Review

Cell wall polysaccharides of Gram positive ovococcoid bacteria and their role as bacteriophage receptors

Katherine Lavelle, Douwe van Sinderen *, Jennifer Mahony *

School of Microbiology & APC Microbiome Ireland, University College Cork, Western Road, Cork T12 YI20, Ireland

Abstract

Gram-positive bacterial cell walls are characterised by the presence of a thick peptidoglycan layer which provides protection from extracellular stresses, maintains cell integrity and determines cell morphology, while it also serves as a foundation to anchor a number of crucial polymeric structures. For ovococcal species, including streptococci, enterococci and lactococci, such structures are represented by rhamnose-containing cell wall polysaccharides, which at least in some instances appear to serve as a functional replacement for wall teichoic acids. The biochemical composition of several streptococcal, lactococcal and enterococcal rhamnose-containing cell wall polysaccharides have been elucidated, while associated functional genomic analyses have facilitated the proposition of models for individual biosynthetic pathways. Here, we review the genomic loci which encode the enzymatic machinery to produce rhamnose-containing, cell wall-associated polysaccharide (Rha cwps) structures of the afore-mentioned ovococcal bacteria with particular emphasis on gene content, biochemical structure and common biosynthetic steps. Furthermore, we discuss the role played by these saccharidic polymers as receptors for bacteriophages and the important role phages play in driving Rha cwps diversification and evolution.

© 2021 Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The Gram-positive bacterial cell wall contains a dense, thick layer of peptidoglycan, a mesh-like structure composed of alternat-
of peptidoglycan outgrowth, which surrounds the cell following peripheral peptidoglycan synthesis (Fig. 1), a mechanism which is not observed in true coccoid species [1,2,3]. Peptidoglycan acts as a scaffold for various secondary cell wall glyco-polymers which decorate the exposed surface of Gram-positive cells. Comprising approximately 50–60 % of the dry weight of the cell wall [4,5], glyco-polymers such as wall teichoic acid, certain capsular polysaccharides (including those of Streptococcus pneumoniae) which are synthesised via the Wzx/Wzy pathway, and not those utilising a synthase-dependent manner [6], and rhamnose (Rha)-containing polysaccharides have been shown to be covalently bound to the peptidoglycan layer through the activity of LytR-CspA-Pst (LCP) transferases [7,8]. In ovococcal species the Rha-containing cell wall polysaccharide (Rha cwps) mediates virulence, host adhesion, immune evasion, phage adsorption and/or antimicrobial resistance, and appears, at least in some cases, to represent a functional replacement for wall teichoic acid (WTA) [9]. Despite the wealth of information pertaining to the physiological function of rhamnose-containing cell wall polymers, the genomic loci which encode the enzymatic machinery for their biosynthesis have not been investigated with the same vigour. A comprehensive review of the Rha-containing polysaccharides of Gram-positive bacteria by Mistou and colleagues highlighted the general aspects of Rha cwps biosynthesis and their physiological functions [9]. However, significant insights regarding the relationship between genomic loci, structural composition and architecture, and biosynthesis and physiological function(s) of the ovococcal Rha cwps have since been established [10–12]. In the current review we will focus on the genome to structure relationship of the Rha cwps of key ovococcal species – streptococci, enterococci and lactococci. We highlight the utility of gene annotation and locus architecture in identifying common elements of the biosynthetic machinery across these clinically and economically important ovococcal species with implications for human and animal disease (streptococci and enterococci) and global dairy fermentation processes (lactococci and dairy streptococci).

1.1. Rhamnose-containing cell wall polysaccharides of ovococcal Gram-positive bacteria

(i) Streptococci

Species of the genus Streptococcus are extensively studied with respect to their cwps structure (and antigenicity) owing to their pathogenesis towards both humans and animals [13,14]. Streptococcal species may be divided into one of two primary groups, the β-haemolytic pyogenic group or the α-haemolytic viridans group [15]. Seminal works by Lancefield [16,17] facilitated the classification of β-haemolytic streptococci based on the presence of antigenic, cell surface-exposed carbohydrate groups (designated groups A through V) [18]. Early studies on the streptococcal antigen identified rhamnose as a commonly occurring monosaccharide with glucose (Glc), galactose (Gal), glucosamine (GlcN) also reported in many cases [19–23]. Furthermore, a common core polysaccharide structure, composed of α1,2- and α1,3-linked rhamnose units was detected in various streptococcal species [24]. As structural and compositional knowledge accumulated, the desire to identify the biosynthetic machinery for this apparently ubiquitous streptococcal polymer increased. A schematic overview of the biochemical structures of the Group specific antigen from selected Lancefield streptococci, including GAC, is presented in Fig. 2.

The non-Lancefield type species Streptococcus mutans emerged as the prototype for such investigations with four serotypes (c, e, f, or k) identified based on the site specific linkage of a glucose side chain on the α1,2-/α1,3-linked rhamnolysyl core of the S. mutans Rha cwps [25–27]. Genes associated with Rha cwps synthesis in streptococci were first identified by Tsukio et al. in S. mutans Rha cwps through the functional characterisation of a three gene locus, the products of which are involved in the biosynthesis of the nucleotide precursor sugar dTDP-L-rhamnose [28]. The identified locus, comprised of rmlA, C and B, is directly involved in the formation of the S. mutans serotype c antigen, is ubiquitous among streptococcal species and is indispensable for completion of the dTDP-L-rhamnose biosynthetic pathway, and thus by extension, the serospecific surface antigen [28,31,32]. Furthermore, a homolog of rmlD, a fourth gene involved in dTDP-L-rhamnose production, was identified in a separate locus, designated rrg, which was also shown to be required for Rha cwps biosynthesis. The S. mutans rhamnose-glucose polysaccharide (rgp) biosynthesis locus is comprised of six core genes, rgpA through to rgpF, which are essential for Rha cwps synthesis [33]. Disruption of particular genes in the rgp locus was shown to cause a reduction in both rhamnose and glucose content of cwps fractions of corresponding mutant strains. Additionally, mutations in rgpE abolished glucose side chain addition to the α1,2/α1,3-linked rhamnolysyl core [33]. A model for the assembly pathway was proposed for the polyrhamnose structure [24] in which RrgA acts as the initiating rhamnolysyltransferase, adding a single Rha residue to N-acetylglucosamine-PP-undecaprenyl (GlcNAc-PP-Und) [9,29]. Subsequently, RgpB and RgpF are proposed to extend and polymerise the rhamnose chain on the glycan through to dTDP-L-rhamnose [30,31]. The identified locus, rpgA-F, gene cluster was later found to be conserved between strains regardless of serotype, which is reflective of the conserved α1,2/α1,3-linked rhamnolysyl rhamnose produced by these strains [34]. Beyond the conserved genes of this rgg cluster, a single, unconnected gene was found to initiate synthesis of the rhamnolysyl core [35]. Named rggG, the gene was found to complement the UDP-GlcNAc:Und-P-GlcNAc-1-P-transferase activity of WeeA, an integral membrane protein which initiates biosynthesis of the
enterobacterial common antigen and O-antigen lipopolysaccharide in Escherichia coli [35,36]. The latter finding suggests that the *rgpG* gene product transfers a GlcNac moiety to the undecaprenyl lipid carrier as the initiating step in rhamnan synthesis. BlastP analysis revealed 62% amino acid similarity between RpgG and TagO, and deletion mutants of RpgG were found to not only affect Rha cwps synthesis, but to induce morphological defects, altered cell division and changes in biofilm formation [37,38].

