

# The N-terminal domain of rhamnosyltransferase EpsF influences exopolysaccharide chain length determination in *Streptococcus thermophilus* 05-34

Guohong Wang<sup>1</sup>, Jiayi Li<sup>1</sup>, Shuxin Xie<sup>1</sup>, Zhengyuan Zhai<sup>1</sup> and Yanling Hao<sup>1,2</sup>

<sup>1</sup> Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Food Science and Nutritional Engineering, China Agricultural University, Beijing, China

<sup>2</sup> Key Laboratory of Functional Dairy, Co-Constructed by Ministry of Education and Beijing Municipality, Beijing, China

## ABSTRACT

Glycosyltransferases are key enzymes involved in the assembly of repeating units of exopolysaccharides (EPS). A glycosyltransferase generally consists of the N-terminal and the C-terminal domain, however, the functional role of these domains in EPS biosynthesis remains largely unknown. In this study, homologous overexpression was employed to investigate the effects of EpsF<sub>N</sub>, a truncated form of rhamnosyltransferase EpsF with only the N-terminal domain, on EPS biosynthesis in *Streptococcus thermophilus* 05-34. Reverse transcription qPCR and Western blotting analysis confirmed the successful expression of *epsF<sub>N</sub>* in 05-34 at the transcription and translation level, respectively. Further analysis showed that the monosaccharide composition and yield of EPS were not affected by the overexpression of *epsF<sub>N</sub>*, whereas the molecular mass decreased by 5-fold. Accordingly, the transcription levels of genes involved in EPS biosynthesis, including chain-length determination gene *epsC*, were down-regulated by 5- to 6-fold. These results indicated that the N-terminal domain of EpsF alone could influence the molecular mass of EPS, probably via lowering the concentration of sugar precursors, which may lead to decreased expression of genes responsible for chain-length determination.

Submitted 5 February 2019  
Accepted 7 January 2020  
Published 12 February 2020

Corresponding author  
Yanling Hao, haoyl@cau.edu.cn

Academic editor  
Joseph Gillespie

Additional Information and  
Declarations can be found on  
page 13

DOI 10.7717/peerj.8524

© Copyright  
2020 Wang et al.

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

**Subjects** Food Science and Technology, Genetics, Microbiology, Molecular Biology

**Keywords** *Streptococcus thermophilus*, Chain length, N-terminal domain of rhamnosyltransferase, Exopolysaccharides

## INTRODUCTION

Microbial exopolysaccharides (EPS) are a wide group of secreted polymers that can be assembled as capsular polysaccharides (CPS) tightly associated with cell surface, or released as extracellular slime in the surrounding of the cell (Kleerebezem *et al.*, 1999; Nwodo, Green & Okoh, 2012). Among microbial EPS, those produced by lactic acid bacteria (LAB) have received increasing attention for conferring textural and rheological properties to fermented products (Lebeer *et al.*, 2009; Caggianiello, Kleerebezem & Spano,

2016). Many of the EPS produced by LAB strains are heteropolysaccharides, which are built up from repeating units consisting of two or more types of monosaccharides, such as galactose, glucose and rhamnose (Remus *et al.*, 2012; Caggianiello, Kleerebezem & Spano, 2016). Heteropolysaccharides are commonly synthesized by the Wzx/Wzy-dependent pathway reported for EPS and O-antigen biosynthesis in Gram-negative bacteria (Whitfield, 1995). In this pathway, first, a priming glycosyltransferase catalyzes the transfer of a sugar-1-phosphate from a nucleotide diphospho-sugar to the undecaprenyl (C55) phosphate lipid moiety. Subsequently, a series of glycosyltransferases catalyze the addition of other sugar residues until the repeat unit is completed. Then, the lipid-linked repeat unit is translocated across the membrane by the flippase Wzx, and subsequently polymerized by the polymerase Wzy (Islam & Lam, 2014). In Gram-positive bacteria, a membrane protein complex Wzd/Wze (designated EpsC/EpsD in LAB) is postulated to be responsible for chain length determination and secretion of the mature EPS (Bentley *et al.*, 2006; Lebeer *et al.*, 2009).

Glycosyltransferases play important roles in the biosynthesis of repeating units of EPS via catalyzing the formation of glycosidic bonds between sugar residues. Donor sugars for these reactions are usually activated in the form of nucleoside diphosphate sugars, such as UDP-Glucose, UDP-Galactose and dTDP-Rhamnose. Transfer of the sugar residue results in either inversion or retention of the anomeric stereochemistry of the donor sugar, as such, glycosyltransferases are classified as inverting or retaining glycosyltransferases, respectively (Lairson *et al.*, 2008). Also, glycosyltransferases are currently classified into 107 families in the CAZy database ([www.cazy.org](http://www.cazy.org)) based on amino acid sequence similarities (Campbell *et al.*, 1997; Coutinho *et al.*, 2003). Moreover, two structural folds, GT-A and GT-B, have been identified for the nucleotide sugar-dependent glycosyltransferases (Breton *et al.*, 2006; Lairson *et al.*, 2008). Both the GT-A and GT-B folds contain two Rossmann-like domains ( $\beta/\alpha/\beta$ ), which are proposed to be involved in substrate binding (Lesk, 1995). Several structural studies of glycosyltransferases showed that domains responsible for donor binding are located on the N-terminal of GT-A glycosyltransferases and the C-terminal of GT-B type, respectively (Martinez-Fleites *et al.*, 2006; Steiner *et al.*, 2010). Additionally, positively charged residues and  $\alpha$ -helices on the N-terminal region of several GT-B glycosyltransferases can also participate in stabilizing the phosphate group of donor sugars (Albesa-Jové *et al.*, 2014). Upon binding of the donor, glycosyltransferases may undergo a conformational change, facilitating the binding of acceptor molecules and the subsequent sugar transfer (Qasba, Ramakrishnan & Boeggeman, 2005).

Construction of N- or C-terminal truncated forms of enzymes has been employed to investigate the structure–function relationships of glucosyltransferase and fucosyltransferase in homopolysaccharides biosynthesis (Bastida, Fernández-Mayoralas & García-Junceda, 2002; Del Moral *et al.*, 2008). It has been revealed that deletion of the C-terminal 65 amino acids of a  $\alpha$ -1,6-fucosyltransferase from *Rhizobium* sp. resulted in its increased affinity to GDP-Fucose donor (Bastida, Fernández-Mayoralas & García-Junceda, 2002). In addition, a  $\alpha$ -1,2-fucosyltransferase lacking its N-terminal domain was reported to lose glycosyltransferase activity in *Helicobacter pylori*

(Wang *et al.*, 1999). More recently, it has been demonstrated that truncation of domain V of the glucosyltransferase GTF180 resulted in increased yield, yet decreased molecular-mass of oligosaccharides produced by *Lactobacillus reuteri* (Meng *et al.*, 2015). These studies suggest possible roles of specific domains of glycosyltransferases in determining enzyme activity, substrate specificity, and properties of EPS produced.

The in situ production of EPS by ropy *Streptococcus thermophilus* strains has been strongly associated with the improved viscosity, water retention and the mouthfeel of yogurts (Broadbent *et al.*, 2003; Purwandari, Shah & Vasiljevic, 2007). Our previous studies have shown that yogurt fermented with an EPS producing strain *S. thermophilus* 05-34 exhibited enhanced physical and sensory properties (Qin *et al.*, 2011). Recently, a typical region encoding genes involved in EPS biosynthesis has been detected in the draft genome sequence of *S. thermophilus* 05-34 (accession number: QFLC00000000). Furthermore, a gene *epsF<sub>N</sub>* (DIS31\_02545), encoding the N-terminal region of a rhamnosyltransferase EpsF, was found within this *eps* cluster. Similarly, the truncated version of EpsF has also been found in several other *S. thermophilus* strains (Li *et al.*, 2018; Wu *et al.*, 2015). In this study, the effects of this N-terminal region of rhamnosyltransferase on the biosynthesis of EPS in 05-34 were investigated.

