcus lactis* subsp. *cremoris* SK110. Journal of bacteriology 172:7126-7130.
19. Räisänen L, Schubert K, Jaakonsaari T, Alatossava T. 2004. Characterization of lipoteichoic acids as *Lactobacillus delbrueckii* phage receptor components. Journal of bacteriology 186:5529-5532.

20. Araki Y, Ito E. 1989. Linkage units in cell walls of Gram-positive bacteria. Critical reviews in microbiology 17:121-135.
21. Neuhaus FC, Baddiley J. 2003. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in Gram-positive bacteria. Microbiology and molecular biology reviews 67:686-723.
22. Percy MG, Gründling A. 2014. Lipoteichoic acid synthesis and function in Gram-positive bacteria. Annual review of microbiology 68:81-100.
23. Fischer W, Rösel P. 1980. The alanine ester substitution of lipoteichoic acid (LTA) in *Staphylococcus aureus*. FEBS letters 119:224-226.
24. Iwasaki H, Shimada A, Ito E. 1986. Comparative studies of lipoteichoic acids from several *Bacillus strains*. Journal of bacteriology 167:508-516.
25. Fischer W. 1988. Physiology of lipoteichoic acids in bacteria, p233-302, Advances in microbial physiology, vol 29. Elsevier.
26. Iwasaki H, Shimada A, Yokoyama K, Ito E. 1989. Structure and glycosylation of lipoteichoic acids in *Bacillus* strains. Journal of bacteriology 171:424-429.
27. Hether NW, Jackson LL. 1983. Lipoteichoic acid from *Listeria monocytogenes*. Journal of bacteriology 156:809-817.
28. Uchikawa K, Sekikawa I, Azuma I. 1986. Structural studies on lipoteichoic acids from four *Listeria* strains. Journal of bacteriology 168:115-122.
29. Heaton MP, Neuhaus FC. 1992. Biosynthesis of D-alanyl-lipoteichoic acid: cloning, nucleotide sequence, and expression of the *Lactobacillus casei* gene for the D-alanine-activating enzyme. Journal of bacteriology 174:4707-4717.
30. Abachin E, Poyart C, Pellegrini E, Milohanic E, Fiedler F, Berche P, Trieu-Cuot P. 2002. Formation of D-alanyl-lipoteichoic acid is required for adhesion and virulence of *Listeria monocytogenes*. Molecular microbiology 43:1-14.
31. Rismondo J, Percy MG, Gründling A. 2018. Discovery of genes required for lipoteichoic acid glycosylation predicts two distinct mechanisms for wall teichoic acid glycosylation. Journal of biological chemistry 293:3293-3306.
32. Delcour J, Ferain T, Deghorain M, Palumbo E, Hols P. 1999. The biosynthesis and functionality of the cell-wall of lactic acid bacteria, p159-184, Lactic acid bacteria: genetics, metabolism and applications. Springer.
33. Brown S, Xia G, Luhachack LG, Campbell J, Meredith TC, Chen C, Winstel V, Gekeler C, Irazoqui JE, Peschel A. 2012. Methicillin resistance in

Staphylococcus aureus requires glycosylated wall teichoic acids. Proceedings of the national academy of sciences 109(46):18909-18914.

34. Xia G, Maier L, Sanchez-Carballo P, Li M, Otto M, Holst O, Peschel A. 2010. Glycosylation of wall teichoic acid in *Staphylococcus aureus* by TarM. Journal of biological chemistry 285(18):13405-13415.
35. Winstel V, Xia G, Peschel A. 2014. Pathways and roles of wall teichoic acid glycosylation in *Staphylococcus aureus*. International journal of medical microbiology 304:215-221.
36. Sobhanifar S, Worrall LJ, Gruninger RJ, Wasney GA, Blaukopf M, Baumann L, Lameignere E, Solomonson M, Brown ED, Withers SG. 2015. Structure and mechanism of *Staphylococcus aureus* TarM, the wall teichoic acid α-glycosyltransferase. Proceedings of the national academy of sciences 112:E576-E585.
37. Lazarevic V, Karamata D. 1995. The tagGH operon of *Bacillus subtilis* 168 encodes a two-component ABC transporter involved in the metabolism of two wall teichoic acids. Molecular microbiology 16:345-355.
38. Allison SE, D'Elia MA, Arar S, Monteiro MA, Brown ED. 2011. Studies of the genetics, function and kinetic mechanism of TagE-the wall teichoic acid glycosyltransferase in *Bacillus subtilis* 168. Journal of biological chemistry:jbc. M111. 241265.
39. De Ambrosini Vm, Gonzalez S, Perdigon G, Holgado Apdr, Oliver G. 1996. Chemical composition of the cell wall of lactic acid bacteria and related species. Chemical and pharmaceutical bulletin 44:2263-2267.
40. Kramer NE, Hasper HE, van den Bogaard PT, Morath S, de Kruijff B, Hartung T, Smid EJ, Breukink E, Kok J, Kuipers OP. 2008. Increased D-alanylation of lipoteichoic acid and a thickened septum are main determinants in the nisin resistance mechanism of *Lactococcus lactis*. Microbiology 154:1755-1762.
41. Braun Jr V, Hertwig S, Neve H, Geis A, Teuber M. 1989. Taxonomic differentiation of bacteriophages of *Lactococcus lactis* by electron microscopy, DNA-DNA hybridization, and protein profiles. Microbiology 135:2551-2560.
42. Breum SØ, Neve H, Heller KJ, Vogensen FK. 2007. Temperate phages TP901-1 and φLC3, belonging to the P335 species, apparently use different pathways

for DNA injection in *Lactococcus lactis* subsp. *cremoris* 3107. FEMS microbiology letters 276:156-164.

43. Ainsworth S, Zomer A, de Jager V, Bottacini F, van Hijum SA, Mahony J, van Sinderen D. 2013. Complete genome of *Lactococcus lactis* subsp. *cremoris* UC509.9, host for a model lactococcal P335 bacteriophage. Genome Announcement 1:e00119-12.
44. Chapot-Chartier MP, Vinogradov E, Sadovskaya I, Andre G, Mistou MY, Trieu-Cuot P, Furlan S, Bidnenko E, Courtin P, Pechoux C, Hols P, Dufrene YF, Kulakauskas S. 2010. Cell surface of *Lactococcus lactis* is covered by a protective polysaccharide pellicle. Journal of biological chemistry 285:10464-71.
45. Liu J, Mushegian A. 2003. Three monophyletic superfamilies account for the majority of the known glycosyltransferases. Protein science 12:1418-1431.
46. Young F, Smith C, Reilly B. 1969. Chromosomal location of genes regulating resistance to bacteriophage in *Bacillus subtilis*. Journal of bacteriology 98:1087-1097.
47. Wendlinger G, Loessner MJ, Scherer S. 1996. Bacteriophage receptors on *Listeria monocytogenes* cells are the N-acetylglucosamine and rhamnose substituents of teichoic acids or the peptidoglycan itself. Microbiology 142:985-992.
48. Spears PA, Havell EA, Hamrick TS, Goforth JB, Levine AL, Abraham ST, Heiss C, Azadi P, Orndorff PE. 2016. *Listeria monocytogenes* wall teichoic acid decoration in virulence and cell-to-cell spread. Molecular microbiology 101:714-730.

Chapter V

Phage resistance conferred by the superinfection exclusion protein $Sie_{2009/LC3}$ and repressor protein Rep_{LC3} of lactococcal P335 phage LC3

Note: Bioinformatic analysis performed in partnership with Dr. Philip Kelleher.

Chapter V contents

5 Abstract

When phage LC3 is present as a lysogenic phage in the chromosome of *Lactococcus lactis* subsp. *cremoris* (*L. lactis*) 3107 it provides resistance against multiple phages belonging to the P335 lactococcal phage group. This phage resistance was shown to be caused by the LC3-prophage encoded superinfection exclusion (Sie) protein $Sie_{2009/LC3,}$ an identical homologue of the previously identified Sie system Sie_{2009}. Furthermore, the LC3-prophage-encoded repressor protein Rep_{LC3} was shown to exhibit superinfection immunity (Sii) against certain phages belonging to the P335 group. Additionally, $Sie_{2009/LC3}$ and other previously identified Sie systems, Sie_{T712}, Sie_{mg2} and Sie_{309} were evaluated for their activity spectrum against 3107-infecting phages belonging to distinct lactococcal phage groups. Results demonstrated that $Sie_{2009/LC3}$, when expressed using a high copy number plasmid and a strong constitutive promoter, elicit a broad activity spectrum in providing resistance against members of the P335, 936 and 949 group phages.

5.1 Introduction

The quality of many fermented dairy products relies on the application of *Lactococcus lactis* starter cultures, though this may be jeopardized by their susceptibility to bacteriophage (phage) infection (1-4). The vast majority of isolated phages infecting industrially important *L. lactis* starter strains (5) either belong to the 936 or c2 phage groups whose members are strictly virulent, or to the P335 phage group which consists of virulent or temperate phages (6, 7). Significant research efforts have been undertaken to define and unravel lactococcal phage-host interactions (8-14), as well as to identify and characterize natural phage defense mechanisms acquired by lactococcal strains (15-21). The expression of specific prophage-encoded genes has been shown to prevent secondary infection by phages in a variety of both Gram-positive and Gram-negative bacteria. Depending on the type of immunity that the lysogenized host confers, these systems can be classified into superinfection immunity (Sii) (22, 23), superinfection exclusion (Sie) (24-26), phage abortive infection (Abi) (27) and adsorption inhibition (Adi) (25, 26, 28). Abi systems use a variety of different mechanisms to achieve killing of the host cell in order to disrupt phage development following phage infection (21, 27). Adi systems alter the cell surface or other cell envelope components thus preventing attachment of homologous phages to the cell (25, 26, 28). Sie proteins interfere with the phage DNA injection process following phage adsorption to the cell (29). Sii systems protect the cell from phage infection by preventing initiation of the lytic life cycle (22, 23).

The first characterized Sii system was the prophage λ-encoded repressor protein which blocks the expression of genes required for lytic development by binding to the P_L and P_R promoters which control expression of CI and Cro, respectively (30-32). Constitutive expression of repressor proteins therefore confers Sii against (typically) homologous phages (22, 23). The repressor of the *Bacillus subtilis* prophage 105 was the first such protein to be characterized in a Gram-positive host (22). The A2 prophage-encoded repressor was also shown to confer self-immunity in lysogenised *Lactobacillus casei* and *Lactobacillus paracasei* strains (23). In addition, a repressor from the lytic P335 phage 31 was shown to confer Sii against a number of related, yet genetically distinct P335 phages including 31, 31.1, 31.2, ul36, ul37, Q30, Q33, Q36, A1, B1, CS and D1 (32).

Examples of prophages encoding Adi systems include the D3 prophage from *Pseudomonas aeruginosa* which harbours a three gene operon encoding proteins that

modify the O-antigen of the cell surface lipopolysaccharide (LPS) of its host (28). This results in an alteration of the D3 receptor, thus preventing this phage to attach to its host and consequently foiling phage infection (33). Similarly, *Escherichia coli* prophages 80 and N15 encode the Cor protein which physically interacts with FhuA, thereby preventing infection by phages that employ FhuA as their receptor (25, 26).

HK97, an *E. coli* prophage, contains gene *15* which encodes an inner membrane protein that blocks (secondary) HK97 DNA entry into the cytoplasm, thus rendering this an Sie system (34). Sie systems encoded by prophages of Gram-positive lactic acid bacteria (LAB) species include the well characterized system Sie_{2009} encoded by the P335 group phage Tuc2009. The sie_{2009} gene is located between the repressor and integrase-encoding genes within the lysogeny module of Tuc2009. Sie_{2009} is a membrane-associated protein that allows phage adsorption to occur while preventing DNA injection of particular phages belonging to the lactococcal 936 group (17). Prophage-encoded Sie systems of lactococcal strains *L. lactis* subsp. *cremoris* MG1363 and *L. lactis* subsp. *lactis* IL1403 exhibit a similar spectrum of activity against (particular) members of the 936 group (35). Interestingly, the *Streptococcus thermophilus* prophage TP-J34 expresses a small lipoprotein, named Ltp, that provides phage resistance by DNA injection blocking against the 936 lytic lactococcal phage P008, as well as self-immunity, suggesting that the molecule targeted by these streptococcal and lactococcal Sie systems is identical (36).