The glucose substitution of the *S. mutans* Rha cwps core structure is achieved by means of a number of distinct enzymatic steps. GluA, a glucose-1-phosphate uridylyltransferase, is responsible for the production of UDP-α-glucose, which then acts as a substrate for RgpE [39]. Ozaki et al. extended the boundary the rgpA-F locus to include three further genes immediately downstream of rgpF. The products of two of these genes, named *rgpH* and *rgpl*, harbour conserved glycosyltransferase domains while it was also determined that *rgpH*, being similar to RgpE, is required for glucose side chain formation, whilst *rgpl* alters the kinetics of RgpE, thus controlling branching frequency [26,40]. The third gene, *orf7*, was later determined to encode a glycerol phosphate transferase which contributes to modification of the Glc side chain structure [41]. The region downstream of *rgpF*, which as mentioned above is responsible for glucose side chain formation, differs significantly between *S. mutans* serotypes *c, e* and *f* [34]. Through serotype conversion studies it was confirmed that this variable region mediates the sero-specificity of *S. mutans* by dictating the presence and position of the glucose linkage [34]. More recent structural analysis of the Rha cwps from *S. mutans* serotypes *c* and *f* has revealed an unexpected fluidity of the structure by identifying minor and major variant polymers for each type [42] (Fig. 3).

These seminal investigations of the biosynthetic assembly of the *S. mutans* Rha cwps have facilitated the unravelling of equivalent pathways in Lancefield streptococci. For example, genes homologous to *rgpG*, *rmlA/C/B/D*, *rgpA* and *rgpB* were identified within the genome of *Streptococcus agalactiae*, a neonate pathogen which expresses the group B carbohydrate (GBC) antigen [43]. The identification of such genes allowed a preliminary assembly pathway to be proposed for the highly complex, multi-antenna structure, which is composed of Rha, galactose (Gal), GlcNac and glucitol residues, linked by phosphodiester bonds [43,44]. *S. agalactiae* was also found to harbour a gene encoding for a TagO/ RgpG like protein named GbcO, which has been shown to initiate GBC synthesis by the transfer of a GlcNac-P moiety to Und-P [45]. Data pertaining to the biosynthesis of group C carbohydrate (GCC) antigen is limited, though a locus with 65% identity to the conserved *rgpA-F* region of *S. mutans* was identified in *Streptococcus dysgalactiae* subsp. *equisimilis* [46]. Cell wall fractions of *S. dysgalactiae* 2023, a clinical bovine isolate, possess two distinct rhamnose-rich polysaccharides termed RRP1 and RRP2. The branched RRP1 is a GCC, while RRP2 is a linear rhamnose-containing polysaccharide [47]. Similarly, the bovine mastitis-causing strain *Streptococcus uberis* 233 possesses two distinct rhamnan structures: a glucose-containing branched rhamnan and an unbranched, linear version [48].

The Group A carbohydrate (GAC) expressed by *Streptococcus pyogenes* has received significant research attention. The GAC structure consists of the α-1,2/α-1,3-linked rhamnan core which is substituted with GlcNac moieties. As with *S. mutans*, the GlcNac substitutions seem to be crucial for the immunogenic response [50,51]. Van Sorge and colleagues identified the genomic locus which encodes the biosynthetic machinery of the GAC structure based on its homology to genes associated with Rha cwps synthesis [52]. This locus encompasses twelve genes, *gacA-L*, and was functionally annotated by mutational analysis. Three predicted rhamnosyltransferase-encoding genes, *gacA*, *gacB* and *gacC* were recalcitrant to disruption and deemed essential for GAC synthesis and viability. Disruption of *gacI*, predicted to encode a glycosyltransferase, *gacJ*, encoding a membrane protein, and *gacK*, which
encodes a Wzx-like transporter [9], resulted in a loss of the GlcNAc-mediated immune response, indicating their involvement in the incorporation of the GlcNAc substitutions [52]. Furthermore, a tarO/gbcO homolog, termed gacO was identified outside of the Rha cwps biosynthesis cluster [52]. Experimental evidence suggests that GacO initiates GAC synthesis through the formation of GlcNAc-P-P-Und, which acts as an acceptor for the first rhamnose moiety of the polyrhamnose core [51]. Further insights into GAC synthesis followed, in which detailed characterisation of the gac genes allowed for the proposition of an assembly pathway [51]. In a similar architecture to that of the S. mutans Rha cwps locus, the leftward end of the gac locus is flanked by an rmlD homolog, gacA [32,51], while the downstream rhamnosyltransferase-encoding genes gacB, gacC and gacG represent functional equivalents of rgpA, rgpB and rgpF, respectively. The ABC-type transport system involved in the export of the polyrhamnose core structure is encoded by gacD and gacE [51]. Notably, the homolog of gacB from S. mutans was found to functionally substitute GacB and restore Rha cwps synthesis during heterologous expression and complementation studies in E. coli harbouring a gacA-GacC-A gacB expressing plasmid, indicating a common initiating step for Rha cwps biosynthesis for species within this genus [29]. Investigations into the enzymatic function of genes associated with GlcNAc side chain formation, gacI, gacJ and gacL determined that GacI harbours GlcNAc-P-Und synthase activity which is enhanced by the membrane protein GacJ. Disruption of the polytopic membrane protein GacL was found to result in an intracellular accumulation of the biosynthetic intermediate β-GlcNAc-P-Und and significant reduction in the amount of GlcNAc in cell wall fractions. It was therefore proposed that GacL transfers a GlcNAc residue to the rhamnose core from GlcNAc-P-Und, the biosynthetic intermediate generated through the activity of GacJ/Gac complex [51]. A subsequent study detected GacH-mediated glycerophosphate modification of the GlcNAc side chain moiety of GAC at a frequency of approximately 25 % [41]. Although rarely discussed in the realm of streptococci, likely due to its non-pathogenic status, the dairy-associated species Streptococcus thermophilus also harbours homologs of genes relating to Rha cwps synthesis, including those of the well characterised S. mutans rgpA-F core synthesis genes [53]. However, as shown by Hols et al., the genetic arrangement and architecture of the cluster is distinct from other species of this genus, as the genes encoding the variable or side-chain structures of the S. thermophilus Rha cwps precede those encoding the core, which are represented by rgpA-F homologs [53]. A presence/absence-based hierarchical clustering of the Rha cwps locus of twenty three industrial strains of S. thermophilus revealed a high level of genetic diversity allowing the assignment of five distinct Rha cwps- associated genotypes, designated A-E [54]. This analysis was expanded by Romero and colleagues to include the Rha cwps locus from S. thermophilus genomes available through NCBI databases in addition to proprietary strains [55]. The boundaries of the locus were also extended with rpoD and radC serving as the relevant start and end flanking genes, noting that these genes play no role in Rha cwps synthesis. 167 strains were found to cluster into one of three major groups (A – C). Group A contains seven subgroups (A1-7), Group B contains six subgroups (B1-6) and Group C, five subgroups (C1-5), with former Groups D and E now represented by A7 and C4, respectively. Furthermore, rgpF was found to
be a genotypic marker for the major groups based on inter- and intra-group sequence identity. A second group-specific attribute was the lack of an rgpE equivalent within strains classified as Group B [55]. Although preliminary functional annotations have been made for genes within the cluster, the overall gene content of the *S. thermophilus* Rha cwps clusters remains poorly characterised with limited experimental proof of the associated protein functions. Interestingly, investigations into the role of the two component signal transduction systems (TCS) TCS06 and TSC07 of *S. thermophilus* LMD-9 in response to bacitracin exposure, indicated that the response regulator of TCS06, encoded by *rr6*, acts as a transcriptional repressor of *rgpA, rgpB* and *rgpC*, and as a transcriptional activator of *rgpl*. The response regulator of TSC07, encoded by *rr7*, is hypothesised to repress transcription of *rmlC*. Taken together, these results suggest a functional role for *S. thermophilus* Rha cwps as an antimicrobial barrier [56]. Furthermore, *tagO* was implicated in an increased acquisition of the integrated conjugative element ICESt3 [57]. This suggests a critical, functional role for *TagO* in maintaining cell wall integrity in *S. thermophilus*, akin to its pathogenic counterparts. A preliminary pathway for Rha cwps core biosynthesis in *S. thermophilus* LMD-9, based on the *S. mutans* model, has been proposed by Thevenard et al. [56].