## MATERIALS AND METHODS

### Bacterial strains, plasmids and growth conditions

The bacterial strains and plasmids used in this study are listed in Table 1. *S. thermophilus* was routinely maintained in de Man–Rogosa–Sharpe (MRS) broth at 37 °C. For EPS isolation, *S. thermophilus* was cultured in 10% (w/v) sterilized reconstituted skim milk (RSM). *Lactococcus lactis* NZ9000, used as a host in cloning experiments, was cultured at 30 °C in M17 broth supplemented with 0.5% (w/v) D-glucose (GM17 medium). When required, chloramphenicol (Cm) was added at 5 µg/mL for both *L. lactis* and *S. thermophilus*.

### Construction of EpsF<sub>N</sub> overexpression strain and Western blot analysis

In order to realize constitutive expression, a 318-bp DNA fragment was synthesized as Fig. S1, which consists of the ribosomal-binding site (RBS) and the constitutive promoter PslpA from the gene encoding the S-layer surface protein of *Lactobacillus acidophilus* (Boot *et al.*, 1996). This DNA fragment was synthesized by Sangon (Beijing, China) and then inserted into the vector pNZ8148 between *Bgl*II and *Nco*I, resulting in the expression vector pSlpA-8148. Subsequently, the *epsF<sub>N</sub>* gene (DIS31\_02545) was amplified by PCR from the genomic DNA of *S. thermophilus* 05-34 with primers F-epsF (5'-CATGCCATGGGAAC AAAAACAGTTTATATCG-3') and R-epsF (5'-CGCGAGCTCCTATGAGCTTGTAGG ACTATTC-3'). Restriction sites used for subsequent cloning are underlined: *Nco*I and *Sac*I for F-epsF and R-epsF, respectively. The amplicon obtained was digested by *Nco*I and *Sac*I, and then inserted into pSlpA-8148. The ligation mixture was introduced into *L. lactis* NZ9000 by electroporation (Holo & Nes, 1989), and transformants were selected on GM17 agar supplemented with Cm. Plasmids were isolated from *L. lactis* transformants using

**Table 1** Bacterial strains and plasmids used in this study.

Strain or plasmid	Relevant characteristics	Source of reference
Strains		
<i>S. thermophilus</i> 05-34	EPS-producing strain isolated from Tibetan kefir grains	CGMCC16303 <i>Qin et al. (2011)</i>
<i>S. thermophilus</i> 05CK	<i>S. thermophilus</i> 05-34 carrying pSlpA-8148	This study
<i>S. thermophilus</i> 05epsF	<i>S. thermophilus</i> 05-34 carrying pSlpA-epsF	This study
<i>S. thermophilus</i> 05Fh <sub>6</sub>	<i>S. thermophilus</i> 05-34 carrying pSlpA-epsFh <sub>6</sub>	This study
<i>L. lactis</i> NZ9000	Plasmid-free derivative of <i>L. lactis</i> MG1363 <i>pepN::nisRK</i>	<i>Kuipers et al. (1998)</i>
Plasmids		
pNZ8148	Cm <sup>r</sup> , inducible expression vector carrying the <i>nisA</i> promoter PnisA	<i>Mierau &amp; Kleerebezem (2005)</i>
pSlpA-8148	Cm <sup>r</sup> , pNZ8148 derivative carrying constitutive promoter PslpA instead of PnisA	This study
pSlpA-epsF	Cm <sup>r</sup> , pSlpA-8148 derivative carrying the <i>epsF<sub>N</sub></i> gene	This study
pSlpA-epsFh <sub>6</sub>	Cm <sup>r</sup> , pSlpA-8148 derivative carrying the <i>epsF<sub>N</sub></i> gene with a 3'-sequence encoding a C-terminal His-tag	This study

QIAGEN Miniprep Spin Kit (Qiagen Inc, Germany), and further verified by restriction analysis and sequencing using the primer R-8148 (5'-CAATCAAAGCAACACGTG-3'). The resulting recombinant plasmid pSlpA-epsF was transformed into *S. thermophilus* 05-34 by electroporation in a 0.2 cm cuvette at 1.5 kV, 25 μF and 200 Ω (*Marciset & Mollet, 1994*), and the recombinant strain was designated as *S. thermophilus* 05epsF.

To confirm the overexpression of EpsF<sub>N</sub> in *S. thermophilus* 05-34, the *epsF<sub>N</sub>* gene was also amplified by PCR using primers F-epsF and R-epsFh<sub>6</sub> (5'-CGCGAGCTCTAATGATGATGATGATGTGAGCTTGTAGGACTATTC-3'), thereby introducing a 6× His tag coding sequence to the 3'-end of *epsF<sub>N</sub>* for Western blot analysis. This amplicon was cloned into pSlpA-8148 as described above, and then transformed into *S. thermophilus* 05-34, yielding the recombinant strain *S. thermophilus* 05Fh<sub>6</sub>. Meanwhile, the plasmid pSlpA-8148 was also introduced into *S. thermophilus* 05-34, resulting in the control strain *S. thermophilus* 05CK. Overnight cultures of *S. thermophilus* 05CK and *S. thermophilus* 05Fh<sub>6</sub> were inoculated into 20 mL of fresh MRS medium containing 5 μg/mL of Cm. Intracellular proteins were extracted as previously described (*Wang et al., 2015*). Aliquots of 50 μg proteins from *S. thermophilus* 05CK and *S. thermophilus* 05Fh<sub>6</sub> were separated by a 15.5% Tricine-SDS-PAGE gel, and then transferred to a nitrocellulose membrane using Bio-Rad Mini Trans-Blot<sup>®</sup> system. The membrane was blocked with TBS + 5% (w/v) skim milk overnight at 4 °C, and then incubated with a 1:5,000 dilution of anti His-Tag mouse monoclonal antibody (CWBiotech, Beijing, China) followed by a 1:5,000 dilution of HRP-conjugated goat Anti-Mouse IgG (CWBiotech, Beijing, China). Protein bands were visualized using a cECL Western Blot Kit (CWBiotech, Beijing, China) according to the manufacturer's instructions.

### RNA isolation and reverse transcription qPCR (RT-qPCR)

*S. thermophilus* 05epsF and *S. thermophilus* 05CK were grown in 10% RSM medium (initial pH 7.0) at 37 °C for 30 h. Each 5 mL of fermentation culture was mixed with 5 mL

