Norwegian University of Life Sciences

Master's Thesis 2021 60 ECTS

Faculty of Chemistry, Biotechnology and Food Science (KBM)

Anchoring of Invasin on the Surface of *Lactiplantibacillus plantarum* Exploiting Homologous N-terminal Transmembrane Proteins

Anna Bakke Ljosland

Master of Molecular Biology

Acknowledgements

The work described in this thesis was performed at the Faculty of Chemistry, Biotechnology and Food Science at the Norwegian University of Life Sciences with Professor Vincent Ejsink, Researcher Geir Mathiesen and Ph. D. fellow Kamilla Wiull as supervisors.

A huge thanks goes to my main supervisors Geir Mathiesen and Kamilla Wiull. I am grateful for all the help you have given me, both during my time at the laboratory and in the process of writing my thesis. Your optimism and enthusiasm have been inspirational.

Thanks also to Professor Vincent Ejsink for valuable feedback during our laboratory meetings. Furthermore, thanks to Professor Torgeir Rhoden Hvidsten for help with performing the bioinformatic analyses.

I also want to thank my fellow master students, Martin and Olga, for good company at the laboratory. Thanks also to the rest of the PEP-group for providing a good working environment.

Thanks also to friends and fellow students for five memorable years at NMBU.

Last but not least, I am grateful for the encouragement and support provided by my family and friends.

Kongsberg, December 2021

Anna Bakke Ljosland

Abstract

Lactic acid bacteria (LAB) are widely used by humans for food production and preservation, and many LAB are natural inhabitants of the human gastrointestinal tract (GIT). For these reasons, LAB are generally recognized as safe for human consumption. The present study is part of a larger research project where the long-term goal is to develop a mucosal vaccine by utilizing the lactic acid bacterium *Lactiplantibacillus plantarum* as a delivery vehicle. This can be achieved by anchoring the relevant antigens on the surface *of L. plantarum*.

The aim of this thesis was to anchor and display Invasin on the cell surface of *L. plantarum* by exploiting homologous N-terminal transmembrane proteins. Invasin, which is a virulence factor from the pathogen *Yersinia pseudotuberculosis,* can also be used as an adjuvant. This is because display of Invasin on the cell surface of LAB that are used as delivery vehicles can give these bacteria an increased possibility to enter organized lymphoid tissues associated with the small intestine. N-terminal transmembrane proteins are composed of a short N-terminal, intracellular region followed by a transmembrane helix and a C-terminal, extracellular region. To gain insight into proteins that could be well-suited as anchors for Invasin, it was performed an *in-silico* analysis of all predicted N-terminal transmembrane proteins in *L. plantarum* WCFS1. Based on the results of these analyses, six candidate proteins were selected to be used to construct anchors for Invasin. Six anchors with the full N-terminal, anchor region of the protein were made. For three of the anchors, it was also constructed truncated anchors which lacked the intracellular region. All the anchor sequences were translationally fused to Invasin and cloned into the inducible pSIP expression system. Six of the nine plasmids were successfully constructed and transformed into *L. plantarum*. The recombinant *L. plantarum* strains were characterized through growth analyses, western blot, and flow cytometry. In the flow cytometry analyses, it was detected clear fluorescent shifts for the three strains pLp_1751Inv, pLp_2132Inv and pLp_1751Short Inv compared to the negative control strain (pEV). This indicated that Invasin was displayed on the cell surface of these strains. After induction of gene expression, all the recombinant strains showed only a low or moderate reduction in growth, except for pLp_1751Inv, which showed a strong reduction in growth. This is also an indication that this strain produced Invasin. Based on the results obtained in this study, pLp_1751Inv appears to be the most promising strain for anchoring and display of Invasin on the cell surface of *L. plantarum.* It can therefore be interesting to investigate pLp_1751Inv in future studies.

Sammendrag

Melkesyrebakterier benyttes i stor grad av mennesker for produksjon og konservering av matvarer, og mange melkesyrebakterier er en naturlig del av menneskers tarmflora. Melkesyrebakterier regnes derfor generelt som trygge å konsumere for mennesker. Denne studien er del av et større forskningsprosjekt hvor det langsiktige målet er å utvikle en slimhinnevaksine for mennesker ved å benytte melkesyrebakterien *Lactiplantibacillus plantarum* som leveringsvektor. Dette kan oppnås gjennom forankring av relevante antigener på overflaten av *L. plantarum*.