In the current chapter, we show that LC3 lysogens of *L. lactis* subsp. *cremoris* 3107 exhibit resistance against lactococcal phages belonging to the P335 group. The lysogeny module of the temperate lactococcal P335 phage LC3 consists of an integrase- (*int*), putative Sie- ($sie_{2009/LC3}$) and repressor (rep_{LC3})-encoding gene. $Sie_{2009/LC3}$ (homologue of Sie_{2009}) and Rep_{LC3} were assessed for their role in the observed phage-resistance of LC3 lysogens. Furthermore, anti-phage activities of $Sie_{2009/LC3}$ and other previously identified Sie systems Sie_{T712}, Sie_{mg2} and Sie_{309} were also assessed in *L. lactis* 3107 using a high-copy, high-expression vector.

5.2 Materials and Methods

5.2.1 Bacterial strains, growth conditions, phage preparations, and plasmids

Strains, phages and plasmids employed in this study are listed in Table 5.1. *L. lactis* strains were grown at 30 °C in M17 broth/agar (Oxoid, U.K.) supplemented with 0.5 % (v/v) glucose (GM17) and 5 μg ml^{-1} chloramphenicol (Cm), where necessary. Lactococcal phages were propagated in GM17 broth cultured with the host strain at an approximate OD_{600nm} of 0.2 and supplemented with 10 mM $CaCl_2$. SM buffer (10 mM $CaCl_2$, 100 mM NaCl, 10 mM $MgSO_4$, 50 mM Tris-HCl at pH 7.5) was used as the diluent in all phage assays.

5.2.2 Phage assays

5.2.2.1 Phage enumeration assays

To determine the phage sensitivity profile of *L. lactis* 3107 and its derivatives, solid and semi-solid agar was prepared using GM17 broth supplemented with 1.5 % (w/v) or 0.7 % (w/v) bacteriological agar, respectively. 10 mM $CaCl_2$ and 0.5 % glycine (w/v) was also incorporated where necessary, and solid and semi-solid agar media were employed according to a previously described method (37). To perform spot assays, an overlay was created by seeding semi-solid agar with the relevant indicator/test strain and spotting 10 μl of 10^8 pfu/ml phage lysate on the overlay. The plates were incubated overnight at 30 °C. Plaque assays were performed by creating an overlay by seeding semi-solid agar with the corresponding indicator/test strain and 10 μl of 10^8 pfu/ml phage lysate from 10-fold phage dilutions that had been prepared in SM buffer. The plates were incubated overnight at 30 °C. The efficiency of plaquing (E.O.P.) was determined as the ratio of the titre obtained from the test strain to that of the control strain (*L. lactis* 3107).

Table 5.1: Bacteria, phages, and plasmids used in this study.

	Description/purpose	Source
Bacterial Strains		
L. lactis subsp. *cremoris*		
3107	Host to TP901-1, LC3, Dub35A, C41431, 58601, 63301, 86501, 50101, 62604, 62601, 66901, 66902, 66903, 949 and WRP3	(38)
3107A / B / C	*L. lactis* 3107 derivatives; lysogenized LC3 in its genome.	This study
NZ9000	MG1363 derivative in which *nisRK* are inserted in the chromosome. Negative control; not infected by 3107-infecting phages.	(39)
Escherichia coli EC101	Cloning host	(40)
Bacteriophages		
TP901-1	Temperate P335 group phage induced from *L. lactis* subsp. *cremoris* 901-1	(41)
LC3	Temperate P335 group phage grown lytically on *L. lactis* 3107	(42)
Dub35A	P335 phage group	(43)
63301	P335 phage group	(43)
58601	P335 phage group	(43)
86501	P335 phage group	(43)
50101	P335 phage group	(43)
C41431	P335 phage group	(43)
62604	936 phage group	(44)
62601	936 phage group	(44)
66901	936 phage group	(44)
66902	936 phage group	(44)
66903	936 phage group	(44)
949	949 phage group	(45)
WRP3	949 phage group	(46)
Plasmids		
pNZ44	*E. coli–L. lactis* high copy number shuttle vector containing the constitutive P44 promoter; Cm^r	(47)
pNZrep_{LC3}	pNZ44 vector harbouring repressor gene, *rep*, from LC3	This study
pNZsie_{2009}	pNZ44 vector harbouring the *sie* gene from *L. lactis* subsp. *cremoris* UC509 prophage Tuc2009	(35)
pNZsie_{T712}	pNZ44 vector harbouring the *sie* gene from *L. lactis* subsp. *cremoris* MG1363 prophage T712	(35)
pNZsie_{mg2}	pNZ44 vector harbouring the *sie* gene from *L. lactis* subsp. *cremoris* MG1363 prophage MG2	(35)
pNZsie_{309}	pNZ44 vector harbouring the *sie* gene from *L. lactis* subsp. *lactis* IL1403 prophage bIL309	(35)

5.2.2.2 Reversible and irreversible adsorption assays

Reversible and irreversible adsorption assays were performed as previously described (48) with the following modifications. Briefly, 10 mL GM17 broth was inoculated with 2 % of the appropriate strain until an approximate OD_{600nm} of 0.5 was reached. The culture was supplemented with 10 mM $CaCl_2$, following this 700 µl of the mixture was added to an equal volume of phage lysate (at a multiplicity of infection [MOI] of 0.01), and the mixture was incubated at 30 °C for 10 minutes. As a control 700 µl GM17 broth supplemented with 10 mM $CaCl_2$ was mixed with 700 µl of the phage lysate and incubated at 30 °C for 10 minutes. Following incubation, irreversible adsorption was determined by diluting samples 1:100 in ice-cold quarter-strength Ringer's solution supplemented with 1 M NaCl. Cell-free supernatants were prepared by centrifugation of the samples at 6,000 × *g* for 1 minute and phage numbers in the supernatants were determined by standard plaque assays using *L. lactis* 3107 as the indicator strain. Following incubation, reversible adsorption was also determined, samples were centrifuged at 6,000 × *g* for 1 minute to pellet cells and the supernatant was diluted 1:100 in quarter strength Ringer's solution supplemented with 1 M NaCl. Phage numbers in the supernatant were determined by standard plaque assays using *L. lactis* 3107 as the host strain. Adsorption was measured as percentage of total phages attached to the host when compared to control phage titre (which lacks bacterial cells) using the following formula: [(Control phage titre – Free phage titre in supernatant) / Control phage titre] x 100.

5.2.3 Generation of *L. lactis* subsp. *cremoris* 3107 mutants resistant to phage LC3.

A single colony of *L. lactis* 3107 was cultured overnight at 30 °C, this culture was stocked in 20 % (v/v) glycerol and stored at -20 °C for further DNA extraction (see section 5.2.4 below for details). *L. lactis* 3107 was subcultured in 10 mL GM17 broth. When the culture reached an OD_{600nm} of 0.2 it was supplemented with 10 mM $CaCl_2$ and 200 µl of a 10^8 pfu/ml LC3 lysate. Every 30 minutes for two hours, 100 µl of the culture was spread-plated on GM17 agar plates following a series of 10-fold dilutions that had been prepared in GM17 broth. The plates were then incubated at 30 °C for 48 h. 200 colonies from each time point were cultured overnight and spot assays were performed against phages LC3 and TP901-1. Isolates that displayed resistance to LC3 and sensitivity to TP901-1 were cultured overnight and stocked in 20 % (v/v) glycerol at -20 °C until required for further analysis. The LC3-resistant isolates were

verified as *L. lactis* 3107 derivatives using multiplex-PCR primers (Table 5.2) which target each of the six 3107 plasmids (described in Chapter II). Following amplification, agarose gel electrophoresis using a 1 % agarose gel (49) was performed to visualise the generated amplicons with each product corresponding to one of the six 3107 plasmids.

Table 5.2: Primers used in this study; restriction sites, where included, are underlined.

Name	Sequence (5' – 3')	Product Size (bp)	Target	Source
p3107_AF	GGAAGGTGGCAGAGCATA	857	3107 plasmid pL3107A	Chapter II
p3107_AR	CGTGTCCTGGTTTGCTAT			
p3107_BF	GGGAGTGAACCAGTAACA	740	3107 plasmid pL3107B	Chapter II
p3107_BR	CCAGGAGAAGAAGTACGA			
p3107_CF	CGTCCCAATACCAACTGT	417	3107 plasmid pL3107C	Chapter II
p3107_CR	GCGGTTCGTTTCTTCTCA			
p3107_DF	GCGTATCTATGGCTGTCA	228	3107 plasmid pL3107D	Chapter II
p3107_DR	GCGTCCTTTGATTCATGAG			
p3107_EF	GCCGTGTCTTGCATTGAT	1,012	pL3107E	Chapter II
p3107_ER	GGCTTGCACCTTTACCTC			
p3107_FF	CCGAAGAGCCACCGAAAG	616	3107 plasmid pL3107F	Chapter II
p3107_FR	GCTTGTAGCTGTTGTCGT			
rep_{LC3}_F	AGCAGCCTGCAGAGGAGGCACTCAATGGAAATTGAACAAATC	861	*rep* gene from LC3	This study
rep_{LC3}_R	AGCAGCTCTAGATTATAAATCAATTTCTAC			
pNZ44_F	CTAATGTCACTAACCTGCCCCG	500	p44 lactococcal promoter and multiple coning site (MCS) of pNZ44	This study
pNZ44_R	GCTTTATCAACTGCTGCT			

5.2.4 Genome sequencing, assembly and bioinformatic analysis.

L. lactis subsp. *cremoris* 3107 and its LC3-resistant derivatives 3107A, 3107B, and 3107C were grown overnight in GM17 broth. 5 µg chromosomal DNA from *L. lactis* 3107 was extracted using phenol-chloroform based extractions as previously described (50). The chromosomal DNA extracted from each strain was sequenced by the commercial sequencing service provider Probiogenomics (University of Parma, Italy) using an Illumina MiSeq platform. Genomic libraries were constructed using the TruSeq DNA PCR-Free LT Kit (Illumina®) and 2.5 µg of genomic DNA, which was fragmented with a Bioruptor NGS ultrasonicator (Diagenode, USA) followed by size evaluation using Tape Station 2200 (Agilent Technologies). Library samples were loaded into a Flow Cell V3 600 cycles (Illumina®) and draft genome Illumina sequencing was performed on a MiSeq genomic platform (Illumina, UK) at GenProbio srl (Parma, Italy). Fastq files of the paired-end reads obtained from the genome sequencing were used as input for genome assemblies through the MEGAnnotator pipeline (51). The MIRA program (version 4.0.2) was used for *de novo* assembly of genome sequence data (52). Following final genome assembly, putative protein-encoding genes were identified using the prediction software Prodigal (version 2.0) (53). Protein-encoding genes were automatically annotated using a BLASTP v2.2.26 (cut-off *E*-value of 0.0001) sequence alignments against the non-redundant protein (nr) database curated by NCBI (ftp://ftp.ncbi.nih.gov/blast/db/).