Compositional analysis of the monosaccharide content of the cell walls from the *S. thermophilus* strains STCH-12 and STCH-15 revealed the presence of rhamnose, galactose and glucose [58]. However, the biochemical structure of an *S. thermophilus* Rha cwps has been elucidated for just a single strain, i.e. ST64987 [59]. The complex structure is composed of a repeating tetrasaccharide core, which is composed of rhamnose and glucose moieties, and which is decorated with tri- and tetra-saccharide side-chains, that possess GlcNAc at their branching points [59].

Overall, Rha cwps synthesis in streptococci is dependent on the presence of the core *rgpA*-F encoded rhamnolysotransferases in addition to the UDP-GlcNAc:Und-P-GlcNAc-1-P-transferase activity encoded by a *tagO* homolog. The individual, strain/species-specific biological characteristics of the Rha cwps structure are underpinned by the diversity of glycolyltransferases encoded by the variable regions of the Rha cwps-associated loci.

(ii) *Lactococcus lactis*

*Lactococcus lactis* is the most extensively applied Lactic Acid Bacterial (LAB) species in commercial and artisanal dairy fermentations. Four subspecies have been characterised – *L. lactis* subsp. *lactis*, *L. lactis* subsp. * cremoris*, *L. lactis* subsp. *hordniae* and *L. lactis* subsp. *tructae* – with subsp. *lactis* and subsp. *cremoris* representing the most important subspecies from a biotechnological perspective [60–62]. *L. lactis* has been subject of focused research attention in recent decades, not only for its ability to impact on the rheological and organoleptic properties of fermented products, for example through plasmid-encoded exopolysaccharide production [63,64], but also for its ability to produce antimicrobial compounds such as nisin, a bacteriocin utilised in food preservation [65,66]. Furthermore, it has become a model Gram-positive organism for the study of phage-host interactions [67–69].

The Rha cwps of lactococci are comprised of two, distinctly synthesised saccharidic components – a highly conserved, neutral, peptidoglycan embedded di- or tri-saccharide rhammam (Rham) and a surface-exposed decoration (previously termed the polysaccharide pellicle (PSP)) which exhibits a high degree of structural variability [10,12,70,71]. The observed structural flexibility of lactococcal Rha cwps is believed to be driven by a need to protect and respond to extra-cellular pressures, in particular phage predation. As such, lactococcal Rha cwps structures appear to be in a state of constant evolution, leading to the emergence of novel features and ultimately divergent cwps groups [10].

The genes responsible for lactococcal Rha cwps biosynthesis are clustered within a single locus, first identified by Dupont et al. in *L. lactis* subsp. *lactis* IL1403 and *L. lactis* subsp. *cremoris* Wg2 through transposon mutagenesis [72]. The leftward end of the lactococcal Rha cwps locus is highly conserved and encompasses genes encoding the enzymatic machinery for the biosynthesis of the rhamnan component including *rmlA-D* and *rgpA-F*, in addition to a lactococcal-specific gene, *wpsJ*. WpsJ is implicated in the transfer of the decorative element to the rhamman core [10,12]. In *silico* characterisation of genes located in the 5′ region allowed the proposition of a model for Rha cwps biosynthesis in *L. lactis*. In keeping with secondary cell wall polysaccharides of other species [9], a TagO ortholog (*limg_1976*) is predicted to initiate rhamnan biosynthesis through its encoded UDP-GlcNAc:Und-P GlcNAc-1-P transferase activity; dTDP-L-Rha is synthesised by the enzymatic activity of the *rmlACBD* gene products; the poly Rha chain is built on the lipid-linked GlcNAc moiety via RgpA, RgpB and RgpF and transported across the membrane by ABC transporter system proteins RgpC and RgpD. Covalent incorporation of the Rha polymer into the peptidoglycan layer is believed to be a function of the LcpA ligase [71].

The rightward (or 3′) end of the locus displays a high level of diversity among different lactococcal strains and encodes the biosynthetic machinery for production of the decorative element or side chain of the Rha cwps [10,73]. The diversity of this region facilitated an initial classification of lactococcal strains into one of three major cwps genotypes – A, B or C [74]. Later studies introduced further diversity through the identification of C type subgroups C1–8 and the novel group D [10,73]. Advances in the structure–function relationship of genes involved in Rha cwps biosynthesis prompted a comparative investigation into the Rha cwps loci of 107 lactococcal strains. The genotypic characterisation of the lactococcal Rha cwps clusters is defined by the variable glycosyltransferase-encoding genetic content of the 3′ region of the cwps locus [10,73]. For example, genotypes A and B lack syn-}
Mutational analysis of genes encoding decorative side chain synthetic functions provided experimental support for the proposed pathway with biochemical analysis of wps gene mutants further corroborating these findings [12].

The side chain subunit structures of distinct strains may be of varying length, while the subunit structures may also be polymerised in some, but not all, cases [10]. Furthermore, the combination and architecture of the component (oligo)saccharides of the Rha cwps is linked to the high degree of specificity of lactococcal phages (this will be discussed in more detail later). Indeed, the rhamnan and side chain structures may each be further decorated through the activity of so-called three component glycosylation systems (TGS) that are encoded by genes located outside the cwps gene cluster [75].

Biological characterisation of the gene pairs – csdAB and csdCD, along with their associated flippase-encoding gene, cflA, determined that they are involved in the glycosylation of rhamnan core and the side chain, respectively. A review of 33 publicly available genomes indicated that csdAB and csdCD are present in nine and ten strains, respectively. The detection of additional genomic loci in particular lactococcal strains, which modify components of the Rha cwps, underpins possibilities for their structural diversification. It also highlights that these modifications are not universal and may be a response to external and environmental pressures [75].