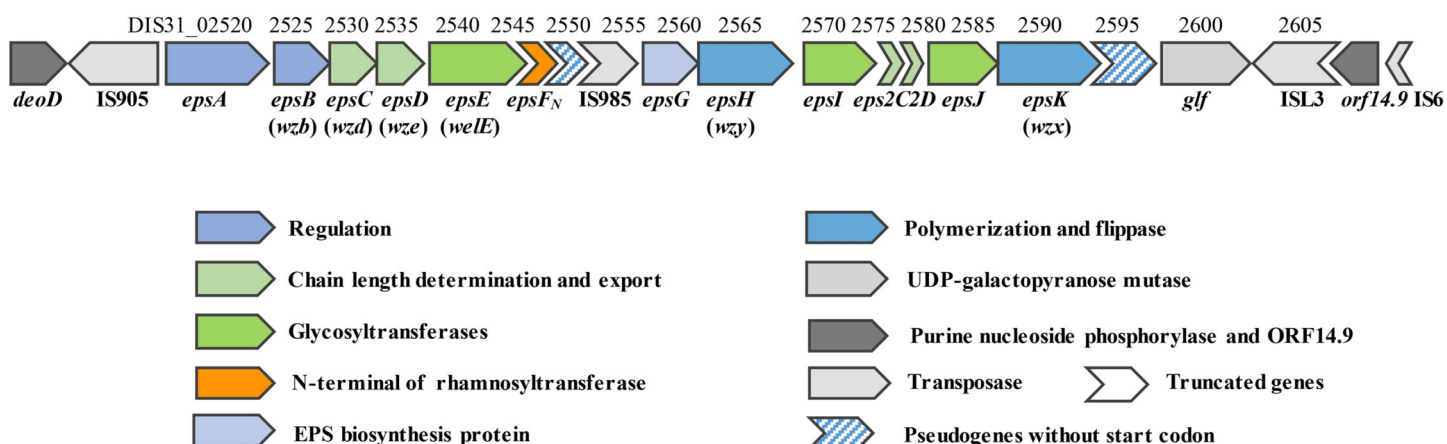
2% (w/v) sodium citrate, 10 mg Pronase E and 20  $\mu$ l  $\beta$ -mercaptoethanol, and incubated in water bath at 50 °C for 3 h. The mixture was centrifuged at 6,000g for 5 min at 4 °C and the supernatant was discarded. The total RNA was isolated using RNAPrep pure cell/bacteria kit following the manufacture's instruction (Tiangen, Beijing, China). Subsequently, reverse transcription was carried out with PrimeScript II 1st strand cDNA synthesis Kit (Takara, Beijing, China), with 750 ng of total RNA as the template. Specific primers were designed using PRIMER V5 software to amplify the 100- to 150-bp regions of interest genes including *epsA*, *epsC*, *eps2C* and *epsG* (Table S1), and their specificity was checked before quantitative analyses. Semi-quantitative PCR was performed using a 1:20 dilution of cDNA as template, and the PCR products were visualized on 2% agarose gels stained with ethidium bromide. Real-Time PCR (qPCR) was performed using SYBR Green assay kit (Tiangen, Beijing, China) in a Roche LightCycler<sup>®</sup> 96 Real-Time PCR System (Roche Applied Science, Rotkreuz, Switzerland). Gene expressions were calculated by the  $2^{-\Delta\Delta CT}$  method (Schmittgen & Livak, 2008) by using the 16S rRNA as the reference gene.

### Isolation and purification of EPS

*S. thermophilus* 05epsF and *S. thermophilus* 05CK were propagated in MRS broth at 37 °C for 24 h, followed by two more precultures inoculated in MRS broth. For EPS isolation, cell cultures were inoculated (2%, v/v) into 10% RSM broth (initial pH 7.0) and grown for 30 h at 37 °C. The extraction of EPS was performed as previously described with some modifications (Cerning *et al.*, 1994). Briefly, cultures were heated at 100 °C for 15 min and centrifuged at 6,000g for 15 min to remove the cells. In order to precipitate proteins, trichloroacetic acid (TCA) was added in the supernatant to a final concentration of 4% (v/v), stored at 4 °C for 2 h, and then centrifuged at 6,000g for 15 min. Crude EPS was precipitated by adding cold ethanol to the supernatant at a ratio of 3:1 (v/v), mixed thoroughly and kept for 24 h at 4 °C. The pellet containing EPS was obtained by centrifugation, dialyzed and freeze-dried. The crude EPS was purified as previously described (Li *et al.*, 2016). The yield of EPS was quantified using the phenol-sulfuric method using glucose as a standard (DuBois *et al.*, 1956). Concentration of protein and nucleic acid of purified fraction was determined spectroscopically at 280 and 260 nm, respectively.

### Molecular mass and monosaccharide composition of EPS

In order to determine the average molecular mass of the EPS sample, an Agilent1100 series HPLC system (Agilent, Santa Clara, CA, USA) was conducted with an RI detector. EPS samples were eluted with deionized water at a flow rate of 0.8 mL/min. Based on the linear regression equation drawn through dextran standards (Sigma–Aldrich, St. Louis, MO, USA), the molecular mass of EPS was analyzed through GPC data processing software. The monosaccharide composition was determined by gas chromatography coupled with mass spectrograph (GC-MS). Some pretreatments were performed as previously described (Li *et al.*, 2016). The treated samples were used for GC-MS with the conditions as follows: initial column temperature was set at 140 °C with a rate of 1.5 °C/min to reach 200 °C and with a rate of 10 °C/min to 250 °C, then the highest temperature was held for 5 min, and samples were injected into the column with N<sub>2</sub> as the carrier gas at a flow



**Figure 1** Genetic organization of the *eps* cluster and flanking regions in *S. thermophilus* 05-34. Gene\_locus tags and gene names are marked above and below the genes, respectively. Gene names in brackets are adopted from the bacterial polysaccharide gene nomenclature (BPGN) system (Reeves et al., 1996). Full-size [DOI: 10.7717/peerj.8524/fig-1](https://doi.org/10.7717/peerj.8524/fig-1)

rate of one mL/min. The monosaccharide composition of EPS was determined by comparison with the retention time of monosaccharide standards.

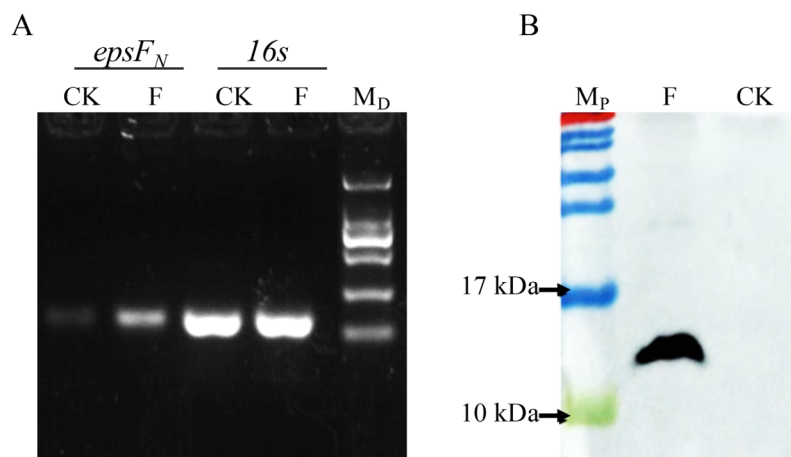
### Statistical analysis

All the experiments were independently conducted in triplicates. Results were presented as the mean value  $\pm$  standard error. When two groups were compared, an unpaired student *t* test with Welch's correction was used to calculate the *p* values.

## RESULTS

### Genetic organization of the *eps* cluster in *S. thermophilus* 05-34

The draft genome of 05-34 has been sequenced by our team recently (accession number: QFLC00000000). A 16.3 kb region encoding enzymes required for EPS biosynthesis was identified on the chromosome (Fig. 1). Based on amino acid sequence and protein structural similarity, putative functions were assigned to 18 genes (DIS31\_02520~DIS31\_02605) within this region (Table S2). Among them, four genes (DIS31\_02520~DIS31\_02535) at the 5'-end are highly conserved in *S. thermophilus* species, and were assigned for regulation (*epsAB*) and chain-length control (*epsCD*) functions in EPS biosynthesis. Two truncated genes *eps2C* and *eps2D* (DIS31\_02575 and DIS31\_02580) encode proteins that show 94% identity to the C-terminus of Lp\_1197 and N-terminus of Lp\_1198, respectively, which were annotated as EPS chain-length controlling proteins in *Lactobacillus plantarum* WCFS1. Genes *epsH* (DIS31\_02565) and *epsK* (DIS31\_02590) encode polymerase Wzy and flippase Wzx of the typical Wzx/Wzy-dependent pathway for EPS biosynthesis, respectively (Table S2). In addition, a pseudogene DIS31\_02595 was also assigned as flippase, however, no translational start codon was detected at its 5'-end. Thus, DIS31\_02595 was considered to be an inactivated gene. The central region of the *eps* cluster encodes an EPS biosynthesis protein EpsG (DIS31\_02560) with unknown function, and several glycosyltransferases. Among them, EpsE (DIS31\_02540) was 100% identical to the priming glycosyltransferase of *S. thermophilus* NCFB 2393, which was shown to transfer glucosyl