The complete genome sequence of *L. lactis* 3107 has previously been determined (54). Here, the *L. lactis* 3107 culture from which the phage-resistant derivatives were generated (as described in section 5.2.3) was sequenced to verify the 3107 published genome sequence, thus creating an updated 3107 reference genome. This was performed in order to ensure the integrity of mutations identified on the genomes of the *L. lactis* 3107 derivatives. Subsequently, using the multiple genome alignment software Mauve v2.4.0 (55) a comparative genome analysis of *L. lactis* 3107 (i.e. the reference genome) and its derivatives was undertaken in order to identify the genomic changes that may have caused the observed phage-resistance. Mauve v2.4.0 (55) was used to identify the position of the integrated LC3 prophage within the chromosome of *L. lactis* 3107. The LC3 genetic *att* elements (attachment sites) had previously been identified as 5'-TTCTTCATG'-3 (56). Consequently, using Artemis v16 (57) the presence of this common region within the 3107 chromosome and inserted LC3 genome was identified.

5.2.5 Construction of pNZ*rep*$_{LC3}$

Primers (Table 5.2) were generated to amplify the gene *LC3p03* from LC3, here renamed *rep*$_{LC3}$ (861 bp). Primers *rep*$_{LC3}$_F and *rep*$_{LC3}$_R, were designed to contain the PstI and XbaI recognition sequence, respectively. The generated amplicon and the high copy, high expression, Cm^r plasmid pNZ44 (3,396 bp) were restricted with PstI and XbaI. Restriction enzyme digestions, DNA ligation and agarose gel electrophoresis using a 1 % agarose gel were performed following standard procedures (49). The ligated construct was introduced into electrocompetent *E. coli* EC101 cells prepared using a previously described protocol (40). *E. coli* EC101 transformants expected to harbour the recombinant vector were selected using LB agar plates (40) supplemented with 5 µg ml^{-1} Cm. Transformants containing the desired recombinant plasmid, designated pNZ*rep*$_{LC3}$ were identified by PCR using primers pNZ44_F and _R (Table 5.2). The sequence integrity of the generated construct was determined by Sanger sequencing of the inserted sequence (performed by MWG Eurofins Genomics, Ebersberg, Germany) and subsequent analysis using the Seqman Pro program (DNAstar v7.0) (58). The verified recombinant plasmid was isolated using a Pure Link DNA Purification (Thermofisher Scientific) kit and introduced into competent cells of *L. lactis* 3107 which were prepared and transformed as described in Chapter II. To confirm the presence of the desired construct, PCR using insert-specific primers (Table 5.2) were performed on the resulting Cm^r isolates.

5.3 Results

5.3.1 Isolation of *L. lactis* 3107 mutants resistant to phage LC3

The dairy starter strain *L. lactis* 3107 is sensitive to the temperate lactoccocal phage LC3. In an attempt to obtain LC3-resistant derivatives of 3107, *L. lactis* 3107 was exposed to LC3. Following LC3 exposure, surviving cells were plated and, using spot assays, eight hundred colonies were then tested for LC3 resistance. Three such isolates, designated here as 3107A, 3107B and 3107C, were shown to possess the desired LC3 phage-resistance phenotype, while they had remained sensitive to infection by the P335 phage TP901-1. As described in Chapter II, *L. lactis* 3107 harbours six plasmids and their associated sequences were used to develop a multiplex PCR system to amplify specific regions of each plasmid. Thus, using this multiplex PCR system on strains 3107A, 3107B and 3107C, these isolates were confirmed to be 3107 derivatives as their plasmid profiles were shown to be identical to that of *L. lactis* 3107 (Fig. 5.1).

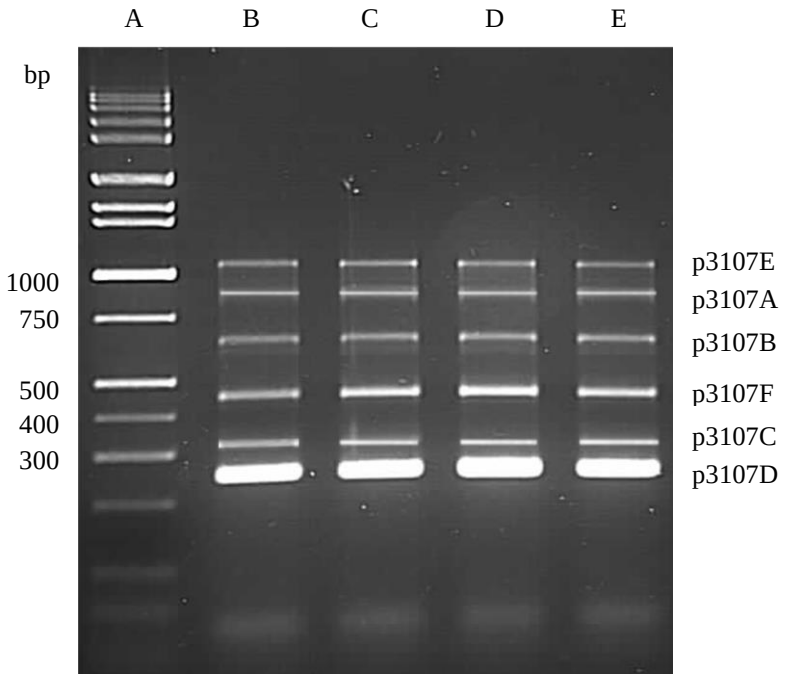


Figure 5.1: Agarose gel image of amplicons generated following multiplex-PCR targeting *L. lactis* 3107 plasmids. Lane A represents a DNA Ladder (relevant sizes are indicated; Fisher BioReagents, Ireland); Lanes B, C, D and E show amplicons from *L. lactis* strains 3107, 3107A, 3107B and 3107C, respectively, representing each of the six *L. lactis* 3107 plasmids: p3107E (amplicon size 1,012 bp), p3107A (amplicon size 857 bp), p3107B (amplicon size 740 bp), p3107F (amplicon size 616 bp), p3107C (amplicon size 417 bp) and p3107D (amplicon size 228 bp).

5.3.2 Genome sequencing and bioinformatic analysis of *L. lactis* 3107 derivatives 3107A, 3107B and 3107C

As outlined above, screening of *L. lactis* 3107 survivors following exposure of this strain to phage LC3 resulted in the isolation of three LC3-resistant 3107 derivatives. The genomes of *L. lactis* 3107A, 3107B and 3107C were sequenced in order to identify which genomic changes were responsible for the observed phage-resistance phenotype. Previous genome analysis of *L. lactis* 3107 revealed a genome length of 2,378,982 bp and a G+C content of 36 %, in addition to six plasmids ranging in size from 2,232 to 60,216 bp (54). Genome analysis of the three sequenced 3107 derivatives revealed different characteristics to that of the parent strain while carrying an identical plasmid complement. The genomes of the three 3107 derivatives are 2,411,154 ± 1,080 bp in length indicating a genomic insertion. Comparative genome analysis using the multiple genome alignment software Mauve v2.4.0 (55) was undertaken to further assess the genomic insertion present in the genomes of 3107A, 3107B and 3107C. This comparative genomic analysis identified the 32,172 ± 1,080 bp insertion to be present at positions 1,295,942-1,295,950 on the 3107A/3107B/3107C chromosome, representing the complete LC3 prophage genome (42). Previously the attachment site for integration of LC3 into the chromosome of *L. lactis* subsp. *cremoris* IMN-C1819 was identified as 5'-TTCTTCATG'-3 (56). Here, following the identification of the LC3 position in the 3107 chromosome, the same 9-bp common core region or attachment site (*attP* or *attB*: 5'-TTCTTCATG-3') was identified in the LC3 and 3107 genomes. This common core region is therefore assumed to have facilitated insertion of the LC3 genome into the 3107 chromosome. The *attB* sequence is located within *L3107_1090* which is predicted to encode a redox-sensing transcriptional repressor Rex (56). Furthermore, the LC3 integrase-encoding gene, which specifies the LC3 site-specific recombinase, is located adjacent to *attP*. Based on these findings we conclude that the observed LC3-resistance phenotype in *L. lactis* strains 3107A, 3107B and 3107C is due to the presence of an integrated LC3 genome and that these three strains therefore represent *L. lactis* 3107 lysogens.

5.3.3 Assessment of the phage-resistance spectrum of the *L. lactis* 3107 lysogens

In order to assess the phage-resistance profile of *L. lactis* 3107A, 3107B and 3107C, quantitative plaque assays were performed using a selection of lactoccocal phages that infect 3107 (Table 5.3). The E.O.P. of the lactococcal phages belonging to

the P335 (LC3, TP901-1, Dub35A, C41431, 50101, 58601, 63301 and 86501), 936 (62604, 62601, 66901, 66902 and 66903), and the 949 groups (949 and WRP3) were tested on 3107 (as a control) and the lysogenised derivatives (Table 5.3). As before, LC3 was unable to form visible plaques on the lysogens. Interestingly, complete phage resistance was also observed for the P335 group phages 50101, 58601 and 86501. Furthermore, phages Dub35A, C41431 and 63301 were shown to infect the LC3 lysogens of 3107, although a low, yet consistent level of phage resistance was observed as the E.O.P. of these phages exhibited a statistically significant reduction (p-values < 0.05) and an obvious visual reduction in plaque size in comparison to the 3107 strain (Table 5.3). These results indicate that the infection pathway of these lactococcal phages is negatively affected by LC3 prophage-encoded phage-resistance mechanism(s). Whether this is due to an Sie or Sii mechanism will be further investigated below (section 5.4). No obvious changes in E.O.P. were observed for the assessed 936 group phages, P335 group phage TP901-1, nor for 949 group phages 949 and WRP3, though a reduction in plaque size was observed for these lactococcal phages, except in the case of WRP3 (Table 5.3).

Table 5.3: Testing P335, 936 and 949 phages on *L. lactis* 3107 and its derivatives; the E.O.P. of phage applied is shown.

	L. lactis subsp. *cremoris*		3107	3107A	3107B	3107C
P335 phages (subgroup)	IIIA	LC3	1 (1.5 mm)	$\leq 10^{-8}$	$\leq 10^{-8}$	$\leq 10^{-8}$
	IIIA	50101	1 (1.5 mm)	$\leq 10^{-8}$	$\leq 10^{-8}$	$\leq 10^{-8}$
	IIIA	63301	1 (1.5 mm)	0.07 ± 0.02 (0.5 mm)	0.08 ± 0.06 (0.5 mm)	0.02 ± 0.12 (0.5 mm)
	IIIA	8650	1 (1 mm)	$\leq 10^{-8}$	$\leq 10^{-8}$	$\leq 10^{-8}$
	III	Dub35A	1 (1.5 mm)	0.05 ± 0.19 (0.5 mm)	0.09 ± 0.12 (0.5 mm)	0.08 ± 0.11 (0.5 mm)
	IIA	TP901-1	1 (1 mm)	0.83 ± 0.47 (0.8 mm)	0.84 ± 0.49 (0.8 mm)	0.88 ± 0.28 (0.8 mm)
	I	C41431	1 (1.5 mm)	0.06 ± 0.08 (0.5 mm)	0.09 ± 0.06 (0.5 mm)	0.01 ± 0.02 (0.5 mm)
	ND	58601	1 (1.5 mm)	$\leq 10^{-8}$	$\leq 10^{-8}$	$\leq 10^{-8}$
936 phages		62601	1 (4.0 mm)	0.86 ± 0.03 (2.5 mm)	0.85 ± 0.10 (2.5 mm)	0.82 ± 0.07 (2.5 mm)
		62604	1 (2.0 mm)	0.86 ± 0.07 (1.0 mm)	0.82 ± 0.02 (1.0 mm)	0.88 ± 0.04 (1.0 mm)
		66901	1 (3.0 mm)	0.80 ± 0.16 (2.0 mm)	0.89 ± 0.10 (2.0 mm)	0.81 ± 0.14 (2.0 mm)
		66902	1 (4.0 mm)	0.85 ± 0.10 (3.0 mm)	0.84 ± 0.10 (3.0 mm)	0.85 ± 0.05 (3.0 mm)
		66903	1 (4.0 mm)	0.89 ± 0.41 (3.0 mm)	0.86 ± 0.60 (3.0 mm)	0.83 ± 0.42 (3.0 mm)
949 phages		949	1 (1.0 mm)	0.76 ± 0.19 (0.5 mm)	0.71 ± 0.02 (0.5 mm)	0.77 ± 0.08 (0.5 mm)
		WRP3	1 (1.0 mm)	0.81 ± 0.08 (1.0 mm)	0.81 ± 0.22 (1.0 mm)	0.85 ± 0.16 (1.0 mm)

ND: not determined

5.3.4 Contribution of the LC3 prophage-contained genes *sie*$_{2009/LC3}$ and *rep*$_{LC3}$, to phage resistance.

The *LC3p02* gene, here renamed *sie*$_{2009/LC3}$ for clarity purpose, is 100 % identical in sequence to *sie*$_{2009}$, being 522 bp in length, specifying a polypeptide of 173 amino acids, and predicted to encode a protein harboring a single amino-terminal transmembrane domain (17). Furthermore, similar to *sie*$_{2009}$, *sie*$_{2009/LC3}$ is located between the genes encoding the integrase and repressor on the lysogeny module of LC3. Sie$_{2009}$ has previously been shown to confer resistance against particular members of the 936 phage group and given the identical sequence of Sie$_{2009}$ to Sie$_{2009/LC3}$, it was expected that Sie$_{2009}$ exerts the same activity spectrum as Sie$_{2009/LC3}$. Additionally, upstream of *sie*$_{LC3}$, is a repressor-encoding gene *rep*$_{LC3}$ specifying a polypeptide of 286 amino acids. Repressor proteins have previously been shown to confer resistance against homologous phages by superinfection immunity (22, 23).