(iii) Enterococcus

The genus Enterococcus was first defined in 1984 following DNA-DNA hybridisation studies, which showed that certain species of enteric streptococci (then named Streptococcus faecalis and Streptococcus faecium) were distantly related to non-enteric streptococci [76]. Although a commensal species in the gut microbiome of mammals and birds, enterococci can also be found in a variety of environmental niches [77]. In recent years, enterococci have gained attention due to the emergence of Enterococcus faecalis and Enterococcus faecium as rapidly evolving opportunistic pathogens, particularly in nosocomial environments [78]. The cell surface of enterococci possesses multiple, secondary glycopolymers [79], including lipoteichoic acids (LTAs), wall teichoic acids (WTAs), variably present capsular polysaccharides (CPS) and the...
ubiquitous enterococcal polysaccharide antigen (EPA), which represents the Rha cwps for this species. The EPA is encoded by a large 40.6 kb cluster of genes arranged in two component modules: the conserved epaA-R region located at the 5′ end, which is involved in the synthesis of the rhamnose-containing EPA core polysaccharide and the 3′ variable region which harbours various glycosyltransferases, epimerases/dehydratases and membrane proteins which encode the biosynthetic machinery for EPA glycans decorations [11].

The presence of an antigenic polysaccharide was first confirmed through the screening of genomic libraries of *E. faecalis* OG1RF and TX52 against patient sera and sera generated from rabbits immunised with *E. faecalis* surface proteins. Seven of the generated cosmids clones were found to only react to patient sera and a representative clone, BO-4861 was insensitive to proteinase K treatment, suggesting the presence of a non-protein antigen such as Rha cwps [80]. Production of the saccharidic antigen was found to be associated with a gene cluster of 43 kb harbouring genes associated with polysaccharide biosynthesis, including the conserved *rml* operon [9,81,82]. Further characterisation defined the region as epaA to epaR and compositional analysis of purified EPA from *E. faecalis* OG1RF confirmed the presence of Glc, Rha, Gal, GlcNAc and N-acetylgalactosamine (GalNAc) residues [83]. Conservation of the epaA-R locus across *E. faecalis* strains was confirmed by hybridisation studies, while variation was observed in the region downstream of epaR between the clinical, vancomycin resistant isolate V583 and OG1RF and in the architecture of the cluster [83,84]. An epa locus harbouring a similar downstream extension of the epaR region was also noted in *E. faecium* TX16. The overall architecture of the core genes was found to differ from the latter *E. faecalis* strain in that epal, epaf and epak were absent, being apparently substituted by epaP and epaQ [85]. Palmer et al. exposed the extent of variation present within the epa loci of both *E. faecalis* and *E. faecium* through a comparative genomics study of enterococcal species [86]. Notably, WTA synthesis–associated genes, tagF, which encodes a teichoic acid polymerase, and tagD, encoding a cytidylyltransferase, which provides activated phosphate for GroP synthesis, were identified within the variable region of both epa loci of *E. faecalis* and *E. faecium* [86]. For a comprehensive schematic overview of epa locus diversity in *E. faecalis* and *E. faecium*, see Palmer et al. [86].

Further divergence of the epa locus was found in both commensal and pathogenic strains of *Enterococcus cecorum*, a bacterium which causes outbreaks of enterococcal spondylitis (ES), a severe disease in broiler chickens characterised by hind limb paralysis and paralytic shell disease in broiler chickens [87,88]. Here, epaA is genetically remote from the epa cluster. The core epa locus for this species is therefore comprised of epaB-H and shares a high level of identity with that of V583. The region downstream of epaH is highly divergent between pathogenic and commensal strains in terms of genetic content and architecture, which is consistent with the variable region of other epa loci [88].

Numerous mutational studies have confirmed that the EPA affects multiple processes including biofilm formation, virulence, host colonisation and immune evasion, conjugative transfer, phage resistance, cell wall integrity, cellular morphology and antimicrobial resistance [83,89–94]. Despite the wealth of knowledge pertaining to the biological functions of EPA, its biochemical structure was not elucidated until 2020 [11]. Detailed NMR analysis of the isolated cell wall fractions of *E. faecalis* VE14089, a plasmid-cured derivative of V583 revealed a highly complex polymer composed of αGlc- and βGlcNAc-substituted rhamnose core onto which teichoic acid decorations are attached by phosphodiester linkages (Fig. 5). Remarkably, the NMR and structural data relating to the teichoic acid decorations were found to be identical to those previously described by Geis et al. as WTA I and WTA II [95]. Interruption of tagB, a glycerophosphate transferase–encoding gene, induced compositional changes to the EPA [95,96] highlighting the interconnected relationship between these secondary glycopolymer structures. Since both WTAs form part of the mature EPA structure, it is logical that ΔtagB mutants not only abolish WTA I and WTA II from the cell surface, but also induced compositional changes in the overall Rha cwps [95]. Of note, both WTA decorations were not detected in an isogenic ΔepaX mutant which is hypothesised to transfer a GalNAc residue to the WTA structure during assembly [11,97]. Additional analysis of whole cells demonstrated that while the WTA decorations are flexible and surface-exposed, the polyrhamnose core is embedded within the cell wall with minimal exposure [11]. Using functional prediction, it has been proposed that the polyrhamnose core is synthesised internally and transported across the membrane by an ABC-type transport system, after which it is modified by the

---

**Fig. 5.** Schematic representation of the complex biochemical structure of the *E. faecalis* V583 EPA in which WTAs are stacked onto the polyrhamnose core. Adapted from Guerardel et al. [11]. Monosaccharide symbols used are based upon the standard Symbol Nomenclature for Glycans (SNFG) [30].
addition of Glc and GlcNAc residues. The WTA decorations are synthesised and modified independently before being linked and transported as a single unit via a Wzx/Wzy pathway. Attachment to the rhamnan core and anchoring to the peptidoglycan layer is believed to be a function of one of five LCP proteins encoded by ef0465, ef1212, ef1569, ef2703 and ef3245, which are located outside of the epa locus, thus producing a mature EPA\[11\]. Very recently, the truncation of galU, a UDP-glucose-1-phosphate uridylyltransferase-encoding gene, has been shown to result in loss of EPA production due to UDP-glucose depletion [98].

**1.2. Shared features of cell wall polysaccharide biosynthesis in ovococcal bacteria**

The elucidation of biosynthesis pathways pertaining to rhamnose-containing cell wall polysaccharides in ovococcal...
species has revealed a number of common enzymatic reactions relating to initiation, elongation and modification. Similarly, it is evident that a modular Rha cwps-associated locus is present in species which produce a complex polysaccharide structure or highly modified linear rhamnan cores [10,12,86]. By focusing on the well characterised Rha cwps biosynthesis pathways of S. pyogenes (GAC), E. faecalis VS83 (EPA) and L. lactis NZ9000, several commonalities emerge with reference to locus architecture, gene function and resulting Rha cwps structure. A schematic overview of the generic Rha cwps biosynthetic pathway is presented in Fig. 6 and specific aspects are discussed further below. Table 1 further highlights the key proteins involved in the synthesis of both the rhamnan core and their respective decorative side chain structures.