Hensikten med denne oppgaven var å forankre og uttrykke Invasin på overflaten av *L. plantarum* ved bruk av homologe N-terminale transmembranproteiner. Invasin, som er en virulensfaktor fra den sykdomsfremkallende bakterien *Yersinia pseudotuberculosis*, kan også fungere som en adjuvant. Årsaken til dette er at Invasin som uttrykkes på overflaten av LAB som brukes som leveringsvektorer kan gi disse bakteriene en større mulighet til å få tilgang til det organiserte lymfesystemet rundt tynntarmen. N-terminale transmembranproteiner inneholder en kort N-terminal, intracellulær del, etterfulgt av en transmembranheliks og en C-terminal, ekstracellulær del. For å få innsikt i hvilke proteiner som kunne være velegnede som ankre for Invasin ble det utført en *in-silico*-analyse av alle predikerte N-terminale transmembranproteiner i *L. plantarum* WCFS1. Basert på resultatene av disse analysene ble det valgt ut seks kandidatproteiner som skulle brukes til å konstruere ankre for Invasin. Det ble laget seks ankre som inneholdt den N-terminale, intracellulære delen av proteinet. For tre av kandidatproteinene ble det også laget kortere ankre som manglet den intracellulære delen. Alle ankersekvensene ble translasjonelt fusjonert til Invasin og klonet inn i det induserbare pSIP-ekspresjonssystemet. Seks av ni plasmider ble suksessfullt konstruert og transformert inn i *L. plantarum*. De rekombinante *L. plantarum*-stammene ble karakterisert gjennom vekstanalyser, western blot og flow cytometri. I flow cytometri-analysene ble det detektert tydelige fluorescenssignaler for de tre stammene pLp_1751Inv, pLp_2132Inv og pLp_1751Short Inv sammenliknet med den negative kontrollstammen (pEV). Dette indikerte at disse stammene uttrykte Invasin på celleoverflaten. Etter induksjon av gen-ekspresjon viste alle de rekombinante stammene kun en lav til moderat vekstreduksjon, med unntak av pLp_1751Inv som viste en kraftig reduksjon i veksten, noe som også tyder på at denne stammen produserte Invasin. Basert på resultatene av denne studien virker pLp_1751Inv som den mest lovende stammen for forankring og uttrykk av Invasin på celleoverflaten av *L. plantarum*. Det kan derfor være interessant å undersøke pLp_1751Inv i fremtidige studier.

Abbreviations

APC	Antigen-presenting cell
ATP	Adenosine triphosphate
BHI	Brain Heart Infusion
BSA	Bovine Serum Albumin
DC	Dendritic cell
DNA	Deoxyribonucleic acid
DTT	DiThioThreitol
FITC	Fluorescein isothiocyanate
GALT	Gut-associated lymphoid tissue
GIT	Gastrointestinal tract
GRAS	Generally recognized as safe
HPK	Histidine protein kinase
HRP	Horseradish peroxidase
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IP	Inducer peptide
LAB	Lactic acid bacteria
LDS	Lithium Dodecyl Sulphate
M-cell	Microfold cell
MHC	Major Histocompatibility Complex
MRS	de Man, Rogosa and Sharpe
OD	Optical Density
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
RR	Response regulator
SDS-PAGE	Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis
SPase	Signal peptidase
TCA	Trichloroacetic acid

Table of contents

1 Introduction

Lactic acid bacteria (LAB) are a diverse group of bacteria that share a common property, which is the ability to produce lactic acid as a metabolite of the fermentation of sugar. These bacteria have been consumed by humans for centuries, and many LAB have a status as generally recognized as safe (GRAS). All bacteria, including LAB, have natural systems for anchoring of proteins to the cell surface. These systems can be utilized to display heterologous proteins on the cell surface of LAB. Cell surface display of proteins has many applications, for example in industry or medicine. Many LAB are ideal candidates to use as delivery vehicles for medically interesting proteins to the mucosal tissues in the human gastrointestinal tract (GIT) (Michon et al., 2016).