To determine if expression of *sie*$_{2009/LC3}$ (sequence 100 % identical to *sie*$_{2009}$) or *rep*$_{LC3}$ is responsible for the phage resistance phenotype observed in the three *L. lactis* 3107-derived LC3 lysogens, the previously constructed plasmid pNZ*sie*$_{2009}$ (35) and newly constructed pNZ*rep*$_{LC3}$ were introduced into 3107. pNZ*sie*$_{2009}$ and pNZ*rep*$_{LC3}$ are expected to overexpress Sie$_{2009/LC3}$ and Rep$_{LC3}$, respectively, due to the high copy number and strong, constitutive P44 promoter of pNZ44 (47). Subsequently, the E.O.P. values of P335 (TP901-1, Dub35A, C41431, 50101, 58601, 63301 and 86501), 936 (62604, 62601, 66901, 66902 and 66903) and 949 group phages (949 and WRP3) were determined for the Sie- and repressor-expressing strains and compared to the E.O.P. of these phages on the 3107 strain (Table 5.4). These phages were also tested against *L. lactis* 3107A harbouring the empty pNZ44 plasmid as a control (Table 5.4).

Expression of *rep*$_{LC3}$ in *L. lactis* 3107 confers resistance against certain P335 group phages as complete resistance against LC3, 50101, 58601, and 86501 was observed (Table 5.4). As expected, Rep$_{LC3}$ elicited no protective effect against tested 936 or 949 group phages (Table 5.4). The expression of Sie$_{2009/LC3}$ exerts complete phage resistance against all of the 3107-infecting phages tested, except for TP901-1 although with a reduced E.O.P. exhibiting a small (10-fold) but statistically significant reduction (p-values < 0.05) in comparison to 3107 harbouring the empty pNZ44 vector (Table 5.4). Thus the low level of phage resistance observed for 3107A when exposed to P335 phages Dub35A, C41431 and 63301 is likely due to activity of Sie$_{2009/LC3}$

(Table 5.4), which is likely expressed at a much lower level in the lysogen as compared to when the *sie*$_{2009/LC3}$ gene is placed under control of the p44 promoter in the multicopy plamid pNZ44. When 936 phages 62604, 62601, 66901, 66902 and 66903, and 949 phages 949 and WRP3 were used to infect 3107A using a plaque assay, a reduction in plaque size was noticed (except for WRP3), yet no phage resistance was observed. However, (over)expression of *sie*$_{2009/LC3}$ when present on plasmid pNZ*sie*$_{2009}$ in 3107 confers complete resistance against all these lactococcal phages (Table 5.4). More than likely the higher level of expression of Sie$_{2009/LC3}$ from the pNZ44 plasmid in comparison to that in 3107A is responsible for the increased level of phage-resistance activity (Table 5.4).

5.3.5 Phage-resistance exerted by Sie$_{T712}$, Sie$_{mg2}$ and Sie$_{309}$ against representatives of the P335, 936 and 949 lactococcal phage groups.

Like Sie$_{2009,}$ previously identified lactococcal prophage-encoded Sie systems including Sie$_{T712}$, Sie$_{mg2}$ and Sie$_{309}$ have demonstrated activity against certain 936 group phages capable of infecting *L. lactis* subsp. *cremoris* NZ9000 (35). To determine if the (over)expression of these Sie proteins exert phage resistance against 3107-infecting phages, plasmids pNZ*sie*$_{T712}$, pNZ*sie*$_{mg2}$ and pNZ*sie*$_{309}$, which (over)express Sie$_{T712}$, Sie$_{mg2}$ and Sie$_{309}$, respectively, were introduced into *L. lactis* 3107. Subsequently, the E.O.P. of P335 (TP901-1, Dub35A, C41431, 50101, 58601, 63301 and 86501), 936 (62604, 62601, 66901, 66902 and 66903) and 949 group phages (949 and WRP3) was determined for these strains (Table 5.4), as well as for *L. lactis* 3107 harbouring the empty pNZ44 plasmid as a control.

Sie$_{mg2}$ is active against the 936 phages 66903, 66901 and 66902. Sie$_{mg2}$ confers complete resistance against phage 66903, while eliciting a remarkably low level of resistance (11-12-fold) against the 936 group phages 66901 and 66902 in comparison to 3107 harbouring the empty pNZ44 vector (Table 5.4). This differs from Sie$_{T712}$ and Sie$_{309}$ which have no impact on the infectivity of any of the 3107-infecting phages tested (Table 5.4).

Table 5.4: Testing P335, 936, and 949 phages on *L. lactis* 3107 and its derivatives; the E.O.P. of phage applied is shown.

	L. lactis subsp. *cremoris*		3107	3107_pNZ44	3107A_pNZ44	3107_pNZ*rep*$_{LC3}$	3107_*pNZsie*$_{2009}$	3107_pNZ*sie*$_{1712}$	3107_pNZ*sie*$_{mg2}$	3107_pNZ*sie*$_{309}$
P335 phages (subgroup)	IIIA	LC3	1 (1.5 mm)	0.87 ± 0.05 (1.5 mm)	$\leq 10^{-8}$	$\leq 10^{-8}$	$\leq 10^{-8}$	0.96 ± 0.55 (1.5 mm)	1.38 ± 0.48 (1.5 mm)	0.98 ± 0.33 (1.5 mm)
	IIIA	50101	1 (1.5 mm)	0.87 ± 0.05 (1.5 mm)	$\leq 10^{-8}$	$\leq 10^{-8}$	$\leq 10^{-8}$	0.96 ± 0.55 (1.5 mm)	1.38 ± 0.48 (1.5 mm)	0.98 ± 0.33 (1.5 mm)
	IIIA	63301	1 (1.5 mm)	0.83 ± 0.04 (1.5 mm)	0.04 ± 0.05 (0.5 mm)	1.87 ± 0.58 (1.5 mm)	$\leq 10^{-8}$	1.16 ± 0.18 (1.5 mm)	1.08 ± 0.93 (1.5 mm)	1.75 ± 1.11 (1.5 mm)
	IIIA	86501	1 (1 mm)	0.87 ± 0.02 (1 mm)	$\leq 10^{-8}$	$\leq 10^{-8}$	$\leq 10^{-8}$	1.14 ± 0.58 (1 mm)	1.37 ± 0.63 (1 mm)	2.01 ± 0.58 (1 mm)
	III	Dub35A	1 (1.5 mm)	0.95 ± 0.06 (1.5 mm)	0.06 ± 0.13 (0.5 mm)	0.99 ± 0.58 (1.5 mm)	$\leq 10^{-8}$	1.26 ± 0.21 (1.5 mm)	0.99 ± 0.90 (1.5 mm)	0.45 ± 0.18 (1.5 mm)
	IIA	TP901-1	1 (1 mm)	0.80 ± 0.02 (1 mm)	0.74 ± 0.06 (0.8 mm)	1.02 ± 0.53 (1 mm)	0.02 ± 0.01 (0.8 mm)	0.43 ± 0.26 (1 mm)	0.39 ± 0.11 (1 mm)	0.85 ± 0.13 (1 mm)
	I	C41431	1 (1.5 mm)	0.84 ± 0.02 (1.5 mm)	0.03 ± 0.01 (0.5 mm)	1.07 ± 0.65 (1.5 mm)	$\leq 10^{-8}$	1.87 ± 0.17 (1.5 mm)	1.18 ± 0.45 (1.5 mm)	1.16 ± 0.59 (1.5 mm)
	ND	58601	1 (1.5 mm)	0.93 ± 0.01 (1.5 mm)	$\leq 10^{-8}$	$\leq 10^{-8}$	$\leq 10^{-8}$	1.07 ± 0.7 (1.5 mm)	1.68 ± 0.94 (1.5 mm)	1.58 ± 0.61 (1.5 mm)
936 phages	62601		1 (4 mm)	0.81 ± 0.03 (4 mm)	0.82 ± 0.10 (2.5 mm)	1.52 ± 1.22 (4 mm)	$\leq 10^{-8}$	1.13 ± 0.61 (4 mm)	1.77 ± 0.43 (4 mm)	2.12 ± 1.09 (4 mm)
	62604		1 (2 mm)	0.87 ± 0.01 (2 mm)	1.00 ± 0.30 (1 mm)	1.95 ± 0.95 (2 mm)	$\leq 10^{-8}$	1.08 ± 0.05 (2 mm)	1.89 ± 0.93 (2 mm)	1.56 ± 0.80 (2 mm)
	66901		1 (3 mm)	0.70 ± 0.02 (3 mm)	0.87 ± 0.06 (2 mm)	1.20 ± 0.41 (3 mm)	$\leq 10^{-8}$	1.82 ± 0.25 (3 mm)	$3.40 \times 10^{-4} \pm 1.76 \times 10^{-4}$ (1 mm)	1.55 ± 0.99 (3 mm)
	66902		1 (4 mm)	0.92 ± 0.06 (4 mm)	0.84 ± 0.34 (3 mm)	1.22 ± 0.21 (4 mm)	$\leq 10^{-8}$	1.65 ± 0.91 (4 mm)	$3.76 \times 10^{-5} \pm 2.90 \times 10^{-5}$ (1.5 mm)	1.90 ± 0.49 (4 mm)
	66903		1 (4 mm)	0.93 ± 0.04 (4 mm)	0.84 ± 0.17 (3 mm)	1.16 ± 1.59 (4 mm)	$\leq 10^{-8}$	1.93 ± 0.91 (4 mm)	$\leq 10^{-8}$	1.67 ± 0.44 (4 mm)
949 phages	949		1 (1 mm)	0.86 ± 0.04 (1 mm)	0.84 ± 0.44 (0.5 mm)	1.68 ± 0.28 (1 mm)	$\leq 10^{-8}$	1.35 ± 0.20 (1 mm)	1.82 ± 1.02 (1 mm)	1.89 ± 0.60 (1 mm)
	WRP3		1 (1 mm)	0.81 ± 0.10 (1 mm)	0.77 ± 0.64 (1 mm)	1.92 ± 0.96 (1 mm)	$\leq 10^{-8}$	1.25 ± 0.51 (1 mm)	1.31 ± 0.08 (1 mm)	1.31 ± 0.18 (1 mm)

ND: not determined

5.3.6 Effect of LC3 prophage-encoded Sie$_{LC3}$ and Rep$_{LC3}$ on phage adsorption

LC3-prophage expressed genes (such as *sie*$_{2009/LC3}$ and *rep*$_{LC3}$) may interfere with the first stages of infection, such as adsorption of the phage to the host, or the DNA injection process of the phage. In order to determine if lysogenized LC3 interferes with adsorption or a subsequent phage infection step, reversible and irreversible adsorption assays were perfomed on *L. lactis* 3107, its derivative 3107A, and *L. lactis* NZ9000 as a negative control (Table 5.5). Adsorption assays were performed using the 3107-infecting lactoccocal phages which are not able to infect the 3107-derived LC3-carrying lysogens (Table 5.3). These phages included the P335 group phages LC3, 50101, 58601 and 86501. The results show that all of these lactococcal phages adsorb reversibly and irreversibly to *L. lactis* 3107 and its LC3 lysogenised derivative 3107A with a high efficiency. Therefore, as expected, the presence of the LC3 prophage does not appear to have a significant impact on the reversible or irreversible adsorption of these phages to their host.