**Figure 2** Detection of the over-expression of *epsF<sub>N</sub>* in *S. thermophilus* by Semi-quantitative PCR (A) and Western blot (B). (A) Agarose electrophoresis of PCR products of *epsF<sub>N</sub>* and 16s rRNA, CK: *S. thermophilus* 05CK; F: *S. thermophilus* 05epsF; M<sub>D</sub>: DNA maker DL2000 (Tiangen). (B) Western blot analysis, CK: *S. thermophilus* 05CK; F: *S. thermophilus* 05Fh<sub>6</sub>; M<sub>P</sub>: PageRuler Prestained Protein Ladder (Thermo Scientific, Waltham, MA, USA). [Full-size](#) DOI: [10.7717/peerj.8524/fig-2](https://doi.org/10.7717/peerj.8524/fig-2)

1-phosphate to the undecaprenyl phosphate (*Almirón-Roig et al., 2000*). In addition, proteins EpsI (DIS31\_02570) and EpsJ (DIS31\_02585) showed 100% identity to  $\beta$ -1,3-glucosyltransferase (QBR99840.1) and 98% identity to GT-2 family glycosyltransferase (QBR99843.1) from *S. thermophilus*, respectively. Notably, the protein encoded by gene *epsF<sub>N</sub>* (DIS31\_02545) was composed of only 87 amino acids, which displayed 99% identity to the N-terminal section of rhamnosyltransferase EpsF with 390 amino acids in *S. thermophilus* EU20 and *S. thermophilus* NCFB2393. Homologous overexpression of *epsF<sub>N</sub>* was carried out to explore its role in EPS biosynthesis in 05-34.

### EpsF<sub>N</sub> was successfully overexpressed in *S. thermophilus* 05-34

The *epsF<sub>N</sub>* gene was amplified by PCR and cloned into plasmid pSlpA-8148. DNA sequencing verified that the length of the amplicon was 264 bp, which showed 100% identity with the *epsF<sub>N</sub>* gene of *S. thermophilus* 05-34. Semi-quantitative PCR revealed the elevated transcription level of *epsF<sub>N</sub>* under the control of promoter pSlpA in strain *S. thermophilus* 05epsF compared to the control strain *S. thermophilus* 05CK (*Fig. 2A*). Then RT-qPCR assay showed that the mRNA level of *epsF<sub>N</sub>* was 370-fold higher in strain *S. thermophilus* 05epsF than that in *S. thermophilus* 05CK. Western blotting assay further revealed the production of a 12-kDa protein in *S. thermophilus* 05Fh<sub>6</sub>, which corresponded to the expected size of EpsF<sub>N</sub>-His<sub>6</sub>, while no corresponding bands were observed in the control strain *S. thermophilus* 05CK (*Fig. 2B*). These results indicated that the gene *epsF<sub>N</sub>* was successfully over-expressed in *S. thermophilus* 05-34.

### The overexpression of EpsF<sub>N</sub> did not affect the yield of EPS in *S. thermophilus*

*S. thermophilus* 05epsF was cultivated in 10% RSM broth (initial pH 7.0) at 37 °C for 30 h, and then EPS from RSM samples was isolated and purified. According to the elution

profile of EPS from *S. thermophilus* 05epsF, there was only one single and relatively symmetrical peak at OD<sub>490</sub> in fraction No. 11, while the peak of the elution profile in the control group appeared in fraction No. 10. These results suggested that overexpression of *epsF<sub>N</sub>* in *S. thermophilus* 05-34 decreased the average molecular mass of EPS. The purified EPS fraction showed no absorption at 260 nm or 280 nm by ultraviolet detection, suggesting that there was no nucleic acids or protein contamination in EPS samples. Based on the glucose standard, the yield of EPS produced by *S. thermophilus* 05epsF was  $124 \pm 4$  mg/L, showing no significant difference with that of the control group ( $131 \pm 2$  mg/L,  $p = 0.18$ ).

### ***S. thermophilus* 05epsF produced EPS with unchanged composition but reduced molecular mass**

The average molecular mass of EPS produced by *S. thermophilus* 05epsF was determined using GPC software to be  $8.8 \times 10^4$  Da (Fig. 3A), while the molecular mass of EPS produced by *S. thermophilus* 05CK was  $4.6 \times 10^5$  Da (Fig. 3B). Based on the retention time of different monosaccharide standards, the monomer analysis by GC-MS indicated that the EPS produced by *S. thermophilus* 05epsF and *S. thermophilus* 05CK were both composed of galactose and glucose in an approximate ratio of 1.0:0.8 with a trace amount of mannose, which was probably derived from cell wall polysaccharides or MRS medium composition (Figs. 4A and 4B). Characteristics of EPS produced by *S. thermophilus* 05CK and 05epsF were summarized in Table 2.

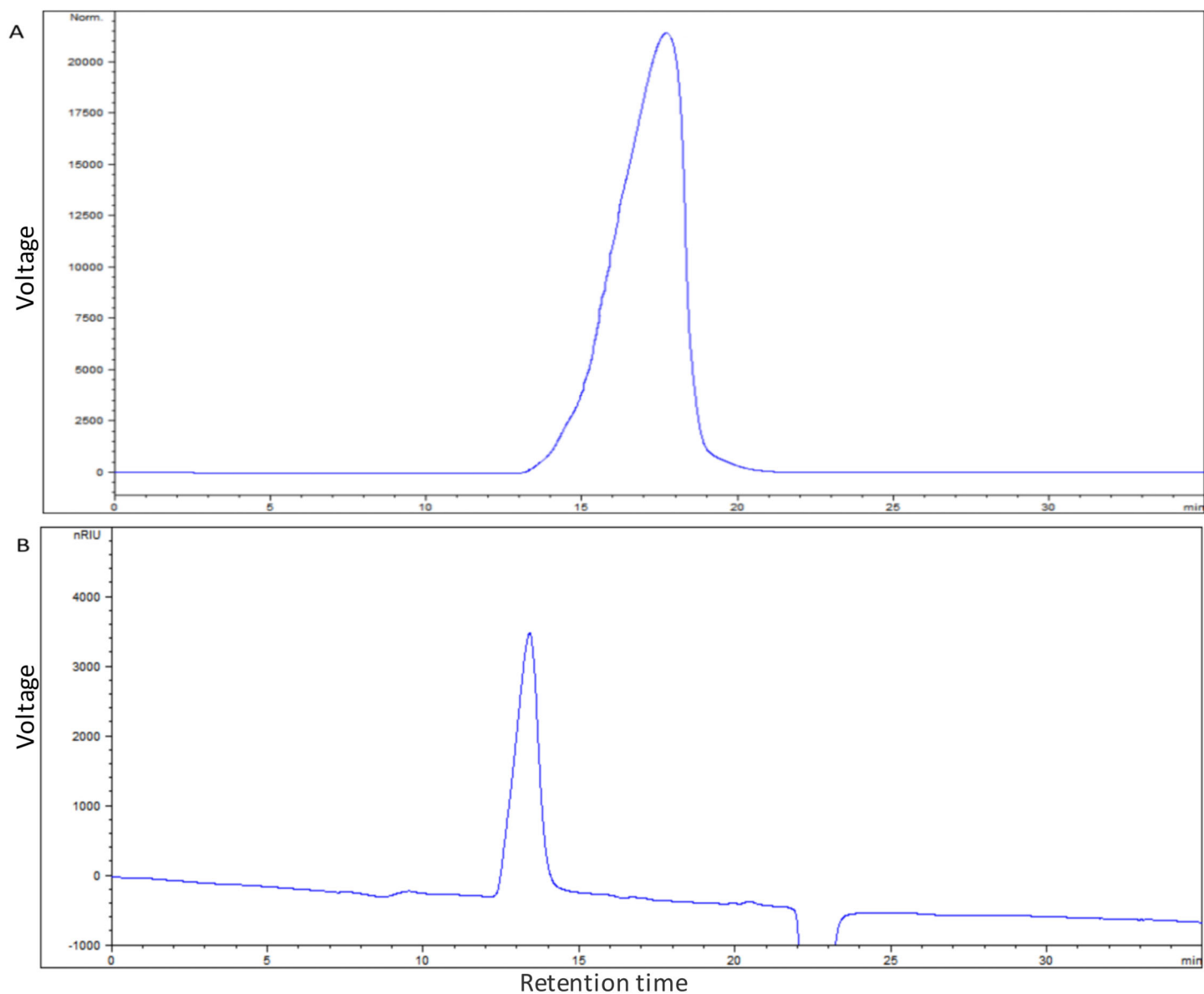
### **The overexpression of *EpsF<sub>N</sub>* caused a down-regulation of *eps* genes in *S. thermophilus***