Many studies have focused on using LAB as delivery vehicles for antigens as a strategy to develop mucosal vaccines (Wells & Mercenier, 2008). However, the antigens used in a vaccine are not always sufficient to give rise to an immune response that result in immunity. This can in some cases be solved by adding an adjuvant to the vaccine. Adjuvants are substances that are added to some vaccines, and that contributes to a stronger immune response in people that receive the vaccine (CDC, 2020).

An example of a potential adjuvant is Invasin, which is a virulence factor from *Yersinia pseudotuberculosis.* This is because expression of Invasin on the cell surface of LAB that are used as vaccine delivery vehicles can give the LAB an increased ability to gain access to the secondary lymphoid tissue surrounding the small intestine.

There are four main strategies for surface display of protein in *Lactobacillaceae*. One of these are N-terminal transmembrane anchoring. This thesis describes the construction of anchors from N-terminal transmembrane proteins in *Lactiplantibacillus plantarum* WCFS1. The goal was to use these anchors to display Invasin on the cell surface of *L. plantarum* WCFS1.

1.1 Lactic acid bacteria

Lactic acid bacteria (LAB) are a diverse group of bacteria that share a common property, which is the ability to produce lactic acid as a metabolite from the fermentation of carbohydrates (Daniel et al., 2011; Wyszyńska et al., 2015). Species of LAB are Gram-positive, non-

sporulating bacteria, with a low G+C content in their genome. Some LAB species are rod-shaped (rods), while others have a round or oval form (cocci). Another characteristic of LAB is that they lack many genes coding for proteins that are involved in several different biosynthetic pathways (Wyszyńska et al., 2015). They can therefore only survive in environments that are rich in pre-formed purines, pyrimidines, amino acids, and B vitamins. In addition, they usually depend on a sugar that they can utilize as an energy and carbon source (Wells & Mercenier, 2008; Wyszyńska et al., 2015). This limits their number of possible habitats. However, they are still present in many different niches, such as plant surfaces, milk, and the oral cavity, gastrointestinal tract, and vagina of vertebrates (Wells & Mercenier, 2008). LAB have a high ability to survive the passage through the stomach because they are resistant to bile and low pH (Wells & Mercenier, 2008).

LAB are aerotolerant, but most LAB prefer an anaerobic environment (Axelsson, 2004). LAB can be divided into two main groups based on the pathway they utilize for carbohydrate fermentation. Bacteria that use the homofermentative pathway ferment glucose through glycolysis and form lactic acid as the main product. On the other hand, fermentation of glucose through the heterofermentative pathway leads to formation of CO_2, ethanol and/or acetic acid in addition to lactic acid (Kandler, 1983).

Species of LAB are found in many different genera, including *Lactococcus, Streptococcus, Enterococcus, Pediococcus, Leuconostoc* and many others (Wyszyńska et al., 2015; Zheng et al., 2020). The term LAB also includes the bacterial species that previously belonged to the genus *Lactobacillus*. This genus was recently divided into 23 new genera (Zheng et al., 2020), as described in more detail in section 1.2.

Due to their fermentative properties, LAB has been widely used by humans for food production and preservation (Daniel et al., 2011; Michon et al., 2016), and this makes them very important economically (de Vos, 2011). For example, some species of LAB have been used to produce cheese and yoghurt from milk (Wells & Mercenier, 2008). Members of the now emended genus *Lactobacillus* are used as starter cultures to produce fermented vegetables, fermented meat, beer, wine and sourdough bread (Sørvig et al., 2003).