Table 5.5: Adsorption (%) of P335 group phages to *L. lactis* 3107 and its derivatives.

	L. lactis subsp. *cremoris*	Adsorption (%)	3107	3107A	NZ9000
Adsorption of P335 phages	LC3	R	99.30 ± 0.45	97.69 ± 1.92	19.12 ± 1.40
		I	97.42 ± 0.92	96.82 ± 1.31	14.18 ± 1.92
	50101	R	85.00 ± 1.13	84.58 ± 1.40	15.48 ± 1.85
		I	84.39 ± 1.72	83.87 ± 1.29	10.75 ± 1.62
	58601	R	84.17 ± 1.40	83.98 ± 1.29	12.48 ± 1.33
		I	83.24 ± 1.33	82.87 ± 1.19	8.75 ± 1.40
	86501	R	86.42 ± 1.19	85.98 ± 1.90	16.48 ± 1.72
		I	85.17 ± 1.92	84.87 ± 1.8	9.82 ± 1.18

R = reversible, I = irreversible

5.4 Discussion

The temperate phage LC3 originates from the dairy-associated *L. lactis* subsp. *cremoris* strain IMN-C1819 (42) and utilizes the common core region 5'-TTCTTCATG'-3 for site-specific integration into the chromosome of its host (56). In this chapter the isolation of three LC3 lysogens of *L. lactis* 3107 is described. These lysogens do not only provide resistance against LC3 itself, but are also resistant to other 3107-infecting phages belonging to the P335 lactococcal phage group. Available whole genome sequences of temperate phages and strains carrying prophages have permitted the identification of Sie- and repressor-encoding genes that confer phage-resistance (17, 22, 23, 35). LC3 encodes the Sie protein $Sie_{2009/LC3}$ which was previously shown to be an identical homologue of Sie_{2009} (35). In addition, LC3 harbours rep_{LC3} directly upstream of $sie_{2009/LC3}$ in its lysogeny module. Here, $Sie_{2009/LC3}$ and Rep_{LC3} were shown to be responsible for the phage-resistance phenotype observed in *L. lactis* 3107 lysogens of LC3. Sie proteins have been shown to interfere with phage DNA injection (35), whereas repressor proteins maintain lysogeny and prevent lytic proliferation of certain superinfecting phages (59). Thus, as expected, adsorption assays (Table 5.5) demonstrated that $Sie_{2009/LC3}$ and Rep_{LC3} proteins interfere with a step in the phage cycle that is beyond host adsorption (17, 22, 23).

$Sie_{2009/LC3}$ was shown to confer complete phage resistance activity against all tested phages belonging to the P335 group (LC3, Dub35A, C41431, 50101, 58601, 63301 and 86501), except for TP901-1. The genetically diverse P335 lactococcal group phages have been classified into nine subgroups (IA/B/C, IIA/B, IIIA/B, IV and V) based on their distinct morphological features and genetic lineages of their tail tip region which is presumed to be involved in interactions with the bacterial host (10, 60). TP901-1 belongs to the P335 subgroup IIA; LC3, Dub35A, 50101, 63301 and 86501 belong to subgroup IIIA, C41431 belongs to subgroup I (60), while the genome of 58601 has not been sequenced to date. $Sie_{2009/LC3}$ appears to be quite effective against P335 subgroup IIIA and I phages, while it did not provide significant levels of resistance against subgroup IIA phages indicating that there may be related DNA injection/infection pathways between subgroup I and III phages that is distinct to those of subgroup II phages, which would be consistent with previous studies (48). The level of phage-resistance has previously been shown to be dependent on the level of expression of Sie proteins (35). This is consistent with the finding that LC3 lysogens of 3107, which carry a single genomic copy of $sie_{2009/LC3}$, exhibit a low level of

resistance against the P335 phages Dub35A, C41431 and 63301. In contrast, plasmid-driven expression of $Sie_{2009/LC3}$ in *L. lactis* 3107 (with $sie_{2009/LC3}$ located on a multi-copy plasmid and transcribed by the strong constitutive p44 promoter) was shown to elicit complete phage resistance against these lactococcal phages.

In addition, $Sie_{2009/LC3}$ was also shown to provide complete phage resistance activity against all tested 936 (62601, 62604, 66901, 66902 and 66903) and 949 group phages (949 and WRP3). This data is comparable with previously reported anti-phage activity of Sie_{2009} against specific phage members of the 936 group (35). Interestingly, LC3 lysogens of 3107 do not elicit resistance against any of the 936 or 949 lactococcal phages tested. This may be consistent with the observation that the level of expression of Sie proteins is positively correlated to the level of phage-resistance activity (35).

Like Sie_{2009}, other lactococcal prophage-encoded Sie systems such as Sie_{T712}, Sie_{mg2} and Sie_{309} have previously been shown to exert activity against a particular subset of the 936 phage group (35). Interestingly, in the current study Sie_{T712} and Sie_{309} did not mediate a phage-resistance phenotype against any of the 3107-infecting phages, whereas Sie_{mg2} was shown to exert phage resistance against 936 group phages 66901, 66902 and 66903. These results demonstrate that the Sie proteins Sie_{T712}, Sie_{309} and Sie_{mg2} are active against specific members of the 936 phage group as previously observed (17, 35).

Here, the LC3 repressor protein Rep_{LC3} was also tested to assess its phage resistance activity against 3107-infecting lactococcal phages. Rep_{LC3} prevented superinfection of phages LC3, 50101, and 86501 which belong to the P335 subgroup IIIA, as well as the P335 phage 58601 (unknown subgroup). Lactococcal phages 50101 and 86501 harbour a repressor gene which is 99 % identical to that of rep_{LC3}, thereby explaining the observed immunity. The rep_{LC3} gene is presumed to be constitutively expressed to maintain the lysogenic state of the temperate LC3 phage genome in 3107. The Rep_{LC3} protein is presumed to provide superinfection immunity against LC3 itself and specific members of the P335 lactococcal group which encode a near identical Rep repressor. The phages that were unaffected by Rep_{LC3} and for which the genome sequence is available (TP901-1, Dub35A, C41431, 63301 and WRP3) do indeed not harbour a close homologue of rep_{LC3}, thereby explaining the lack of immunity.

In conclusion, the activity of a lactococcal prophage-encoded Sie and repressor from the same prophage was assessed against 15 lactococcal phages belonging to

either the P335, 936 or 949 group. The data in this current study suggests that the phage-resistance phenotype of the LC3 lysogens of 3107 against phage members of the P335 groups is largely due to superinfection exclusion (mediated by $Sie_{2009/LC3}$) rather than superinfection immunity (mediated by Rep_{LC3}). The current study has provided further insights into the importance and activity of Sie systems and repressor genes as potential phage-control mechanisms in the dairy fermentation industry. Starter cultures that harbour prophages expressing *sie* and/or repressor-encoding genes may therefore be protected against superinfecting phages, though they may at the same time be under threat by such prophages in case they become induced. Using genetic engineering, starter cultures harbouring desired *sie* or repressor-encoding systems could be constructed, thus providing phage resistance and consequently preventing phage infection of dairy starter cultures.

5.5 References

1. Mahony J, Murphy J, van Sinderen D. 2012. Lactococcal 936-type phages and dairy fermentation problems: from detection to evolution and prevention. Frontiers in microbiology 3:335.
2. Garneau JE, Moineau S. 2011. Bacteriophages of lactic acid bacteria and their impact on milk fermentations. Microbial cell factories 10:S20.
3. Bebeacua C, Lai L, Vegge CS, Brondsted L, van Heel M, Veesler D, Cambillau C. 2013. Visualizing a complete *Siphoviridae* member by single-particle electron microscopy: the structure of lactococcal phage TP901-1. Journal of Virology 87:1061-8.
4. Stockdale SR, Mahony J, Courtin P, Chapot-Chartier MP, van Pijkeren JP, Britton RA, Neve H, Heller KJ, Aideh B, Vogensen FK, van Sinderen D. 2013. The lactococcal phages Tuc2009 and TP901-1 incorporate two alternate forms of their tail fiber into their virions for infection specialization. Journal of biological chemistry 288:5581-90.
5. van Hylckama Vlieg JE, Rademaker JL, Bachmann H, Molenaar D, Kelly WJ, Siezen RJ. 2006. Natural diversity and adaptive responses of *Lactococcus lactis*. Current opinion in biotechnology 17:183-190.
6. Moineau S, Borkaev M, Holler BJ, Walker SA, Kondo JK, Vedamuthu ER, Vandenbergh PA. 1996. Isolation and characterization of lactococcal bacteriophages from cultured buttermilk plants in the United States. Journal of dairy science 79:2104-2111.
7. Neve H, Kemper U, Geis A, Heller K. 1994. Monitoring and characterization of lactococcal bacteriophages in a dairy plant. Kieler milchwirtschaftliche forschungsberichte (Germany).
8. Mahony J, Martel B, Tremblay DM, Neve H, Heller KJ, Moineau S, Sinderen D. 2013. Identification of a new P335 subgroup through molecular analysis of lactococcal phages Q33 and BM13. Applied and environmental microbiology 79(14), 4401-4409.
9. Mc Grath S, Neve H, Seegers JF, Eijlander R, Vegge CS, Brondsted L, Heller KJ, Fitzgerald GF, Vogensen FK, van Sinderen D. 2006. Anatomy of a lactococcal phage tail. Journal of bacteriology 188:3972-82.
10. Mahony J, Oliveira J, Collins B, Hanemaaijer L, Lugli GA, Neve H, Ventura M, Kouwen TR, Cambillau C, van Sinderen D. 2017. Genetic and functional

characterisation of the lactococcal P335 phage-host interactions. BMC genomics 18:146.

11. Mahony J, Van Sinderen D. 2012. Structural aspects of the interaction of dairy phages with their host bacteria. Viruses 4:1410-1424.
12. Mahony J, Ainsworth S, Stockdale S, van Sinderen D. 2012. Phages of lactic acid bacteria: the role of genetics in understanding phage-host interactions and their co-evolutionary processes. Virology 434:143-150.
13. Valyasevi R, Sandine W, Geller B. 1991. A membrane protein is required for bacteriophage c2 infection of *Lactococcus lactis* subsp. *lactis* C2. Journal of bacteriology 173:6095-6100.
14. McCabe O, Spinelli S, Farenc C, Labbé M, Tremblay D, Blangy S, Oscarson S, Moineau S, Cambillau C. 2015. The targeted recognition of *Lactococcus lactis* phages to their polysaccharide receptors. Molecular microbiology 96:875-886.
15. Nyengaard N, Vogensen FK, Josephsen J. 1995. Restriction-modification systems in *Lactococcus lactis*. Gene 157:13-18.
16. Coffey A, Ross RP. 2002. Bacteriophage-resistance systems in dairy starter strains: molecular analysis to application, p 303-321, Lactic acid bacteria: genetics, metabolism and applications. Springer.
17. McGrath S, Fitzgerald GF, Sinderen Dv. 2002. Identification and characterization of phage-resistance genes in temperate lactococcal bacteriophages. Molecular microbiology 43:509-520.
18. Tükel Çl, Şanlibaba P, Özden B, Akcelik M, Akçelik M. 2006. Identification of adsorption inhibition, restriction/modification and abortive infection type phage resistance systems in *Lactococcus lactis* strains. Acta biologica hungarica 57:377-385.
19. Murray NE. 2000. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiology and molecular biology reviews 64:412-434.
20. Pingoud A, Fuxreiter M, Pingoud V, Wende W. 2005. Type II restriction endonucleases: structure and mechanism. Cellular and molecular life sciences 62:685.
21. Chopin M-C, Chopin A, Bidnenko E. 2005. Phage abortive infection in lactococci: variations on a theme. Current opinion in microbiology 8:473-479.