(i) Rhamnan core biosynthesis

Rhamnan polysaccharides are believed to be a functional replacement of WTA [9], the synthesis of which is initiated by the UDP-GlcNAc:Und-P-GlcNAc-1-P transferase activity of TagO [99]. As discussed above, it is now evident that ovococcocal species which lack WTA as the major cell wall polymer, have repurposed the function of TagO to initiate synthesis of a Rha-based cwps. Variability exists primarily in the genomic position of the initiating gene. For example, the enterococcal equivalent to tagO, embodied by epaA, is the first gene of the E. faecalis epa locus, yet is remote from the locus in E. ecoum [11,88]. Similarly, the L. lactis tagO homolog, and that of multiple streptococcal species, is positioned at a distinct genetic location from that of the primary Rha cwps locus [11,12,37]. The dTDP-L-rhamnose required for synthesis of the polyrhamnose core is provided by the functions of the rml locus [9]. In the case of L. lactis and E. faecalis, the rmlA-D genes are clustered together at the 5’, conserved end of the cwps locus. In contrast, a distinct architecture is observed in streptococcal species in which rmlA, C, and B are located at an unconnected genomic location [9,11,12,29]. Despite these variations, it appears that all species utilise a common pathway for the establishment of a dTDP-L-Rha pool in the early stages of Rha cwps synthesis.

The addition of the primary rhamnose unit to the undecaprenyl-PP-linked GlcNAc foundation is carried out by the first rhamnosyltransferase encoded within the associated cwps-associated locus of each species. For streptococcal Lancefield Groups A, B, C and G, this rhamnosyltransferase is highly conserved α-D-GlcNAc-β-1-4-L-rhamnosyltransferase [29], whereas in E. faecalis and L. lactis α-1-3/α-1-2 rhamnosyl transferases are predicted to complete the respective rhamnose transfer [11,12]. Extension and polymerisation of the rhamnose chain is a function of the remaining glycosyltransferases within the 5’ conserved region of the respective loci and in all cases an ABC-type transporter system is used to transfer the polyrhamnose structure across the membrane [11,12,51,100].

(ii) Rhamnan core decoration

In the case of the S. pyogenes GAC, the rhamnan backbone component is modified by the addition of an antigenic GlcNAc moiety. Detailed experimental evidence has demonstrated that the UDP-GlcNAc:Und-P-GlcNAc transferase activity of GacI, aided by the membrane protein GacJ, is required for this modification [51,100]. Rush et al. identified GacJ homologs in GBS and E. faecalis, while a similar protein was identified in L. lactis NZ9000 [12]. A review of the gene annotations from both E. faecalis and L. lactis Rha cwps-associated clusters confirms the presence of a gene encoding a DUF2304 domain, typical of GacJ-like proteins, immediately downstream of the gacJ homolog [11,12].

In the case of E. faecalis, this gene pair is embodied by epal-epaJ, and akin to their GAC counterparts, their activity produces a lipid-linked GlcNAc moiety which is transferred to the EPA rhamnan core by one of two membrane proteins EpaP or EpaQ [9]. The modifying Glc residue is provided by the activity of distinct genes and not the epal-epaJ gene pair [11]. For L. lactis the equivalent gene pair, named wpsA-wpsB, is present in all known L. lactis cwps operons [10] and functions as the initiator of decorative side chain synthesis [12]. Overall, it may be concluded that these species utilise a homologous gene pair to produce a lipid-linked GlcNAc moiety, the subsequent transfer of which is a function of species/strain-specific Rha cwps assembly, wherein the GlcNAc moiety may be directly attached to the rhamnan core or be extended to form longer decorative structures.

1.3. Shared features of cell wall polysaccharide biosynthesis in ovococcoid bacteria

For many of the above discussed species, elucidation of the biochemical structures of the Rha cwps is not accompanied by an in-depth characterisation of the variable region of the locus. This substantial knowledge gap leads to difficulty in establishing common structure–function relationships between species with particular reference to those species which produce highly complex rhamnan polysaccharide decorations. Nevertheless, detailed functional assignment of individual genes within the epa and cwps loci of E. faecalis VS83 and L. lactis NZ9000, respectively, have allowed the identification of common biochemical steps (Fig. 7).

For example, the transfer of individual sugar moieties to the growing precursor subunit of the decorative side chain is primarily a function of cytosolic glycosyltransferases containing a PF00535 domain. Furthermore, both systems utilise a Wzx-dependent mechanism for export of the complex decorative structures; in cases where the decorative subunits are polymerised, an encoded Wzy polymerase is employed, and the presence of LicD family proteins is associated with the transfer of Rbo-P units in the case of E. faecalis and Gro-P units for L. lactis cwps types A and B [10,11].

Mahony et al. [10] have recently shown that the gene content of the variable region of the L. lactis Rha cwps locus can be used to predict structural features including decoration, composition and polymerisation, and the presence of unique features such as Gro-P or Galf residues [10]. Therefore, by applying this current dogma to additional species, in conjunction with shared functional predictions based upon bioinformatic data analysis and gene annotation, it is now possible to predict various characteristics of the Rha cwps structure, such as the order of their incorporation, the presence of unique modifications, and the presence of oligo– polysaccharide side chain structures. Informative gene annotations has been a bottleneck for genome-based research for the past two decades; however, such predictive abilities will be transformative to a deeper understanding of the diversity of the associated chemical structures and their implications for strain-specific phenotypes, phage-host interactions among other environmental and industrial considerations. It is well established that phages infecting lactic acid bacteria typically exhibit high host-recognition specificity. The diversity of cwps composition and architecture, the introduction of modifications on both the rhamnan backbone and side chain structures alludes to an external pressure that drives this need for constant modification and diversification. In all ecological niches and in an industrial context in particular, phages are a major driver of host evolution. Below we summarise a number of model phage-host systems in which Rha cwps is defined to mediate phage binding and infection.
1.4. Rhamnose cell wall polysaccharides as phage receptors in ovococci

(i) Streptococcus

The nature of the host-encoded phage receptor of haemolytic streptococci has not been studied in detail to date. Early studies of phage-host interactions of haemolytic streptococci uncovered a group antigen-specific pattern of infection in which purified Group C antigen was found to neutralise Group C-infecting phages, but not those which infect Group A strains. In contrast, purified Group A carbohydrate was unable to inactivate Group A infecting phages [101] and phage A25, which infects Group A strains is hypothesised to require intact cell wall and peptidoglycan fractions for optimal adsorption [102]. The temperate phage, P9, has been reported to adsorb to the purified Group C antigen of both S. equi and S. zooepidemicus, yet is unable to infect or replicate on the latter [103,104]. The Group C antigen has also been identified as the probable receptor for the prototype phage C1 with specific reference to the importance of the GalNAc moiety [105,106]. In contrast to the above, the choline residues of TA have been identified being pivotal for infection of S. pneumoniae by phage Dp-1 [107,108].

Significant data defining the role of the surface antigen in phage adsorption for the non-haemolytic species S. mutans has been provided. Phage M102 is specific to serotype c strains of S. mutans and cannot adsorb to serotypes e, f or k, and the Glc side chain of the serotype c antigen has been shown to be critical for both adsorption and infection by this phage [26]. Unequivocal proof was provided through sero-conversion experiments, in which the exchange of a S. mutans type e sero-specific locus with that of an S. mutans type c locus (identified as the region between rggF and orf12) led to a dramatic increase in phage adsorption [26]. In addition, phage M102 is also unable to adsorb to strains that carry mutations in the rggA, rggB or rggC genes [26]. The host-encoded receptor for phages of S. thermophilus are also believed to be saccharidic in nature [109,110]. Preliminary evidence suggests that mutations within a glycosyltransferase encoding gene of the Rha cwps-associated locus confers a phage resistance phenotype, yet this remains to be experimentally confirmed [54,58].