LAB have been consumed by humans for thousands of years, and many species in this group are natural inhabitants of the gastrointestinal tract (GIT) of humans and animals (Daniel et al., 2011). For these reasons, many LAB have been given a GRAS status, meaning that they are **generally recognized as safe** to consume for humans (Diep et al., 2009). Some LAB strains

are also marketed as probiotics (Daniel et al., 2011; de Vos, 2011). The Food and Agriculture Organization (FAO) and the World Health Organization (WHO) have defined probiotics as: "Live microorganisms which when administered in adequate amounts confer a health benefit on the host" (FAO/WHO, 2001). However, there are also some species of LAB that are pathogenic, for example *Streptococcus pyogenes*, *Streptococcus pneumoniae* and some species of *Enterococcus* (Fisher & Phillips, 2009; Henriques-Normark & Tuomanen, 2013; Walker et al., 2014).

As mentioned, many LAB have a GRAS status and are able to survive the harsh conditions in the human GIT (Michon et al., 2016). This, in combination with many other properties of LAB, make them promising delivery vehicles of medically interesting proteins, including vaccines (Kleerebezem et al., 2003; Mohamadzadeh et al., 2009).

1.2 *Lactiplantibacillus plantarum*

Lactiplantibacillus plantarum is a versatile bacterium that can adapt to many different niches in the environment. It is found in many plant and vegetable fermentations, meat, and dairy products (Kleerebezem et al., 2003; Zheng et al., 2020). *L. plantarum* is also a natural inhabitant of the human gastrointestinal tract (GIT) and have a close association with the intestinal epithelium (Mohamadzadeh et al., 2009). Some strains of *L. plantarum* have also been reported to have probiotic effects on human health (Gareau et al., 2010).

Until recently, this bacterium was named *Lactobacillus plantarum* and belonged to the genus *Lactobacillus*. This was a very large genus; it comprised 261 bacterial species that were genetically, phenotypically, and ecologically diverse. Therefore, in the spring 2020, it was done changes in the taxonomy that involved a division of the genus *Lactobacillus* into 23 new genera. The bacterial species previously known as *Lactobacillus plantarum* now belongs to one of these new genera, called *Lactiplantibacillus*, and has therefore been given the new name *Lactiplantibacillus plantarum*. Members of the genus *Lactiplantibacillus* are homo-fermentative, non-sporeforming and non-motile rods that can ferment many types of carbohydrates. *L. plantarum* is the type species of this genus (Zheng et al., 2020).

Among the bacterial species that were members of the now emended genus *Lactobacillus*, *Lactiplantibacillus plantarum* is one of the species that is most studied and best understood (Diep et al., 2009). *L. plantarum* WCFS1 is a single colony isolate of the strain NCIMB8826 that originally was isolated from human saliva (Kleerebezem et al., 2003). This strain was the first *Latobacillus* to be sequenced (Kleerebezem et al., 2003).

L. plantarum WCFS1 has the largest genomes of all sequenced lactic acid bacteria (Siezen & van Hylckama Vlieg, 2011) (Chevallier et al., 1994), and this can explain why this bacterium has such a high ecological flexibility (Kleerebezem et al., 2003). *L. plantarum* has many properties that makes it ideal as a delivery vehicle of molecules to mucosal tissues. It is GRAS, has a high survival rate when it passes through the stomach (Mohamadzadeh et al., 2009; Wells & Mercenier, 2008), and can survive for up to 6-7 days in the human GIT (Kleerebezem et al., 2003; Michon et al., 2016). Furthermore, there has been developed genetic tools for *L. plantarum*, including gene expression systems (Kleerebezem et al., 2003; Sørvig et al., 2003), that can make it easier to genetically modify this bacterium in experiments.

1.3 Gene expression systems in lactic acid bacteria (LAB)

Species of LAB can be genetically engineered to express a heterologous protein of interest. To achieve this, the gene coding for that protein must be integrated into the DNA of the LAB, either on the chromosome or on a plasmid. There are two main types of gene expression systems in LAB: Constitutive expression systems and inducible expression systems. Which of the two systems that is most suitable, will vary depending on the aim of the experiment.

When using a constitutive system for gene expression, the level of gene expression can be fine-tuned to an optimal level (Rud et al., 2006). This is because different constitutive promoters vary in strength.

1.3.1 Inducible gene expression systems

In an inducible gene expression system, the expression of the gene of interest is driven by an inducible promoter (Sørvig et al., 2005). This type of expression system is very suitable in experiments where overproduction of a protein to a high level is desired (Rud et al., 2006).