22. Dhaese P, Seurinck J, De Smet B, Van Montagu M. 1985. Nucleotide sequence and mutational analysis of an immunity repressor gene from *Bacillus subtilis* temperate phage ø105. Nucleic acids research 13:5441-5455.
23. Ladero V, García P, Bascarán V, Herrero M, Alvarez MA, Suárez JE. 1998. Identification of the repressor-encoding gene of the *Lactobacillus* bacteriophage A2. Journal of bacteriology 180:3474-3476.
24. Nesper J, Blaß J, Fountoulakis M, Reidl J. 1999. Characterization of the major control region of vibrio cholerae bacteriophage K139: Immunity, exclusion, and integration. Journal of bacteriology 181:2902-2913.
25. Uc-Mass A, Loeza EJ, de la Garza M, Guarneros G, Hernández-Sánchez J, Kameyama L. 2004. An orthologue of the *cor* gene is involved in the exclusion of temperate lambdoid phages. Evidence that Cor inactivates FhuA receptor functions. Virology 329:425-433.
26. Vostrov AA, Vostrukhina OA, Svarchevsky AN, Rybchin VN. 1996. Proteins responsible for lysogenic conversion caused by coliphages N15 and phi80 are highly homologous. Journal of bacteriology 178:1484-1486.
27. Snyder L. 1995. Phage-exclusion enzymes: a bonanza of biochemical and cell biology reagents? Molecular microbiology 15:415-420.
28. Newton GJ, Daniels C, Burrows LL, Kropinski AM, Clarke AJ, Lam JS. 2001. Three-component-mediated serotype conversion in *Pseudomonas aeruginosa* by bacteriophage D3. Molecular microbiology 39:1237-1247.
29. Bondy-Denomy J, Davidson AR. 2014. When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness. Journal of microbiology 52:235-242.
30. Johnson AD, Poteete AR, Lauer G, Sauer RT, Ackers GK, Ptashne M. 1981. lambda repressor and cro-components of an efficient molecular switch. Nature 294:217-23.
31. Sauer R, Ross M, Ptashne M. 1982. Cleavage of the lambda and P22 repressors by RecA protein. Journal of biological chemistry 257:4458-4462.
32. Durmaz E, Madsen SM, Israelsen H, Klaenhammer TR. 2002. *Lactococcus lactis* lytic bacteriophages of the P335 group are inhibited by overexpression of a truncated CI repressor. Journal of bacteriology 184:6532-6544.
33. Kuzio J, Kropinski AM. 1983. O-antigen conversion in *Pseudomonas aeruginosa* PAO1 by bacteriophage D3. Journal of bacteriology 155:203-212.

34. Cumby N, Edwards AM, Davidson AR, Maxwell KL. 2012. The bacteriophage HK97 gp15 moron element encodes a novel superinfection exclusion protein. Journal of bacteriology 194:5012-5019.
35. Mahony J, McGrath S, Fitzgerald GF, van Sinderen D. 2008. Identification and characterization of lactococcal-prophage-carried superinfection exclusion genes. Applied and environmental microbiology 74:6206-6215.
36. Sun X, Göhler A, Heller KJ, Neve H. 2006. The *ltp* gene of temperate Streptococcus thermophilus phage TP-J34 confers superinfection exclusion to *Streptococcus thermophilus* and *Lactococcus lactis*. Virology 350:146-157.
37. Lillehaug D. 1997. An improved plaque assay for poor plaque-producing temperate lactococcal bacteriophages. Journal of applied microbiology 83:85-90.
38. Braun Jr V, Hertwig S, Neve H, Geis A, Teuber M. 1989. Taxonomic differentiation of bacteriophages of *Lactococcus lactis* by electron microscopy, DNA-DNA hybridization, and protein profiles. Microbiology 135:2551-2560.
39. Kuipers OP, de Ruyter PG, Kleerebezem M, de Vos WM. 1998. Quorum sensing-controlled gene expression in lactic acid bacteria. Journal of biotechnology 64:15-21.
40. Sambrook J, Russell DW. 2006. Transformation of *E. coli* by electroporation. Cold spring harbor protocols 2006:pdb. prot3933.
41. Christiansen B, Johnsen M, Stenby E, Vogensen F, Hammer K. 1994. Characterization of the lactococcal temperate phage TP901-1 and its site-specific integration. Journal of bacteriology 176:1069-1076.
42. Lillehaug D, Lindqvist B, Birkeland N. 1991. Characterization of phiLC3, a *Lactococcus lactis* subsp. *cremoris* temperature bacteriophage with cohesive single-stranded DNA ends. Applied and environmental microbiology 57:3206-3211.
43. Oliveira J, Mahony J, Lugli GA, Hanemaaijer L, Kouwen T, Ventura M, van Sinderen D. 2016. Genome sequences of eight prophages isolated from *Lactococcus lactis* dairy strains. Genome announcements 4:e00906-16.
44. Oliveira J. 2017. Biodiversity of lactococcal bacteriophages and associated phage-host interactions: PhD thesis, University College Cork.

45. Samson JE, Moineau S. 2010. Characterization of *Lactococcus lactis* phage 949 and comparison with other lactococcal phages. Applied and environmental microbiology 76:6843-6852.
46. Mahony J, Randazzo W, Neve H, Settanni L, Sinderen D. 2015. Lactococcal 949 group phages recognize a carbohydrate receptor on the host cell surface. Applied and environmental microbiology 81:3299-305.
47. McGrath S, Fitzgerald GF, van Sinderen D. 2001. Improvement and optimization of two engineered phage resistance mechanisms in *Lactococcus lactis*. Applied and environmental microbiology 67:608-616.
48. Breum SØ, Neve H, Heller KJ, Vogensen FK. 2007. Temperate phages TP901-1 and φLC3, belonging to the P335 species, apparently use different pathways for DNA injection in *Lactococcus lactis* subsp. *cremoris* 3107. FEMS microbiology letters 276:156-164.
49. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold spring harbor laboratory press.
50. Sambrook J, Russell DW. 2006. Purification of nucleic acids by extraction with phenol: chloroform. Cold spring harbor protocols 2006:pdb. prot4455.
51. Lugli GA, Milani C, Mancabelli L, van Sinderen D, Ventura M. 2016. MEGAnnotator: a user-friendly pipeline for microbial genomes assembly and annotation. FEMS microbiology letters 363.
52. Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Müller WE, Wetter T, Suhai S. 2004. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome research 14:1147-1159.
53. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC bioinformatics 11:119.
54. Garzon AE, Mahony J, Bottacini F, Kelleher P, van Sinderen D. 2019. Complete genome sequence of *Lactococcus lactis* subsp. *cremoris* 3107, host for the model lactococcal P335 bacteriophage TP901-1. Microbiology Resource Announcement 8:e01635-18.
55. Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome research 14:1394-1403.

56. Lillehaug D, Birkeland N. 1993. Characterization of genetic elements required for site-specific integration of the temperate lactococcal bacteriophage phi LC3 and construction of integration-negative phi LC3 mutants. Journal of bacteriology 175:1745-1755.
57. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. Bioinformatics 16:944-5.
58. Swindell SR, Plasterer TN. 1997. SeqMan: Contig assembly. Sequence data analysis guidebook:75-89.
59. Berngruber TW, Weissing FJ, Gandon S. 2010. Inhibition of superinfection and the evolution of viral latency. Journal of virology 84:10200-10208.
60. Kelleher P, Mahony J, Schweinlin K, Neve H, Franz CM, van Sinderen D. 2018. Assessing the functionality and genetic diversity of lactococcal prophages. International journal of food microbiology 272:29-40.

Thesis Discussion

Dairy fermentation facilities are under constant threat by phage infections. It is now widely accepted that phage proliferation in fermentation processes will continue to be a problem as phages cannot be completely eliminated, though their multiplication and dispersal may be controlled. In order to better comprehend how a lactococcal phage recognizes and infects its host, it is imperative that functional genomic studies and structural characterisation of both the phage and host components involved in these interactions is undertaken (1-9).

The overall aim of the work described in this thesis was to identify bacterial and phage-encoded protein(s) required for, or interfering with, the infection process of the dairy starter strain *Lactococcus lactis* subsp. *cremoris* (*L. lactis*) 3107 by lactococcal phages TP901-1 and LC3, both of which belong to the so-called P335 group of lactococcal phages. Previously, the injection processes of TP901-1 and LC3 have been suggested to be distinct (10). The interaction of phage TP901-1 with its host, *L. lactis* 3107, commences with the adsorption of the phage to the polysaccharide pellicle (PSP) component of the cell wall polysaccharide (CWPS) of its host (2). The ability to genetically manipulate a strain as well as the availability of its sequenced genome are key elements to facilitate studies of the interactions between a phage and its host at a molecular level. *L. lactis* 3107 displays a low efficiency of transformation using currently available methods (11, 12). Therefore, the lack of refined protocols for the efficient transfer of DNA into this dairy-associated strain has limited its genetic manipulation. This thesis describes an optimised transformation protocol for *L. lactis* 3107, thereby improving its genetic manipulation, while we furthermore determined its genome sequence to facilitate further exploration of the molecular interactions between phages TP901-1 and LC3 with their host *L. lactis* 3107.

Following attachment of phage TP901-1 to the host-specified saccharidic receptor, DNA injection of this phage is triggered by the product of *gtfA* or, more likely based on our current data, its glycosylated substrate. Currently, the identity of the substrate for the predicted membrane-associated glycosyltransferase GtfA is not known, nor is it clear how GtfA is involved in triggering DNA injection of TP901-1. Previous phage-host interaction studies involving Gram-positive infecting phages have demonstrated that phages may recognize various host-encoded receptor

molecule(s), such as a CWPS, teichoic acid (TA) and/or a surface-exposed protein (13, 14). GtfA may therefore act as an injection trigger by itself or indirectly through the saccharidic product of its predicted transferase activity. Based on the predicted function of GtfA, this thesis proposes that the DNA injection trigger of TP901-1 is the (as yet unknown) cell envelope-associated substrate for GtfA. Following confirmation that neither the PSP nor rhamnan is the substrate for GtfA, possible GtfA substrates include LTA, WTA, peptidoglycan or a glycoprotein. The analysis of TAs as possible substrates for GtfA performed as part of this thesis was compromised by the lack of knowledge regarding the glycosylation process (if any) of LTAs of *L. lactis* 3107 and our ignorance regarding (the existence of) WTA in *L. lactis* strains. Analysis of the composition of lactococcal cell wall components is therefore necessary for the identification of components acting as DNA injection triggers required by their infecting phages. Analysis of the biosynthetic pathway of each component of the *L. lactis* 3107 cell envelope can determine if GtfA is involved in the glycosylation process of one of these components. If the substrate for GtfA does not represent the DNA injection trigger for TP901-1, it may be that (presumed) membrane-associated GtfA acts directly as a proteinaceous injection trigger. For example, following initial attachment, the model *Bacillus subtilis* siphophage SPP1 irreversibly binds to the ectodomain of the membrane-associated YueB protein, thereby triggering DNA injection (13). Similar to YueB, GtfA exhibits an ectodomain which may represent the actual injection trigger for TP901-1. In order to test this notion, expression and isolation of the ectodomain of GtfA will need to be pursued, followed by an assessment of the ability of such a protein to falsely trigger DNA injection of TP901-1 phages.