Since the molecular mass of EPS produced by *S. thermophilus* 05epsF decreased 5-fold than that of the control strain, we analyzed the transcription level of chain-length determining genes *epsC* and *eps2C*, regulatory gene *epsA*, and gene *epsG* encoding the EPS biosynthesis protein. RT-qPCR results revealed that the transcription levels of these genes in strain *S. thermophilus* 05epsF were 5- to 6-fold lower than that of the control strain (Fig. 5), indicating that the over-expression of *EpsF<sub>N</sub>* led to a down-regulation of genes within the *eps* cluster in *S. thermophilus* 05-34. Therefore, the reduced molecular mass of EPS may result from the decreased level of these genes, especially the down-regulation of *epsC* in *S. thermophilus* 05epsF.

## **DISCUSSION**

Nucleotide sugar-dependent glycosyltransferases play important roles in the assembly of repeating units of EPS in bacteria. In this study, the functional role the 87-amino acid fragment of rhamnosyltransferase EpsF in EPS biosynthesis was investigated. Notably, the overexpression of this N-terminal region (*EpsF<sub>N</sub>*) did not result in the incorporation of rhamnose residues into the repeating units of EPS produced by 05-34, despite this strain harbors the four enzymes RmlABCD (gene locus DIS31\_09015 ~ 09025, DIS31\_00470) required for the biosynthesis of dTDP-Rhamnose, the substrate of EpsF

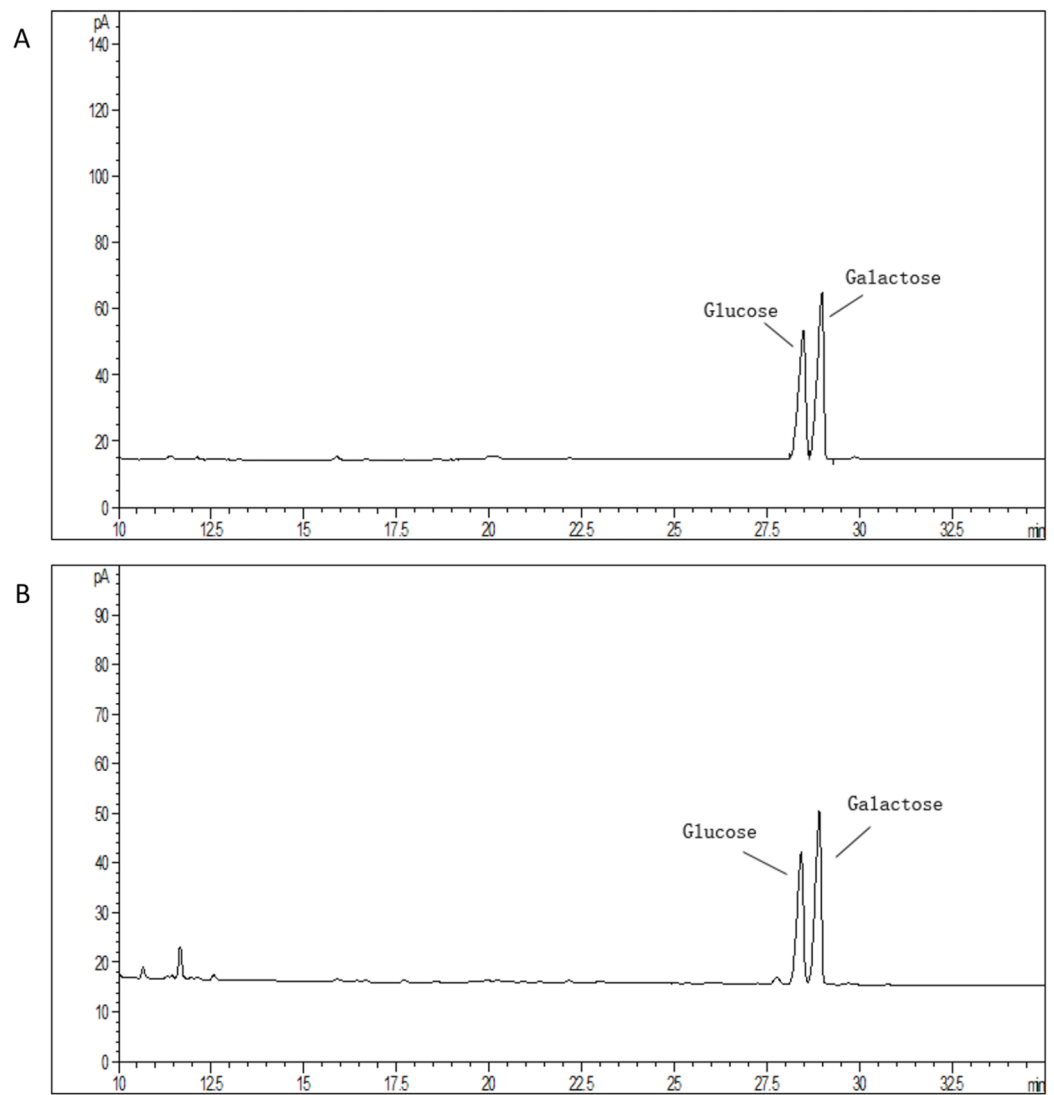




**Figure 3** Molecular mass of EPS produced by *S. thermophilus* 05epsF (A) and *S. thermophilus* 05CK (B). According to the linear regression equation drawn through dextran standards, the molecular mass of EPS samples was determined based on their retention time (X-axis) using gel permeation chromatography (GPC). [Full-size !\[\]\(fcc3264021d438d9732560e78099f674\_img.jpg\) DOI: 10.7717/peerj.8524/fig-3](https://doi.org/10.7717/peerj.8524/fig-3)

(Zeidan *et al.*, 2017). Thus, it is proposed that EpsF<sub>N</sub> lost the catalytic function as a rhamnosyltransferase due to rearrangements leading to loss of C-terminal domain.

Protein sequence analysis of the full length version EpsF (GenBank Accession No: [ABD96549.1](https://www.ncbi.nlm.nih.gov/nuclot/ABD96549.1), EpsF from *S. thermophilus* EU20) showed that it belongs to the glycosyltransferase family 4 (GT-4) and displays a GT-B type fold. Secondary structure analysis showed that EpsF<sub>N</sub> contains two  $\alpha$ -helices and three  $\beta$ -strands, which form the conserved N-terminal segment of Rossmann folds that responsible for nucleotides binding (Rao & Rossmann, 1973). Further sequence alignment analysis of EpsF<sub>N</sub> with homologous rhamnosyltransferases from other *Streptococcus* species revealed the presence of an