Two important, inducible gene expression systems for use in LAB are the NICE-system and the pSIP-system. Both systems are developed using regulatory promoters and genes that naturally exists in LAB, and that are involved in bacteriocin production. Bacteriocins are ribosomally synthesized, antimicrobial peptides that is secreted by LAB to kill closely related bacterial species in the environment (Diep et al., 2009). The production of bacteriocins is in many cases strictly regulated via quorum-sensing mechanisms that involve a secreted peptide pheromone (Inducer peptide, IP), a sensor located in the membrane called a histidine protein kinase (HPK) and a cytoplasmic response regulator (RR) (Diep et al., 2009) (Eijsink et al., 2002). The genes coding for the IP, HPK and RR are usually located on the same, regulatory

operon (Diep et al., 2009; Sørvig et al., 2003). The production of bacteriocins is controlled through a series of steps, which is illustrated in detail in figure 1.1.

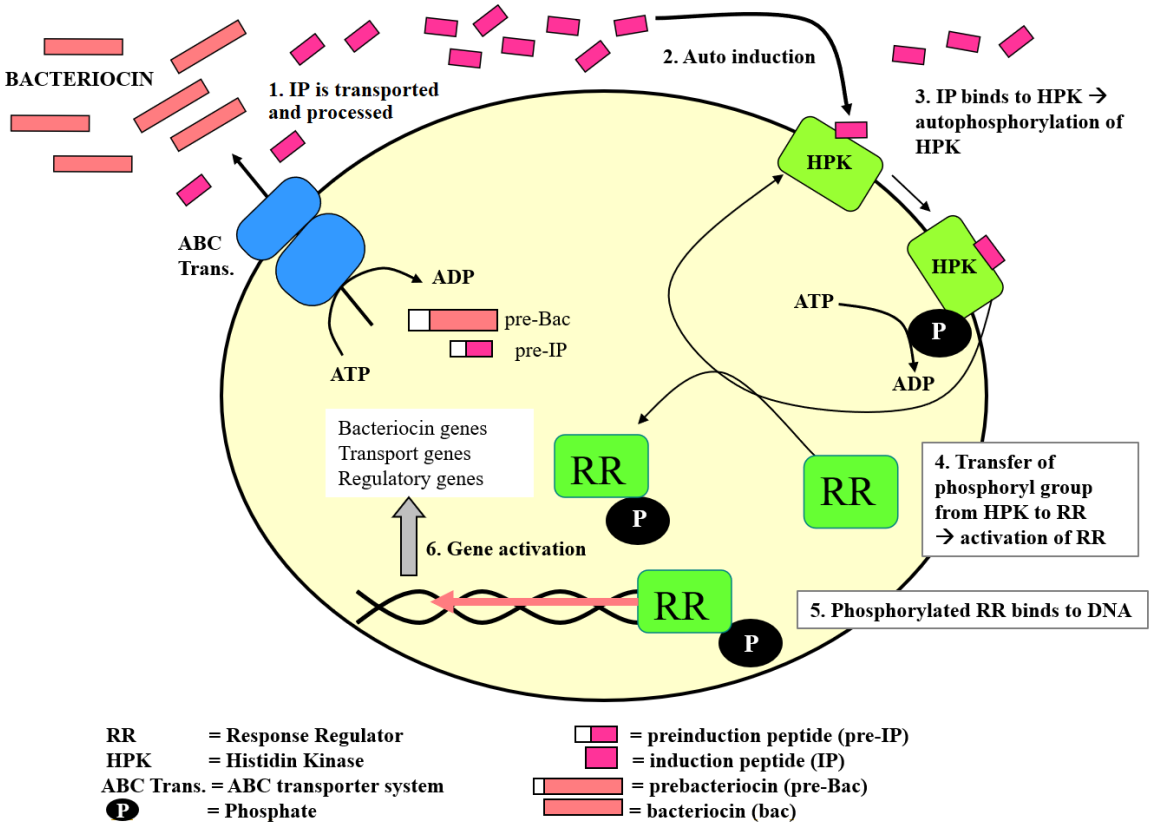


Figure 1.1. Illustration of pheromone-regulated production of a class II bacteriocin. During cell growth, there is a low constitutive, production of the peptide pheromone (IP). The IP is transported out of the cell through an ABC transporter (1). The concentration of IP in the surrounding medium increases and will reach a threshold concentration. This makes the IP bind to its membrane-located sensor, the histidine protein kinase (HPK) (2). This leads to autophosphorylation of the HPK (3). Subsequently, the phosphoryl group is transferred from HPK to the response regulator (RR), which then becomes activated (4). The activated, phosphorylated RR binds to DNA on regulated promoters that drives the expression of operons involved in bacteriocin production (5). This activates transcription of all operons involved in the production of bacteriocins, including the regulatory operon itself (6) (Sørvig et al., 2003). The figure is modified from (Mathiesen, 2004).

1.3.2 The pSIP-system