(ii) Lactococcus

Early studies of phage-host interactions in L. lactis identified rhamnose-containing cell wall components as the receptor for phage kh [111], while mutations within genes of the Rha cwps locus of lactococcal strains cause phage insensitivity due to adsorption deficiency as mentioned above. Remarkably, the host range of the prolific 936 lactococcal phage species (now termed the Skuviruses) and their associated receptor binding phylogeny, can be directly correlated to the Rha cwps genotype [74]. Genotype swapping of the C2 variable region of L. lactis 3107 with that of the C1 type strain L. lactis NZ9000 confirmed that the locus also mediates infection by the P335 phages φLc3 and TP901-1 [73]; however, a direct correlation to host range and Rha cwps genotype could not be established for this genetically heterogenous phage species [112]. The less commonly encountered lactococcal phage groups including 949 and P087, have also been shown to recognise moieties encoded by the Rha cwps-associated locus [113], while the Ceduoviridae (formerly termed c2 phages) recognise and bind to both an as yet undefined saccharidic component, and a protein (phage infection protein, PIP) receptor on the cell surface [114].

(iii) Enterococcus

The EPA has been linked to phage infection in siphophage φNPV1, which is unable to form plaques on Δ epaB, M, E and N derivatives of the host strain [83]. A later study by Chatterjee et al. involving transposon mutants identified the glycosyltransferases epaOX and epaOX2 of E. faecalis OG1RF as critical for adsorption of φVPE25 and noted that disruption of the O-antigen ligase, epaOY significantly affects plaquing efficiency [115]. Furthermore, an independent transposon mediated disruption of an encoded LytR-type response regulator was found to downregulate transcription of epaOX, epaOX2 and epaOY, resulting in a 40 % reduction of the absorbance of φVPE25 [115]. The role of the EPA in host attachment is not limited to Siphoviridae phages. Adsorption of phage Idefix, which belongs to the Podoviridae family, is dependent on epaX, encoding a glycosyltransferase which is
Table 2
Summary of EPA associated proteins which have been shown to be critical for phage infection.

| Protein | Function                          | Location   | Phage       |
|---------|-----------------------------------|------------|-------------|
| EpaA    | Initiation of rhamnan biosynthesis| Conserved  | NPV1        |
| EpaB    | GT – rhamnan elongation           | Conserved  | NPV1        |
| EpaR    | UDP Glc phosphotransferase        | Conserved  | NPV1, phi4, phi51, phi47, phi19 |
| EpaX    | GT – GalNAc decoration of TA chain| Variable   | phi4        |
| EpaY    | Membrane protein – polymerisation TA blocks| Variable   | phi4        |
| EpaM    | ABC transporter – translocation of rhamnan| Conserved  | NPV1        |
| EpaN    | GT – rhamnose elongation          | Conserved  | NPV1        |
| EpaE    | rmLA – Glc-1-P thymidylytransferase| Conserved  | NPV1        |
| EpaOX   | GT - GalNAc decoration of TA chain| Variable   | VPE25       |
| EpaOX2  | GT                               | Variable   | VPE25       |
| EpaOY   | Membrane protein – polymerisation TA blocks| Variable   | Phi4, phi47, phi19 |
| EpaS    | GT - GalNAc decoration of TA chain| Variable   | VPE25       |
| EpaW    | Epimerase – conversion to UDP-Rha | Variable   | Phi17       |
| EpaAC   | Epimerase – UDP-GlcNAc to UDP-GalNAc | Variable | Phi4, phi19, phi47 |

hypothesised to add a GalNAc moiety to the teichoic acid decorative chain of the EPA [117,118], whilst infection of myophages EF1TV, phi17 and phi19 also involves EPA-associated genes [117,118].

Duerkop et al. [119] additionally identified a conserved integral membrane protein in E. faecalis V583 which is essential for infection by siphophages ΦVPE25 and ΦFW. The protein, named PIP<sub>3P</sub>, due to homology to its lactococcal counterpart [114], was found to contain a central, hypervariable region of 160 aa which directly correlates to host range. Investigations into phage resistant isolates which harboured mutations within PIP<sub>3P</sub> determined that it facilitates DNA injection, while it is not involved in adsorption to the host [119]. Thus overall, successful infection by ΦVPE25 is a two-step process of initial EPA-mediated attachment followed by PIP<sub>3P</sub>-activated DNA injection [115,119]. Similarly, ΦNPV1 requires both PIP<sub>3P</sub> and a functional epdR for successful infection of E. faecalis OG1RF [120]. Of note, it has recently been suggested that additional and, as yet unidentified DNA injection triggers exist for E. faecalis phages [117,121]. A summary of EPA-associated proteins shown to be complicit in phage infection is shown in Table 2.

2. Summary and outlook

Recent studies have provided significant insights into the biosynthesis of Rha cwps for key ovococcal species. The identification of a direct relationship between gene content and architecture, shared gene function and Rha cwps structure, provides an in silico template for predicting key structural features for strains or species where biochemical data is lacking. We have highlighted significant commonalities for both gene function and biosynthesis of ovococcal Rha cwps. It has recently been proposed that the nomenclature of the S. mutans serotype C Rgp, and its associated genes, be changed to SCC (sero-specific carbohydrate C) and SCC, respectively, to better reflect their function [29]. As common intra-species functions have been identified for Rha cwps biosynthetic genes, a future harmonisation of the nomenclature across both pathogenic and non-pathogenic ovococcal species will further our understanding of ovococcal Rha cwps structure, diversity, biosynthesis, and functionalities. Furthermore, the evolution and genetic diversity of Rha cwps loci is heavily influenced by external pressures, in particular, phage predation. Longitudinal and comparative mapping of phage induced mutations in Rha cwps-associated genes may serve as a tool for elucidating critical, or indeed, shared elements of biosynthesis in ovococcal species, which mediate phage sensitivity. In addition, the rapid response of phage, at times by single point mutation, to overcome and adapt to Rha cwps modifications allows for the identification and/or characterisation of phage-encoded host recognition machinery which specifically recognise Rha cwps on the host cell surface.

CRediT authorship contribution statement
Katherine Lavelle: Writing – original draft. Douwe van Sinderen: Writing – review & editing. Jennifer Mahony: Writing – review & editing.

Declaration of Competing Interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements
JM is in receipt of a Starting Investigator Research Grant (SIRG) (Ref. No. 15/SIRG/3430) funded by Science Foundation Ireland (SFI). DvS is supported by a Principal Investigator award (Ref. No. 13/IA/1953) through SFI and by APC Microbiome Ireland (under Science Foundation Ireland (SFI) grant number: SFI/12/RC/2273-P1 and SFI/12/RC/2273-P2).

Funding
This publication has emanated from research conducted with the financial support of supported in part by a grant from Science Foundation Ireland under Grant numbers 12/RC/2273-P2 and 15/SIRG/3430. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References
[1] Zapun A, Vernet T, Pinho MG. The different shapes of cocci. FEMS Microbiol Rev. 2008;32(2):345–60.
[2] Higgins ML, Shockman GD. Model for cell wall growth of Streptococcus faecalis. J. Bacteriol. 1970;101(2):643–8.
Chapot-Chartier M-P. Interactions of the cell-wall glycopolymers of lactic acid bacteria and their bacteriophages. Front. Microbiol. 2014;5(236).