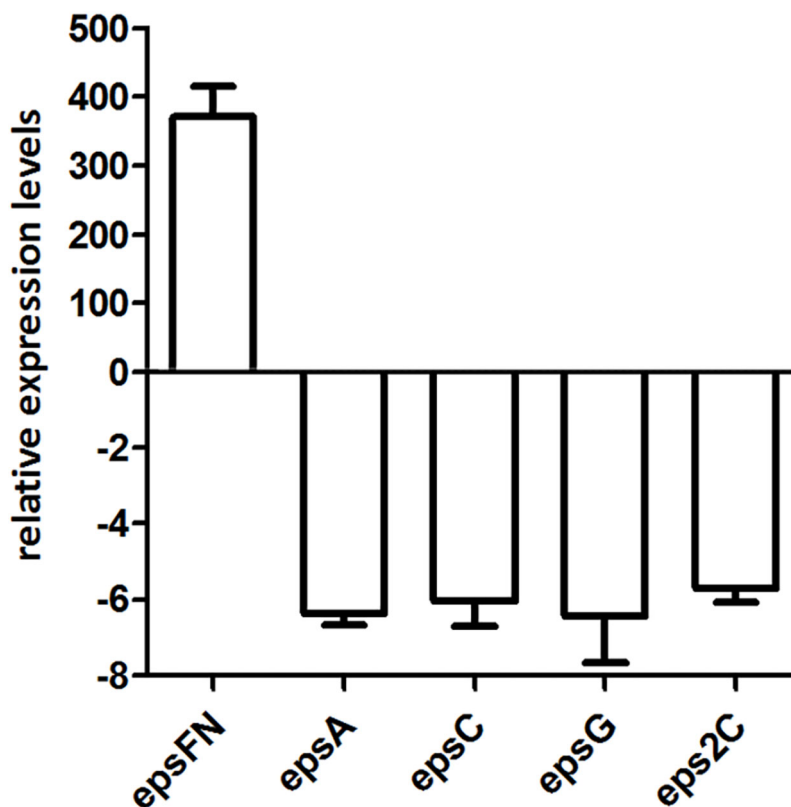
**Figure 4** Monosaccharide composition profile of hydrolyzed EPS from *S. thermophilus* 05epsF (A) and *S. thermophilus* 05CK (B). [Full-size !\[\]\(1663bb69f307a960345edb0e712f8c02\_img.jpg\) DOI: 10.7717/peerj.8524/fig-4](https://doi.org/10.7717/peerj.8524/fig-4)

EX<sub>7</sub>E motif (residues 21–29) and a DxD motif (residues 33–35) (Fig. 6), which were proposed to be involved in the stabilizing of donor substrates and the interaction with phosphate groups of nucleotides, respectively (Coutinho *et al.*, 2003; Breton *et al.*, 2006). Taken together, it is speculated that EpsF<sub>N</sub> may still have the ability to interact with nucleotide sugar molecules, and thus to affect the EPS biosynthesis in 05-34.

In the EpsF<sub>N</sub> overexpression strain, the molecular mass of EPS decreased from  $4.6 \times 10^5$  Da to  $8.8 \times 10^4$  Da, and the transcription level of genes within the *eps* cluster including *epsC*, *eps2C*, *epsA* and *epsG* also showed a remarkable down-regulation. A previous study by our team revealed that when the molecular mass of EPS increased from  $2.5 \times 10^4$  Da to  $4.7 \times 10^5$  Da in 05-34, the corresponding transcription level of *epsC* showed a 2.7-fold up-regulation (Li *et al.*, 2016). Protein structural homologs analysis revealed a high similarity (99.9%; *E*-value =  $3.90\text{E}-25$ ) of EpsC to the chain-length regulation protein

**Table 2** EPS characteristics produced by *S. thermophilus* 05CK and 05epsF.

	<i>S. thermophilus</i> 05CK	<i>S. thermophilus</i> 05epsF
Yield (mg/L)	131 ± 2	124 ± 4
Monosaccharide composition	galactose and glucose	galactose and glucose
Molar ratio (Gal/Glc)	1.25	1.25
Molecular mass (Da)	4.6 × 10 <sup>5</sup>	8.8 × 10 <sup>4</sup>
Elution time (fraction No.)	10	11

**Figure 5** RT-qPCR analysis of the expression levels of *epsFN*, *epsA*, *epsC*, *epsG* and *eps2C* in *S. thermophilus* 05epsF. The fold changes calculated are relative to the transcript levels in *S. thermophilus* 05epsF compared to those in *S. thermophilus* 05CK. Error bars represent the standard errors of the results obtained by three independent experiments ( $p < 0.05$ ).

Full-size DOI: 10.7717/peerj.8524/fig-5

WzzE of *Escherichia coli* (Table S2). The impact of the expression level of WzzE homologs on lipopolysaccharide O-antigen chain length distribution has been reported in *Shigella* (Carter *et al.*, 2007) and *Pseudomonas* (Huszczynski *et al.*, 2019). In addition, Bender, Cartee & Yother (2003) showed that inactivation of *cps2C* or *cps2D* led to the production of only short-chain polymers in *S. pneumoniae*. Therefore, it is assumed that there is a correlation between the transcription level of chain-length determining genes and the molecular mass of EPS in *S. thermophilus* 05-34. Furthermore, the concentration of nucleotide sugars has also been shown to modulate the chain length of CPS in



shorter-chain length polysaccharide was released, the assembly of a new chain would begin immediately. Thus, we speculated that *S. thermophilus* 05epsF may produce increased number of shorter-chain length EPS, resulting in a total yield similar to that of the control strain.

## CONCLUSIONS

The current study describes the effect of EpsF<sub>N</sub>, an 87-amino acid N-terminal domain of EpsF, on EPS biosynthesis. Our results showed that overexpression of this N-terminal domain did not change the yield or monosaccharide composition of EPS produced by *S. thermophilus* 05-34, whereas the molecular mass decreased from  $4.6 \times 10^5$  Da to  $8.8 \times 10^4$  Da. Accordingly, the transcription level of genes within the *eps* cluster showed a remarkable down-regulation in the EpsF<sub>N</sub> overexpression strain, suggesting that EpsF<sub>N</sub> has an impact on the transcription level of *eps* cluster, potentially via nonspecific interactions with sugar precursors and lowering their intracellular concentration.

## ACKNOWLEDGEMENTS

We thank Professor Willem M. de Vos (Wageningen University) for the gift of plasmid pNZ8148.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This work was supported by the National Natural Science Foundation of China (No. 21476250). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:  
National Natural Science Foundation of China: 21476250.

### Competing Interests

The authors declare that they have no competing interests.

### Author Contributions

- Guohong Wang conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Jiayi Li performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Shuxin Xie performed the experiments, prepared figures and/or tables, and approved the final draft.
- Zhengyuan Zhai analyzed the data, prepared figures and/or tables, and approved the final draft.
- Yanling Hao conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

## Data Availability

The following information was supplied regarding data availability:

The draft genome sequence of *Streptococcus thermophilus* 05-34 is available at GenBank: [QFLC00000000](https://www.ncbi.nlm.nih.gov/nuclseq/05-34).

## Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.8524#supplemental-information>.