The pSIP-system is a set of vectors that were constructed for inducible gene expression in *Lactiplantibacillus plantarum* and *Latilactobacillus sakei* (Fig. 1.2). This expression system is a one-plasmid system that is based on regulatory genes and promoters that are naturally involved in the production of the class II bacteriocins sakacin A and sakacin P in *L. sakei* (Fig.

1.1) (Sørvig et al., 2003; Sørvig et al., 2005). The expression of the genes coding for these bacteriocins are driven by strong, regulated promoters (Diep et al., 2009).

Class II bacteriocins, like sakacin A and sakacin P, do not function as peptide pheromones. Bacterial strains that produce these types of bacteriocins must therefore also produce and secrete a separate peptide pheromone. The gene encoding the peptide pheromone is in most cases co-transcribed with the genes coding for the histidine protein kinase (HPK) and the response regulator (RR) (Eijsink et al., 2002; Sørvig et al., 2005).

The pSIP-expression vectors that were used in this study contain the regulatory genes involved in production of the bacteriocin sakacin P (the *spp* gene cluster) (Sørvig et al., 2003; Sørvig et al., 2005). These vectors contain the promoter that naturally drives the expression of sakacin P, called P_{sppA}, and the HPK and RR that are responsible for activating this promoter when induced by a peptide pheromone (IP) (Fig. 1.2). In these pSIP-vectors, the promoter P_{sppA} is used to achieve strictly regulated expression of a gene of interest (Sørvig et al., 2003; Sørvig et al., 2005).

In the pSIP-vectors used in this study, the gene coding for the inducer peptide, *sppIP*, has been deleted (Sørvig et al., 2003). This ensures that the gene of interest only will be expressed after external addition of IP. The pSIP-vectors are built up of cassettes, making it easy to exchange all elements through restriction enzyme digestion and ligation (Sørvig et al., 2003). In addition, the vectors contain replication determinants for *E. coli*, *L. sakei* and *L. plantarum* and the expression system can therefore function in all these bacterial species.