Neuhaus PC, Baddiley J. A continuum of anionic charge: structures and functions of D-alanyl teichoic acids in gram-positive bacteria. Microbiol Mol Bacteriol. 2003;67(4):686–723.

Choi EH, Zhang F, Lu Y-J, Malley R. Capsular polysaccharide (CPS) release by serotype 3 pneumococcal strains reduces the protective effect against type-3 CPS antibodies. Clin Exp Immunol. 2016;2;162–7.

Kawai Y, Marles-Wright J, Cleverley RM, Emmens R, Ishikawa S, Kuwano M, et al. 2011. A widespread family of bacterial cell wall assembly proteins. EMBO J 30(24):4931–41.

Chao AC, Frank J, Dengler V, Schneewind O, Missiakas D. 2013. Staphylococcus aureus mutants lacking the LytR-CysP-A-Psr family of enzymes release cell wall teichoic acids into the extracellular medium. J Bacteriol. 195(20):4650–9.

Mistou M-Y, Sutcliffe IC, van Sorge NM. Bacterial glycobiology: rhombo-hemolytic cell-wall polysaccharides in Gram-positive bacteria. FEMS Microbiol Rev. 2016;40(4):464–79.

Mahony J, Frantzien C, Vinogradov E, Sadowskaya I, Theodorou I, Kelleher P, et al. The CWPS Rubik's cube: Linking diversity of cell wall polysaccharide structures with the encoded biosynthetic machinery of selected Lactococcus lactis strains. Mol Microbiol. 2020;114(4):582–96.

Guerrardel Y, Sadowskaya I, Maes E, Furlan S, Chapot-Chartier M-P, Mesnage S, et al. Complete sequence of the enterococcal polysaccharide antigen (EPa) or vancomycin-resistant Enterococcus faecalis V583 reveals that EPA decorations are teichoic acids covalently linked to a rhamnopolysaccharide backbone. mBio. 2020;11(2):e02677–e220.

Theodorou I, Courin P, Palaisserie S, Kulakauskas S, Bidnenko E, Chabauty C, et al. A dual-chain assembly pathway generates the high structural diversity of cell-wall polysaccharides in Lactococcus lactis. J. Biol. Chem. 2019;294:17612–23.

Hardie JM, Whiley RA. Classification and overview of the genera Streptococcus and Enterococcus. J. Appl. Microbiol. 1987;53:11–115.

Sitkiewicz I, Hryniewicz W. Pyogenic streptococci-danger of re-emerging pathogens. J. Med. Microbiol. 1997;46:1563–64.

McCarty M, Lancefield RC. 1955. Variation in the group-specific carbohydrate of beta-hemolytic streptococci by latex agglutination. J. Clin Microbiol. 1978;8(3):326–8.

Slade HD, Slamp WC. Cell-wall composition and the grouping antigens of Streptococcus and Enterococcus. J. Appl. Micrbiol. 1997;83(S1):1S–11S.

Sutcliffe IC, Black GW, Harrington DJ. Bioinformatic insights into the pathophysiology, morphogenesis and cellular division. Mol. Microbiol. 2019;112(3):944–58.

van der Beek SL, Le Breton Y, Ferenbach AT, Chapman RN, van Aalten DMF, et al. Characterization of streptococcal rgp genes required for rhamnan synthesis in Streptococcus mutans. Infect. Immun. 2015;25(12):1323–4.

van Sorge NM, Cole JN, Kuipers K, Henningham A, Aziz RK, Kasiere-Friese A, et al. The classical lancedefier antigen of group A Streptococcus is a virulence determinant with implications for vaccine design. Cell Host Microbe. 2014;15(6):729–40.

Holst P, Hancy F, Fontaine L, Grossiord B, Prozzi D, Leblond-Bourget N, et al. New insights in the molecular biology and physiology of Streptococcus mutans revealed by comparative genomics. FEMS Microbiol Rev. 2005;29(3):435–63.

Szymczak P, Rau MH, Monteiro JM, Pinho MG, Filipe SR, Vogensen FK, et al. A comparative genomic approach for identifying putative counterparts of carbohydrate biosynthesis loci in Streptococcus thermophilus. Microb. Genomics. 2019;5(1):1–12.

Rush JS, Edgar RJ, Deng P, Chen J, Zhu H, van Sorge NM, et al. A novel gene required for rhamnose-glucose polysaccharide synthesis in Streptococcus mutans. J. Bacteriol. 1999;181(20):6556–9.

Leher J, Vigeant KA, Tataf LD, Valvano MA. Functional characterization and membrane topology of an rhamnose polylipopolysaccharide. J. Biol. Chem. 2007;189(7):2618–28.

De A, Liao S, Bitoun JP, Roth K, Beatty WL, Wu H, et al. Deficiency of RgpG causes major defects in cell division and biofilm formation, and deficiency of LytR-CpsA-Psr family proteins leads to accumulation of cell wall antigens in culture medium by Streptococcus mutans. Appl. Environ. Microbiol. 2017;83(17):e00927–e1017.

Kovacs OJ, Faustofeter RC, Bischer AP, Quivey Jr RC. Streptococcus mutans requires mature rhamnose-glucose polysaccharides for proper pathophysiology, morphogenesis and cellular division. Mol. Microbiol. 2019;112(3):944–58.

Yamashita Y, Tsukoiwa Y, Nakano Y, Tomishita K, Oho T, Koga T. Biological functions of UDP-glucose-synthesis in Streptococcus mutans. Microbiology (Reading, England). 1998;144(5):1233-45.

Ozaki K, Shibata Y, Yamashita Y, Nakano Y, Tsuda H, Kido N, Ohma M, et al. A novel mechanism for group antigens and side chains in the biosynthesis of rhamnose-glucose polysaccharide synthesis. FEBS Lett. 2002;532(1–2):159–63.

Edgar RJ, van Hensbergen VP, Ruda A, Turner AG, Deng P, Le Breton Y, et al. Discovery of glyceral phosphate modification on streptococcal rhamnose polysaccharides. Nat. Chem. 2019;11(5):639–66.

Michon F, Brisson JR, Dell A, Kasper DL, Jennings HJ, Multianzymatic group-specific polysaccharide of group B Streptococcus. Biochemistry 1988;27(14):3341–51.

St Michael F, Yang Q, Cairns C, Vinogradov E, Fleming P, Hayes AC, et al. Investigating the candidacy of the serotype specific rhamn polysaccharide based glycoconjugates to prevent disease caused by the dental pathogen Streptococcus mutans. Glycoconjugate J. 2018;35(1):135–44.

Colant E, Dramsi S, Chapot-Chartier M-P, Courtin P, Kulakauskas S, Pichoux C, et al. Role of the group B antigen of Streptococcus agalactiae: A peptidoglycan-anchored polysaccharide involved in cell wall biogenesis. PLoS Pathog. 2012;8(6):e1002756.

Van der Beek SL, de Groot S, Kinkae T, Miyoshi-Aikyama T. Complete genome sequence of Streptococcus dysgalactiae subsp. equisimilis 167 carrying Lancefield group C antigen and comparative genomics of S. dysgalactiae subsp. equisimilis strains. Genome Biol. Evol. 2013;5(9):1644–51.

Nakano K, Nomura R, Oshima T. Streptococcus mutans and cariovascuar disease. Jpn. Dent. Sci. Rev. 2008;44(1):29–37.