Three TP901-1 resistant derivatives of *L. lactis* 3107, known as E119, E121 and E126, were isolated of which two (E119 and E126) were complemented by *gtfA*, demonstrating the involvement of GtfA in DNA injection of TP901-1. GtfA is hypothesized to be part of an extracellular glycosylation pathway (as shown in Fig. 1, which was previously discussed in Chapter IV) which typically involves (i) an enzyme required to produce an UDP-sugar intermediate (possibly the gene upstream of *gtfA*, which possesses such a predicted function), (ii) an enzyme that transfers the sugar from its UDP-linked form to membrane-associated undecaprenyl, (iii) a flippase enzyme that re-orients this undecaprenyl-sugar from a cytoplasmic- to an extracellular-facing membrane location, and (iv) an enzyme, in this case GtfA, which transfers the sugar from the undecaprenyl-carrier to its presumed cell wall-associated

substrate. If the TP901-1 DNA injection trigger is the GtfA substrate, it follows that mutations in the genes for the enzymes that are responsible for steps (i) through to (iv) will cause insensitivity to TP901-1. Since two of the mutants (i.e. E191 and E126) were defective in step (iv) we hypothesize that the third mutant (E121) carries a mutation in a gene responsible for step (i), (ii) or (iii). Candidates for such genes (for steps i-iii) have been identified and future work will focus on determining their involvement in TP901-1 DNA injection (for example by cloning/expressing such candidates into the E121 mutant). The generation of TP901-1 escape mutants, being capable of infecting all three phage-resistant 3107 derivatives of which two harbour a defective *gtfA* gene and one harbours a functional *gtfA* gene, further suggests that a component of the biosynthetic pathway involved in the production of the DNA injection trigger for TP901-1 other than GtfA is mutated in E121. Future studies could focus on optimising the recombineering system for application in *L. lactis* 3107 to allow targeted mutagenesis in this strain (15-17). Improving the genetic manipulation of *L. lactis* 3107 is an important element in identifying genes, in addition to *gtfA*, as a genetic component of *L. lactis* 3107 responsible for triggering DNA injection of phage TP901-1.

This thesis also showed that subsequent adaptation of phage TP901-1 to phage-resistant host cells is achieved through acquisition of particular mutations in the *tal* gene. Currently it is not clear how these *tal* mutations in TP901-1 allow the corresponding mutant phages to overcome the barriers presented by GtfA-negative *L. lactis* 3107 derivatives. Perhaps the modified Tal is capable of recognizing or attaching to a modified injection trigger, thereby opening the DNA ejection conduit through conformational changes in the tail tube allowing DNA release. Future studies could focus on introducing targeted mutations to the sequence of *tal* in order to determine the exact nucleotide alteration(s) that is (are) responsible for the TP901-1 DNA injection trigger. Generation of TP901-1 escape mutants reinforces the fact that lactococcal phages can adapt to phage-resistant dairy host strains.

This thesis revealed that *gtfA* is ubiquitous among *Lactococcus lactis* genomes and that it is specifically involved in TP901-1 DNA injection as well as that of other lactococcal phages such as certain members of the rare 949 phages, i.e. phages 949 and WRP3. This study was limited by the number of available lactoccocal P335 group phages capable of infecting *L. lactis* NZ9000, thus making it impossible to assess the impact of a non-functional *gtfA* in this genetically accessible strain. It has previously

been demonstrated that following reversible adsorption, a number of c2-like lactococcal phages irreversibly bind to PIP (phage infection protein), triggering ejection of the phage genome (18). Further screening studies of lactococcal phages capable of infecting *L. lactis* NZ9000 are necessary to facilitate studying the role of *gtfA* in phage DNA injection.

Significant research efforts have been undertaken to identify and characterize natural phage defense mechanisms acquired by lactococcal strains in order to identify potential phage-control mechanisms in the dairy sector (19-25). This thesis also aimed to generate LC3-resistant derivatives of *L. lactis* 3107 and to identify gene(s) involved in the acquired phage resistance. Following the exposure of LC3 to *L. lactis* 3107 we obtained lysogens of LC3 which exhibit resistance to a number of P335 phages. The LC3-prophage-encoded superinfection exclusion (Sie) protein $Sie_{2009/LC3}$ and repressor protein Rep_{LC3} were shown to be responsible for such LC3 lysogens to confer resistance to various P335 phages. In addition, the plasmid-driven expression of $Sie_{2009/LC3}$ did not only result in complete resistance against phages belonging to the P335 group, but also against phage members of the lactococcal 936 or 949 groups. This thesis further demonstrates the importance of lactococcal Sie and/or repressor proteins as mechanisms that could be introduced in the dairy fermentation industry to help control phage contamination. It would be interesting to determine if escape mutants can be generated which are capable of overcoming Sie or superinfection immunity in order to assess the robustness of such phage resistance systems. The TP-J34 prophage in *Streptococcus thermophilus* has previously been shown to encode the lipoprotein Ltp_{TP-J34}, which prevents (secondary) infection by 936 group phages. This streptococcal Sie system is predicted to directly interact with and block entry of the tape measure protein (TMP) into the bacterial membrane from the phage tail tube, thus inhibiting DNA injection (26). Escape mutant phages overcoming Ltp_{TP-J34} exhibit mutations within the *tmp* gene which resulted in the assembly of shorter tails (27). The TMP is important for tail assembly, ion-channel formation, and DNA passage into the cell, thus, like TP-J34, TP901-1 may need to mutate its TMP to overcome the $Sie_{2009/LC3}$ system. Starter cultures could be genetically engineered to construct strains harbouring desired *sie* or repressor-encoding systems, thus preventing phage contamination and consequently avoiding the negative implications of phage infection of dairy starter cultures. However, such genetically engineered strains will as yet not

find their way to the food fermentation industry due to restrictive regulations with regards to the use of genetically modified organisms (GMOs) in food products.

In conclusion, this thesis analysed how the lactococcal P335 phage TP901-1 interacts with its *L. lactis* 3107 host at a molecular level, and has introduced new possibilities for future studies focused on such phage-host interactions. The importance of identifying host-encoded genes involved in phage-host interactions is abundantly clear as genetic manipulation of such genes can help reduce phage contamination in the dairy fermented sector. In addition, the importance of prophage-associated *sie* and *rep* genes as phage-defence mechanisms for phage contamination has been highlighted.

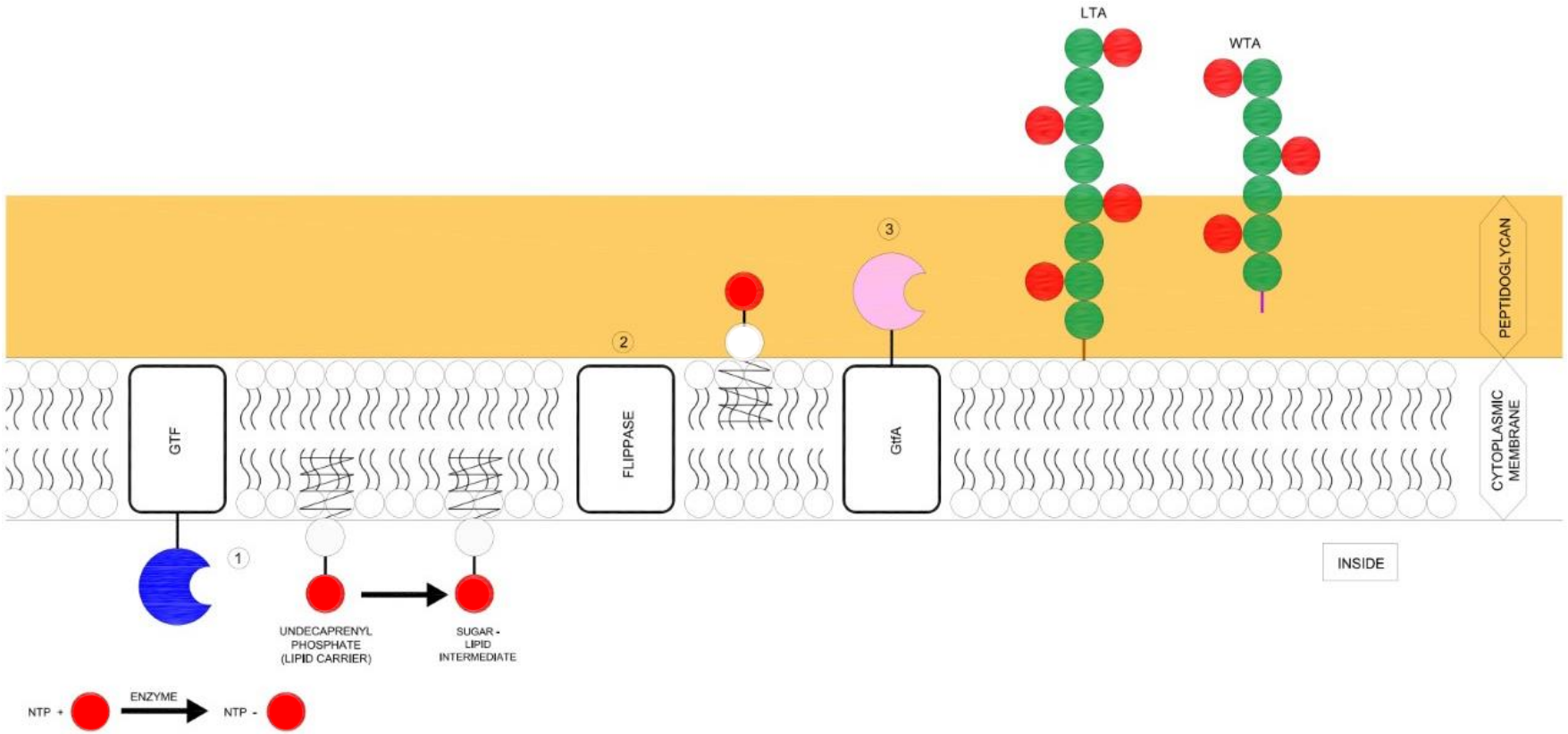


Figure 1: Model for TA (LTA or WTA) glycosylation in *L. lactis*. The LTA or WTA glycosylation process in *L. lactis* initially requires an, as yet, unidentified cytoplasmic GTF (blue) which transfers a sugar residue (red) onto an undecaprenyl-phosphate which acts as a lipid carrier (Step 1). The sugar-lipid intermediate is then transported across the membrane by an unknown flippase enzyme (Step 2). The sugar residue is then transferred by a second extracellular GTF proposed to be GtfA (pink) onto the GroP chain (green) of the LTA which is anchored in the cytoplasmic membrane by means of a lipid attachment (brown) or WTA which is covalently attached to peptidoglycan (purple) (Step 3).