## REFERENCES

- Albesa-Jové D, Giganti D, Jackson M, Alzari PM, Guerin ME. 2014. Structure–function relationships of membrane-associated GT-B glycosyltransferases. *Glycobiology* **24**(2):108–124 DOI [10.1093/glycob/cwt101](https://doi.org/10.1093/glycob/cwt101).
- Almirón-Roig E, Gasson MJ, Mulholland F, Griffin AM. 2000. The complete *cps* gene cluster from *Streptococcus thermophilus* NCFB, 2393 involved in the biosynthesis of a new exopolysaccharide. *Microbiology* **146**(11):2793–2802 DOI [10.1099/00221287-146-11-2793](https://doi.org/10.1099/00221287-146-11-2793).
- Bastida A, Fernández-Mayoralas A, García-Junceda E. 2002. C-Terminal truncation of  $\alpha$  1,6-fucosyltransferase from *Rhizobium* sp. does not annul the transferase activity of the enzyme. *Bioorganic & Medicinal Chemistry* **10**(3):737–742 DOI [10.1016/S0968-0896\(01\)00327-3](https://doi.org/10.1016/S0968-0896(01)00327-3).
- Bender MH, Cartee RT, Yother J. 2003. Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in *Streptococcus pneumoniae*. *Journal of Bacteriology* **185**(20):6057–6066 DOI [10.1128/JB.185.20.6057-6066.2003](https://doi.org/10.1128/JB.185.20.6057-6066.2003).
- Bentley SD, Aanensen DM, Mavroidi A, Saunders D, Rabinowitsch E, Collins M, Donohoe K, Harris D, Murphy L, Quail MA, Samuel G, Skovsted IC, Kalltoft MS, Barrell B, Reeves PR, Parkhill J, Spratt BG. 2006. Genetic analysis of the capsular biosynthetic locus from all 90 Pneumococcal serotypes. *PLOS Genetics* **2**(3):e31 DOI [10.1371/journal.pgen.0020031](https://doi.org/10.1371/journal.pgen.0020031).
- Boot HJ, Kolen CP, Andreadaki FJ, Leer RJ, Pouwels PH. 1996. The *Lactobacillus acidophilus* S-layer protein gene expression site comprises two consensus promoter sequences, one of which directs transcription of stable mRNA. *Journal of Bacteriology* **178**(18):5388–5394 DOI [10.1128/JB.178.18.5388-5394.1996](https://doi.org/10.1128/JB.178.18.5388-5394.1996).
- Bouazzaoui K, LaPointe G. 2006. Use of antisense RNA to modulate glycosyltransferase gene expression and exopolysaccharide molecular mass in *Lactobacillus rhamnosus*. *Journal of Microbiological Methods* **65**(2):216–225 DOI [10.1016/j.mimet.2005.07.011](https://doi.org/10.1016/j.mimet.2005.07.011).
- Breton C, Šnajdrová L, Jeanneau C, Koča J, Imberty A. 2006. Structures and mechanisms of glycosyltransferases. *Glycobiology* **16**(2):29R–37R DOI [10.1093/glycob/cwj016](https://doi.org/10.1093/glycob/cwj016).
- Broadbent JR, McMahan DJ, Welker DL, Oberg CJ, Moineau S. 2003. Biochemistry, genetics, and applications of exopolysaccharide production in *Streptococcus thermophilus*: a review. *Journal of Dairy Science* **86**(2):407–423 DOI [10.3168/jds.S0022-0302\(03\)73619-4](https://doi.org/10.3168/jds.S0022-0302(03)73619-4).
- Caggianiello G, Kleerebezem M, Spano G. 2016. Exopolysaccharides produced by lactic acid bacteria: from health-promoting benefits to stress tolerance mechanisms. *Applied Microbiology and Biotechnology* **100**(9):3877–3886 DOI [10.1007/s00253-016-7471-2](https://doi.org/10.1007/s00253-016-7471-2).
- Campbell JA, Davies GJ, Bulone V, Henrissat B. 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochemical Journal* **326**(3):929–939 DOI [10.1042/bj3260929u](https://doi.org/10.1042/bj3260929u).
- Carter JA, Blondel CJ, Zaldivar M, Alvarez SA, Marolda CL, Valvano MA, Contreras I. 2007. O-antigen modal chain length in *Shigella flexneri* 2a is growth-regulated through

- RfaH-mediated transcriptional control of the *wzy* gene. *Microbiology* **153**(10):3499–3507  
DOI [10.1099/mic.0.2007/010066-0](https://doi.org/10.1099/mic.0.2007/010066-0).
- Cerning J, Renard CMGC, Thibault JF, Bouillanne C, Landon M, Desmazeaud MJ, Topisirovic L. 1994.** Carbon source requirements for exopolysaccharide production by *Lactobacillus casei* CG11 and partial structure analysis of the polymer. *Applied and Environmental Microbiology* **60**(11):3914–3919 DOI [10.1128/AEM.60.11.3914-3919.1994](https://doi.org/10.1128/AEM.60.11.3914-3919.1994).
- Coutinho PM, Deleury E, Davies GJ, Henrissat B. 2003.** An evolving hierarchical family classification for glycosyltransferases. *Journal of Molecular Biology* **328**(2):307–317  
DOI [10.1016/S0022-2836\(03\)00307-3](https://doi.org/10.1016/S0022-2836(03)00307-3).
- De Vuyst L, Degeest B. 1999.** Heteropolysaccharides from lactic acid bacteria. *FEMS Microbiology Reviews* **23**(2):153–177 DOI [10.1111/j.1574-6976.1999.tb00395.x](https://doi.org/10.1111/j.1574-6976.1999.tb00395.x).
- Del Moral S, Olvera C, Rodriguez M, Munguia A. 2008.** Functional role of the additional domains in inulosucrase (IsIA) from *Leuconostoc citreum* CW28. *BMC Biochemistry* **9**(1):6  
DOI [10.1186/1471-2091-9-6](https://doi.org/10.1186/1471-2091-9-6).
- DuBois M, Gilles KA, Hamilton JK, Rebers PA, Smith F. 1956.** Colorimetric method for determination of sugars and related substances. *Analytical Chemistry* **28**(3):350–356  
DOI [10.1021/ac60111a017](https://doi.org/10.1021/ac60111a017).
- Forsee WT, Cartee RT, Yother J. 2009.** A kinetic model for chain length modulation of *Streptococcus pneumoniae* celluburonan capsular polysaccharide by nucleotide sugar donor concentrations. *Journal of Biological Chemistry* **284**(18):11836–11844  
DOI [10.1074/jbc.M900379200](https://doi.org/10.1074/jbc.M900379200).
- Grangeasse C. 2016.** Rewiring the pneumococcal cell cycle with serine/threonine- and tyrosine-kinases. *Trends in Microbiology* **24**(9):713–724 DOI [10.1016/j.tim.2016.04.004](https://doi.org/10.1016/j.tim.2016.04.004).
- 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**(12):3119–3123 DOI [10.1128/AEM.55.12.3119-3123.1989](https://doi.org/10.1128/AEM.55.12.3119-3123.1989).
- Huszczynski SM, Coumoundouros C, Pham P, Lam JS, Khursigara CM. 2019.** Unique regions of the polysaccharide copolymerase Wzz<sub>2</sub> from *Pseudomonas aeruginosa* are essential for O-specific antigen chain length control. *Journal of Bacteriology* **201**(15):e00165-19  
DOI [10.1128/JB.00165-19](https://doi.org/10.1128/JB.00165-19).
- Islam ST, Lam JS. 2014.** Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. *Canadian Journal of Microbiology* **60**(11):697–716 DOI [10.1139/cjm-2014-0595](https://doi.org/10.1139/cjm-2014-0595).
- Kleerebezem M, Van Kranenburg R, Tuinier R, Boels IC, Zoon P, Looijesteijn E, Hugenholtz J, De Vos WM. 1999.** Exopolysaccharides produced by *Lactococcus lactis*: from genetic engineering to improved rheological properties? In: Konings WN, Kuipers OP, Veld JHJH, eds. *Lactic Acid Bacteria: Genetics, Metabolism and Applications*. Dordrecht: Springer, 357–365.
- Kuipers OP, De Ruyter PGGA, Kleerebezem M, De Vos WM. 1998.** Quorum sensing-controlled gene expression in lactic acid bacteria. *Journal of Biotechnology* **64**(1):15–21  
DOI [10.1016/S0168-1656\(98\)00100-X](https://doi.org/10.1016/S0168-1656(98)00100-X).
- Lairson LL, Henrissat B, Davies GJ, Withers SG. 2008.** Glycosyltransferases: structures, functions, and mechanisms. *Annual Review of Biochemistry* **77**(1):521–555  
DOI [10.1146/annurev.biochem.76.061005.092322](https://doi.org/10.1146/annurev.biochem.76.061005.092322).
- Lebeer S, Verhoeven TLA, Francius G, Schoofs G, Lambrichts I, Dufrene Y, Vanderleyden J, De Keersmaecker SCJ. 2009.** Identification of a gene cluster for the biosynthesis of a long, galactose-rich exopolysaccharide in *Lactobacillus rhamnosus* GG and functional analysis of the priming glycosyltransferase. *Applied and Environmental Microbiology* **75**(11):3554–3563  
DOI [10.1128/AEM.02919-08](https://doi.org/10.1128/AEM.02919-08).