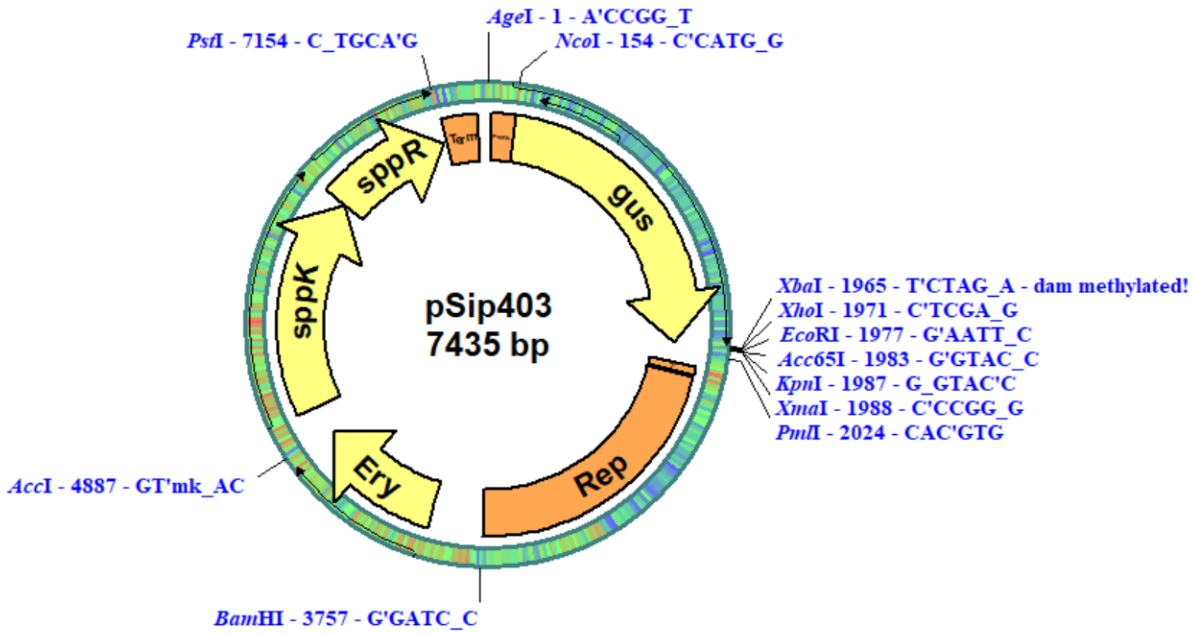


Figure 1.2. The pSIP-vector expression system. Illustration of the pSIP403 plasmid, which is one of several plasmids that were constructed in the development of the pSIP-system for inducible gene expression (Sørvig et al., 2003). The sppK gene encodes the histidine protein kinase (HPK) and the sppR encodes the response regulator (RR). Ery is the erythromycin resistance marker. The *gus* gene encodes β-glucuronidase from *E. coli*, which was used as a reporter enzyme to evaluate the functionality of the pSIP-vectors (Sørvig et al., 2003). The orange region preceding the *gus* gene is the P_{sppA} promoter. The *gus* gene can easily be exchanged with a gene of interest through restriction enzyme digestion and ligation. The Rep region contains replication determinants, and a multiple cloning site (MCS) is located between *gus* and Rep. The plasmid map is constructed using pDraw32.

1.4 Secretion and anchoring of proteins in Gram-positive bacteria

Gram-positive bacteria have only a single cell membrane which is covered by a thick cell wall composed of peptidoglycan. In contrast, Gram-negative bacteria have only a thin layer of peptidoglycan in their cell wall. In addition to an inner membrane, they also have an outer membrane that covers the peptidoglycan layer. Compared to Gram-negative bacteria, protein secretion is easier in Gram-positive bacteria since they only have a single cell membrane (Green & Mecsas, 2016).

Proteins that have a function outside of the bacterial cell must be translocated across the membrane. This includes proteins that are completely secreted and proteins that are anchored to the cell membrane or cell wall. These proteins are generally synthesized as preproteins with an N-terminal signal peptide sequence which informs the cell about the correct destination of the protein (Tjalsma et al., 2004). In general, signal peptides consist of three main domains: N, H and C. The N region often contains one to three positively charged amino acids, such as

arginine and lysine (Driessen & Nouwen, 2008; Tjalsma et al., 2004). The N domain is followed by the H domain, consisting of 10-15 residues (Driessen & Nouwen, 2008). Since most of the residues in this domain are hydrophobic, it can form an alpha helix in the membrane (Tjalsma et al., 2004; van Roosmalen et al., 2004). The H domain is followed by the C domain, which is more polar and contains a cleavage site for signal peptidase (Driessen & Nouwen, 2008).

There are seven different secretion pathways that have been characterized in Gram-positive bacteria, and one of these is the secretion pathway (sec). Most proteins that are going to be secreted, are translocated across the membrane via this pathway. Proteins that must be retained in the cell membrane or cell wall, are also translocated via the Sec pathway (Tjalsma et al., 2004). The main components in the sec translocase machinery are a membrane channel for protein transport, secYEG, and a motor protein, SecA, that is driven by ATP (Prabudiansyah et al., 2015) (Fig. 1.3).