References

1. Mahony J, Ainsworth S, Stockdale S, van Sinderen D. 2012. Phages of lactic acid bacteria: the role of genetics in understanding phage-host interactions and their co-evolutionary processes. Virology 434:143-150.
2. Ainsworth S, Sadovskaya I, Vinogradov E, Courtin P, Guerardel Y, Mahony J, Grard T, Cambillau C, Chapot-Chartier MP, Sinderen D. 2014. Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. MBio 5:e00880-14.
3. Farenc C, Spinelli S, Vinogradov E, Tremblay D, Blangy S, Sadovskaya I, Moineau S, Cambillau C. 2014. Molecular insights on the recognition of a *Lactococcus lactis* cell wall pellicle by the phage 1358 receptor binding protein. Journal of virology 88:7005-15.
4. Legrand P, Collins B, Blangy S, Murphy J, Spinelli S, Gutierrez C, Richet N, Kellenberger C, Desmyter A, Mahony J. 2016. The atomic structure of the phage Tuc2009 baseplate tripod suggests that host recognition involves two different carbohydrate binding modules. MBio 7:e01781-15.
5. Mahony J, Randazzo W, Neve H, Settanni L, Sinderen D. 2015. Lactococcal 949 group phages recognize a carbohydrate receptor on the host cell surface. Applied and environmental microbiology Microbiol 81:3299-305.
6. Andre G, Kulakauskas S, Chapot-Chartier M-P, Navet B, Deghorain M, Bernard E, Hols P, Dufrêne YF. 2010. Imaging the nanoscale organization of peptidoglycan in living *Lactococcus lactis* cells. Nature communications 1:27.
7. Chapot-Chartier MP, Vinogradov E, Sadovskaya I, Andre G, Mistou MY, Trieu-Cuot P, Furlan S, Bidnenko E, Courtin P, Pechoux C. 2010. Cell surface of *Lactococcus lactis* is covered by a protective polysaccharide pellicle. Journal of biological chemistry 285:10464-71.
8. Vinogradov E, Sadovskaya I, Courtin P, Kulakauskas S, Grard T, Mahony J, van Sinderen D, Chapot-Chartier M-P. 2018. Determination of the cell wall polysaccharide and teichoic acid structures from *Lactococcus lactis* IL1403. Carbohydrate research 462:39-44.
9. Vinogradov E, Sadovskaya I, Grard T, Murphy J, Mahony J, Chapot-Chartier M-P, van Sinderen D. 2018. Structural studies of the cell wall polysaccharide from *Lactococcus lactis* UC509.9. Carbohydrate research 461:25-31.

10. Ostergaard Breum S, Neve H, Heller KJ, Vogensen FK. 2007. Temperate phages TP901-1 and phiLC3, belonging to the P335 species, apparently use different pathways for DNA injection in *Lactococcus lactis* subsp. *cremoris* 3107. FEMS microbiology letters 276:156-64.
11. Holo H, Nes IF. 1989. High-frequency transformation, by electroporation, of *Lactococcus lactis* subsp. *cremoris* grown with glycine in osmotically stabilized media. Applied and environmental microbiology 55:3119-3123.
12. Landete JM, Arqués JL, Peirotén Á, Langa S, Medina M. 2014. An improved method for the electrotransformation of lactic acid bacteria: a comparative survey. Journal of microbiological methods 105:130-133.
13. São-José C, Baptista C, Santos MA. 2004. *Bacillus subtilis* operon encoding a membrane receptor for bacteriophage SPP1. Journal of bacteriology 186:8337-8346.
14. Valyasevi R, Sandine W, Geller B. 1991. A membrane protein is required for bacteriophage c2 infection of *Lactococcus lactis* subsp. *lactis* C2. Journal of bacteriology 173:6095-6100.
15. Stockdale SR, Mahony J, Courtin P, Chapot-Chartier MP, van Pijkeren JP, Britton RA, Neve H, Heller KJ, Aideh B, Vogensen FK, van Sinderen D. 2013. The lactococcal phages Tuc2009 and TP901-1 incorporate two alternate forms of their tail fiber into their virions for infection specialization. Journal of biological chemistry 288:5581-90.
16. van Pijkeren J-P, Neoh KM, Sirias D, Findley AS, Britton RA. 2012. Exploring optimization parameters to increase ssDNA recombineering in *Lactococcus lactis* and *Lactobacillus reuteri*. Bioengineered 3:209-217.
17. van Pijkeren J-P, Britton RA. 2012. High efficiency recombineering in lactic acid bacteria. Nucleic acids research 40:e76-e76.
18. Monteville MR, Ardestani B, Geller BL. 1994. Lactococcal bacteriophages require a host cell wall carbohydrate and a plasma membrane protein for adsorption and ejection of DNA. Applied and environmental microbiology 60:3204-3211.
19. Nyengaard N, Vogensen FK, Josephsen J. 1995. Restriction-modification systems in *Lactococcus lactis*. Gene 157:13-18.

20. Coffey A, Ross RP. 2002. Bacteriophage-resistance systems in dairy starter strains: molecular analysis to application, p 303-321, Lactic acid bacteria: genetics, metabolism and applications. Springer.
21. McGrath S, Fitzgerald GF, Sinderen Dv. 2002. Identification and characterization of phage-resistance genes in temperate lactococcal bacteriophages. Molecular microbiology 43:509-520.
22. Tükel Çl, Şanlibaba P, Özden B, Akcelik M, Akçelik M. 2006. Identification of adsorption inhibition, restriction/modification and abortive infection type phage resistance systems in *Lactococcus lactis* strains. Acta biologica hungarica 57:377-385.
23. Murray NE. 2000. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiology and molecular biology reviews 64:412-434.
24. Pingoud A, Fuxreiter M, Pingoud V, Wende W. 2005. Type II restriction endonucleases: structure and mechanism. Cellular and molecular life sciences 62:685.
25. Chopin M-C, Chopin A, Bidnenko E. 2005. Phage abortive infection in lactococci: variations on a theme. Current opinion in microbiology 8:473-479.
26. Sun X, Göhler A, Heller KJ, Neve H. 2006. The *ltp* gene of temperate *Streptococcus thermophilus* phage TP-J34 confers superinfection exclusion to *Streptococcus thermophilus* and *Lactococcus lactis*. Virology 350:146-157.
27. Bebeacua C, Lorenzo Fajardo JC, Blangy S, Spinelli S, Bollmann S, Neve H, Cambillau C, Heller KJ. 2013. X-ray structure of a superinfection exclusion lipoprotein from phage TP-J 34 and identification of the tape measure protein as its target. Molecular microbiology 89:152-165.

Abbreviations

Abi - Abortive infection

Adi - Adsorption inhibition

Ala - Alanine

BIM - Bacteriophage insensitive mutant

BLAST - Basic local alignment search tool

BP - Baseplate

CFSN - Cell free supernatant

CFU – Colony forming unit

Cm - Chloramphenicol

COG - Clusters of orthologous groups

CWPS - Cell wall polysaccharides

DHB - 2,5-dihydroxy-benzoic acid

ds-DNA - Double-stranded DNA

E.O.P. - Efficiency of plaquing

Em - Erythromycin

EPS - Extracellular polysaccharides

Gal - Galactose

Gal-t - Galactose-1-phosphate-uridylyltransferase

GC - Gas chromatography

GC-MS - Gas chromatography-mass spectrometry

GlcNac - N-acetyl-glucosamine

GMO - Genetically modified organism

GTF - Glycosyltransferase

HF - Hydrofluoric acid

HPAEC-PAD - High performance anion exchange chromatography coupled with pulse-amperometric detection

HPLC - High-performance liquid chromatography

ICTV - International committee on taxonomy of viruses

IS - Insertion sequence

LAB - Lactic acid bacteria

LPS - Lipopolysaccharides

LTA - Lipoteichoic acid

MALDI-TOF MS - Matrix-assisted laser desorption ionization-time of flight mass spectrometry

MCP - Minor capsid protein

MTase - Methyltransferase

MOI - Multiplicity of infection

MTP - Major tail protein

MurNac - N-acetyl-muramic acid

N/A - Not applicable

NCBI - National center for biotechnology information

ND - Not determined

NGS - Next generation sequencing

Nt - N-terminal

ORF - Open reading frame

PIP - Phage infection protein

PMF - Proton motive force

PS - Polysaccharide

PSP - Polysaccharide pellicle

RBP - Receptor-binding protein

RCR - Rolling circle replication

REase - Restriction endonuclease

R-M - Restriction-modification

rRNA - Ribosomal RNA

SEC - Size exclusion chromatography

Sie - Superinfection exclusion

Sii - Superinfection immunity

SMRT - Single molecule real time

SNP - Single nucleotide polymorphism

SRA - Sequence Read Archive

TA - Teichoic acid

Tal - Tail-associated lysin

TCA - Trichloroacetic acid

TFA - Trifluoroacetic acid

TMP - Tape measure protein

tRNA - Transfer RNA

UDP - 5'-diphosphate

UMP - Uridine 5'-monophosphate

Usp - Universal stress protein

WT - Wild type

WTA - Wall teichoic acids

Acknowledgements

Firstly I would like to thank Douwe for giving me the opportunity to undertake this PhD and be part of this research project. Douwe, thank you for being a magnificent supervisor, your support, guidance and expertise throughout my PhD has been priceless. I greatly appreciate your encouragement to always think outside the box and aim for perfection. Thank you so much for your time and patience with me, it has been a pleasure to work under your supervision.

Jen, you have been a wonderful mentor and I'm so grateful to have been able to learn so much from you. I really appreciate that you always made time for me whenever I needed guidance, and that you always made me feel like I could come to you not only professionally but also as a friend. I will always be so grateful to you for listening to me and giving me advice when I became so overwhelmed with the workload, your advice got me through my 3rd year. I would also like to take this opportunity to show my immense gratitude to you both, Douwe and Jen. Dealing with my grandmother's illness and passing away was very challenging and your support and advice meant a lot to me, so for that, I will always be grateful.

My PhD thesis would not have been completed without the input of all the international collaborators, Dr. Horst Neve, Prof. Finn Vogensen, Dr. Christian Cambillau, Dr. Marie-Pierre Chapot-Chartier, Dr. Simon Palussière and Dr. Pascal Courtin, who I would like to give a big thank you to. I would also like to say thanks to James, Philip, Eoghan, Juan and Joana for being there for me whenever I needed guidance, and to Stephen, Katherine, Tom, Mona, Marine and Ilias for sharing this

experience with me. I would like to give an enormous thank you to James for guiding me through the beginning of my PhD and for being a great mentor to me. I would also like to give a huge thank you to Philip for teaching me everything that I know about Bioinformatics and for always being more than happy to help me, it was a pleasure sharing leg space with you! Stephen, Katherine and Joanna thank you for taking the time to get to know me and for always being there for me. My girl next door, Jenna, it was a pleasure sharing my under- and post-graduate experience with you, can't wait to see where life takes us next! To all the people in lab 340, 339 and 338 who made this experience very enjoyable and always had me laughing, thank you so much and I wish you all the best.

To my very first mentors, Selena and David, working under your supervision as a Research Assistant was an experience that I will always cherish. You were fantastic mentors, thanks to you both I became passionate about research. You inspired me to a PhD and for that I will always be thankful.

To my closest friends, you have no idea what a stress relief your company has been during this PhD. Alana, Laura, Megan, Alison, Kevin and Chrissy whenever I'm with you suddenly I feel 16 again. Thank you for your friendship all these years and for always reminding me that life is nothing without laughing, singing and dancing! Aoife, thank you for always thinking of me and for all those times you helped me unwind by taking me out, showing me how beautiful Cork is with all those nature walks, and for always sharing your work-out plans with me! Jenna, it

was so nice having a familiar face around, thank you for always being so willing to escape reality with me!

To my little sister, Aisha, thank you for always looking up to me and thinking that I'm the best. To my Godmother, Nicola, thank you so much for teaching me English and for helping me with my homework in secondary school, I would've never made it this far without all your help. Thank you for always being there to listen to me and give me advice whenever I need it. To my dad, thank you for being there when I need you and for always being so proud of me.

To my sweetheart Colin, throughout this PhD you have been my rock and there is no words to describe how thankful I am. The older we get the more I realise what an amazing man you've turned into and how lucky I am to have you. Thank you for always being there for me and for making my PhD your priority, your support has been everything!

And finally, to my amazing Mami who is a strong and gentle soul. Thank you for encouraging me to take on this PhD, without your encouragement and continued support I would not have made it this far. I am so lucky to have you as a mother, you made the sacrifice to move to Ireland so I can have more opportunities in life. You have always given us the best of everything and put our studies first and for that I will always be grateful and proud to call you my Mami. Like your dad taught you, you've taught me to work like I'm going to live forever and enjoy life like there is no tomorrow, and that's what we will continue doing. Yeah! I'm finally finished.... ☺