- Lesk AM. 1995.** NAD-binding domains of dehydrogenases. *Current Opinion in Structural Biology* **5(6)**:775–783 DOI [10.1016/0959-440X\(95\)80010-7](https://doi.org/10.1016/0959-440X(95)80010-7).
- Li B, Ding X, Evivie SE, Jin D, Meng Y, Huo G, Liu F. 2018.** Short communication: genomic and phenotypic analyses of exopolysaccharides produced by *Streptococcus thermophilus* KLDS SM. *Journal of Dairy Science* **101(1)**:106–112 DOI [10.3168/jds.2017-13534](https://doi.org/10.3168/jds.2017-13534).
- Li D, Li J, Zhao F, Wang G, Qin Q, Hao Y. 2016.** The influence of fermentation condition on production and molecular mass of EPS produced by *Streptococcus thermophilus* 05-34 in milk-based medium. *Food Chemistry* **197**:367–372 DOI [10.1016/j.foodchem.2015.10.129](https://doi.org/10.1016/j.foodchem.2015.10.129).
- Marciset O, Mollet B. 1994.** Multifactorial experimental design for optimizing transformation: electroporation of *Streptococcus thermophilus*. *Biotechnology and Bioengineering* **43(6)**:490–496 DOI [10.1002/bit.260430609](https://doi.org/10.1002/bit.260430609).
- Martinez-Fleites C, Proctor M, Roberts S, Bolam DN, Gilbert HJ, Davies GJ. 2006.** Insights into the synthesis of lipopolysaccharide and antibiotics through the structures of two retaining glycosyltransferases from family GT4. *Chemistry & Biology* **13(11)**:1143–1152 DOI [10.1016/j.chembiol.2006.09.005](https://doi.org/10.1016/j.chembiol.2006.09.005).
- Meng X, Dobruchowska JM, Pijning T, Gerwig GJ, Kamerling JP, Dijkhuizen L. 2015.** Truncation of domain V of the multidomain glucansucrase GTF180 of *Lactobacillus reuteri* 180 heavily impairs its polysaccharide-synthesizing ability. *Applied Microbiology and Biotechnology* **99(14)**:5885–5894 DOI [10.1007/s00253-014-6361-8](https://doi.org/10.1007/s00253-014-6361-8).
- Mierau I, Kleerebezem M. 2005.** 10 years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*. *Applied Microbiology and Biotechnology* **68(6)**:705–717 DOI [10.1007/s00253-005-0107-6](https://doi.org/10.1007/s00253-005-0107-6).
- Nwodo U, Green E, Okoh A. 2012.** Bacterial exopolysaccharides: functionality and prospects. *International Journal of Molecular Sciences* **13(12)**:14002–14015 DOI [10.3390/ijms131114002](https://doi.org/10.3390/ijms131114002).
- Purwandari U, Shah NP, Vasiljevic T. 2007.** Effects of exopolysaccharide-producing strains of *Streptococcus thermophilus* on technological and rheological properties of set-type yoghurt. *International Dairy Journal* **17(11)**:1344–1352 DOI [10.1016/j.idairyj.2007.01.018](https://doi.org/10.1016/j.idairyj.2007.01.018).
- Qasba PK, Ramakrishnan B, Boeggeman E. 2005.** Substrate-induced conformational changes in glycosyltransferases. *Trends in Biochemical Sciences* **30(1)**:53–62 DOI [10.1016/j.tibs.2004.11.005](https://doi.org/10.1016/j.tibs.2004.11.005).
- Qin QQ, Xia BS, Xiong Y, Zhang SX, Luo YB, Hao YL. 2011.** Structural characterization of the exopolysaccharide produced by *Streptococcus thermophilus* 05-34 and its in situ application in yogurt. *Journal of Food Science* **76(9)**:C1226–C1230 DOI [10.1111/j.1750-3841.2011.02397.x](https://doi.org/10.1111/j.1750-3841.2011.02397.x).
- Rao ST, Rossmann MG. 1973.** Comparison of super-secondary structures in proteins. *Journal of Molecular Biology* **76(2)**:241–256 DOI [10.1016/0022-2836\(73\)90388-4](https://doi.org/10.1016/0022-2836(73)90388-4).
- Reeves PR, Hobbs M, Valvano MA, Skurnik M, Whitfield C, Coplin D, Kido N, Klena J, Maskell D, Raetz CRH, Rick PD. 1996.** Bacterial polysaccharide synthesis and gene nomenclature. *Trends in Microbiology* **4(12)**:495–503 DOI [10.1016/S0966-842X\(97\)82912-5](https://doi.org/10.1016/S0966-842X(97)82912-5).
- Remus DM, Van Kranenburg R, Van Swam II, Taverne N, Bongers RS, Wels M, Wells JM, Bron PA, Kleerebezem M. 2012.** Impact of 4 *Lactobacillus plantarum* capsular polysaccharide clusters on surface glycan composition and host cell signaling. *Microbial Cell Factories* **11(1)**:149 DOI [10.1186/1475-2859-11-149](https://doi.org/10.1186/1475-2859-11-149).
- Robert X, Gouet P. 2014.** Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* **42(W1)**:W320–W324 DOI [10.1093/nar/gku316](https://doi.org/10.1093/nar/gku316).
- Schmittgen TD, Livak KJ. 2008.** Analyzing real-time PCR data by the comparative C<sub>T</sub> method. *Nature Protocols* **3(6)**:1101–1108 DOI [10.1038/nprot.2008.73](https://doi.org/10.1038/nprot.2008.73).



- Steiner K, Hagelueken G, Messner P, Schäffer C, Naismith JH. 2010.** Structural basis of substrate binding in WsaF, a rhamnosyltransferase from *Geobacillus stearothermophilus*. *Journal of Molecular Biology* **397(2)**:436–447 DOI [10.1016/j.jmb.2010.01.035](https://doi.org/10.1016/j.jmb.2010.01.035).
- Wang G, Boulton PG, Chan NWC, Palcic MM, Taylor DE. 1999.** Novel *Helicobacter pylori*  $\alpha$ 1,2-fucosyltransferase, a key enzyme in the synthesis of Lewis antigens. *Microbiology* **145(11)**:3245–3253 DOI [10.1099/00221287-145-11-3245](https://doi.org/10.1099/00221287-145-11-3245).
- Wang G, Li D, Ma X, An H, Zhai Z, Ren F, Hao Y. 2015.** Functional role of *oppA* encoding an oligopeptide-binding protein from *Lactobacillus salivarius* Ren in bile tolerance. *Journal of Industrial Microbiology & Biotechnology* **42(8)**:1167–1174 DOI [10.1007/s10295-015-1634-5](https://doi.org/10.1007/s10295-015-1634-5).
- Whitfield C. 1995.** Biosynthesis of lipopolysaccharide O antigens. *Trends in Microbiology* **3(5)**:178–185 DOI [10.1016/S0966-842X\(00\)88917-9](https://doi.org/10.1016/S0966-842X(00)88917-9).
- Wu Q, Tun HM, Leung FC-C, Shah NP. 2015.** Genomic insights into high exopolysaccharide-producing dairy starter bacterium *Streptococcus thermophilus* ASCC 1275. *Scientific Reports* **4(1)**:133 DOI [10.1038/srep04974](https://doi.org/10.1038/srep04974).
- Zeidan AA, Poulsen VK, Janzen T, Buldo P, Derkx PMF, Øregaard G, Neves AR. 2017.** Polysaccharide production by lactic acid bacteria: from genes to industrial applications. *FEMS Microbiology Reviews* **41(Suppl. 1)**:S168–S200 DOI [10.1093/femsre/fox017](https://doi.org/10.1093/femsre/fox017).
- Zhang Y. 2009.** I-TASSER: Fully automated protein structure prediction in CASP8. *Proteins-structure Function and Bioinformatics* **77(S9)**:100–113 DOI [10.1002/prot.22588](https://doi.org/10.1002/prot.22588).