Colgan JE, Kundi TJ, Krause RM. Structure of the streptococcal group A, A-variant and C carbohydrates. Immunochimistry. 1978;15(10):755–60.

Rush JS, Edgar RJ, Deng P, Chen J, Zhu H, van Sorg NM, et al. The molecular mechanism of N-acetylglucosamine side-chain attachment to the Lancefield A carbohydrate in Streptococcus pyogenes. J. Biol. Chem. 2017;292(47):19441–57.

van Sorge NM, Cole JN, Kuipers K, Henningham A, Aziz RK, Kasiere-Friese A, et al. The classical lancedefier antigen of group A Streptococcus is a virulence determinant with implications for vaccine design. Cell Host Microbe. 2014;15(6):729–40.

Holst P, Hancy F, Fontaine L, Grossiord B, Prozzi D, Leblond-Bourget N, et al. New insights in the molecular biology and physiology of Streptococcus thermophilus revealed by comparative genomics. FEMS Microbiol Rev. 2005;29(3):435–63.

Szymczak P, Rau MH, Monteiro JM, Pinho MG, Filipe SR, Vogensen FK, et al. A comparative genomic approach for identifying putative counterparts of carbohydrate biosynthesis loci in Streptococcus thermophilus. Microb. Genomics. 2019;5(1):1–12.

Romero DA, Magill D, Millen A, Horvath P, Fremaux C. Dairy lacticoccal and streptococcal phage-host interactions: an industrial perspective in an evolving phage landscape. FEMS Microbiol. Rev. 2020;44(6):30532. I. Shevendar B, Besset C, Choinard S, Fourcassé P, Boyaval P, Monnet V, et al. Response of S. thermophilus LMD-9 to bacitracin: Involvement of a CeresA/BA-like module and of the rhamnose-glucose polysaccharide synthesis pathway. Int. J. Food Microbiol. 2014;177:89–97.
Szymczak P, Filipe SR, Covas G, Vogensen FK, Neves AR, Janzen T. Cell wall glycans mediate recognition of the dairymaker Streptococcus thermophilus by bacteriophages. Appl. Environ. Microbiol. 2019;85(2):e02109–e02117.

McDonnell B, Hanemaaijer L, Bottacini F, Kelleher P, Lavelle K, Sadowski I, et al. A cell wall-associated polysaccharide is required for bacteriophage adsorption to the Streptococcus thermophilus cell surface. Mol. Microbiol. 2020;114(1):31–45.

Parapouli M, Delbès-Paus C, Koukkou A-I, Montel M-C, Samelis J. Comparative genomic analysis identifies divergent genomic features of pathogenic Enterococcus cecorum including a Type IV CRISPR-Cas system, a capsular locus, an ecp-like locus, and putative host cell binding proteins. PLoS ONE 2015;10(4):e0121294.

Dahmane N, Robert E, Deschamps J, Meylheuc T, Delorme C, Briandet R, et al. Cell surface of L. lactis Trialu, P. et al. Cell surface of Lactococcus lactis subsp. lactis, subsp. lactis biovar diacetylactis isolated from a mixed culture of cheese. Int. J. Food Microbiol. 2013;158(2–3):275–84.

Legrand P, Collins B, Blangy S, Murphy J, Spinelli S, Gutierrez C, et al. The cell wall polysaccharides of Enterococcus faecium DT104 are required for binding to the human immune system. PLoS Pathog. 2019. 15(5):e1007730-e.

Singh KV, Murray BE. Loss of a major enterococcal polysaccharide antigen (Epa) by Enterococcus faecalis is associated with increased resistance to ceftriaxone and carbapenems. Antimicrob. Agents and Chemotherapy. 2019;63:e00481–e00519.

Krause RM. Studies on the bacteriophages of hemolytic streptococci : I. Factors influencing the interaction of phage and susceptible host cell. J. Exp. Med. 1957;106(3):365–84.

Giraud MF, Naismith JH. The rhamnose pathway. Curr Opin Struct Biol. 2002;12(4):467–72.
[108] Garriss G, Henriques-Normark B. Lysogeny in Streptococcus pneumoniae. Microorganisms 2020;8(10):1546.

[109] Binetti AG, Quiberoni A, Reinheimer JA. Phage adsorption to Streptococcus thermophilus. Influence of environmental factors and characterization of cell-receptors. Food Res. Int. 2002;35(1):73–83.

[110] Quiberoni A, Stiefel J, Reinheimer JA. Characterization of phage receptors in Streptococcus thermophilus using purified cell walls obtained by a simple protocol. J. Appl. Microbiol. 2000;89(6):1059–65.

[111] Valyasevi R, Sandine WE, Geller BL. The bacteriophage kh receptor of Lactococcus lactis subsp. cremoris K11 is the rhamnose of the extracellular wall polysaccharide. Appl Environ Microbiol. 1990;56(6):1882–9.

[112] Mahony J, Oliveira J, Collins B, Hanemaaijer L, Lugli GA, Neve H, et al. Genetic and functional characterization of the lactococcal P335 phage-host interactions. BMC Genomics 2017;18(1):1–16.

[113] Mahony J, Randazzo W, Neve H, Settanni L, van Sinderen D. Lactococcal 949 group phages recognize a carbohydrate receptor on the host cell surface. Appl. Environ. Microbiol. 2015;81(10):3299–305.

[114] Valyasevi R, Sandine WE, Geller BL. A membrane protein is required for bacteriophage c2 infection of Lactococcus lactis subsp. lactis C2. J. Bacteriol. 1991;173(19):6095.

[115] Chatterjee A, Willett JLE, Nguyen UT, Monogue B, Palmer KL, Dunny GM, et al. Parallel genomics uncover novel Enterococcal-bacteriophage interactions. mBio. 2020;11(2):e03120–e3219.

[116] Lossouarn J, Briet A, Moncaut E, Furlan S, Bouteau A, Son O, et al. Enterococcus faecalis countermeasures defeat a virulent picovirinae bacteriophage. Viruses. 2019;11(1).

[117] Chatterjee A, Johnson CN, Luong P, Hullahalli K, McBride SW, Schubert AM, et al. Bacteriophage resistance alters antibiotic-mediated intestinal expansion of Enterococci. Infect. Immun. 2019;87(6):e00085–e119.

[118] D’Andrea MM, Frezza D, Romano E, Marmo P, Henrič De Angelis L, Perini N, et al. The lytic bacteriophage vB_EfaH_EF1TV, a new member of the Hereleviridae family, disrupts biofilm produced by Enterococcus faecalis clinical strains. J. Glob. Antimicrob. Resist. 2020;21:68–75.

[119] Duerkop BA, Huo W, Bhardwaj P, Palmer KL, Hooper LV. Molecular basis for lytic bacteriophage resistance in Enterococci. mBio. 2016;7(4):e01304–e1316.

[120] Ho K, Huo W, Pas S, Dao R, Palmer KL. Loss-of-function mutations in epaR confer resistance to eNPV1 infection in Enterococcus faecalis OG1RF. Antimicrob. Agents. Chemother. 2018;62(10).

[121] Al –Zubidi M, Widziolek M, Court EK, Gains AF, Smith RE, Ansbro K, et al. Identification of novel bacteriophages with therapeutic potential that target Enterococcus faecalis. Infect. Immun. 2019;87(11):e00512–19.