The transport of proteins to the sec translocase machinery can happen in several ways. In some bacterial species, the protein is guided to the sec translocase machinery by a chaperone (Fig 1.3). The protein is translocated through the membrane in an unfolded state (Michon et al., 2016). The motor protein, SecA, delivers energy (ATP) that is required for translocating the protein through the SecYEG channel.

Proteins that are going to be completely secreted, contain a cleavage site with the consensus motif A – X – A located in the C domain of their signal peptide. These proteins are recognized and cleaved by a membrane-bound signal peptidase I (van Roosmalen et al., 2004). The SPase cleaves off the signal peptide from the protein during, or shortly after, translocation across the membrane. The signal peptide is subsequently degraded and removed from the membrane (Tjalsma et al., 2004). The mature protein is released from the membrane and can fold into its native conformation (Tjalsma et al., 2004).

Proteins that do not have a SPase cleavage site in their N-terminal signal peptide sequence will remain uncleaved after translocation through the SecYEG channel. The absence of a SPase cleavage site signals to the cell that the protein should be retained in the membrane (Tjalsma et al., 2004).

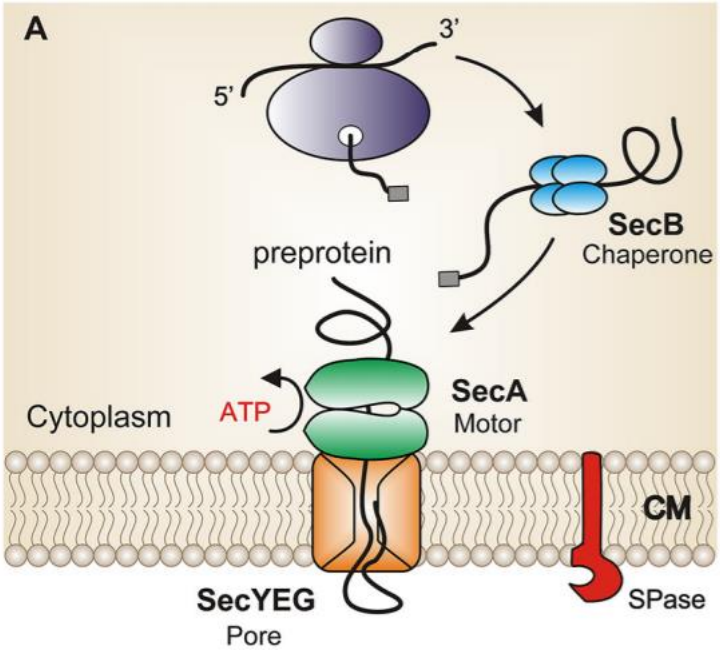


Figure 1.3 The Sec translocase system. Illustration of the sec translocase machinery in its minimal form. Proteins that must be translocated over the membrane are synthesized as preproteins with an N-terminal signal peptide. Most of these proteins are translocated via the Sec pathway. After the preprotein has been synthesized in the ribosome (shown in purple) it can be guided to the sec translocase machinery in several ways, for example with help from the chaperone SecB. The preprotein is translocated through the SecYEG channel in an unfolded state, using energy (ATP) that is provided by the motor protein SecA. A membrane-bound signal peptidase recognizes and cleaves the preproteins that contain a SPase cleavage site in their N-terminal signal peptide. The figure is from an article written by (Prabudiansyah et al., 2015), which is an open access article distributed under the terms of the Creative Commons Attribution License: https://creativecommons.org/licenses/by/4.0/.

1.5 Anchoring of proteins to the surface of *Lactobacillaceae*

Many proteins have a function in the cell wall or the cell membrane and must therefore be retained in those locations. Proteins that are anchored to the cell membrane or cell wall are for example cell wall hydrolases, penicillin-binding proteins that are involved in peptidoglycan synthesis, substrate-binding proteins, DNases, and RNases (Tjalsma et al., 2004).

Proteins that are anchored to the cell membrane or the cell wall in *Lactobacillaceae* can potentially be utilized to anchor and express heterologous proteins on the cell surface. There is an increasing interest in displaying heterologous proteins on the surface of LAB since this is often necessary in studies where these bacteria are used as mucosal delivery vehicles of medically interesting proteins (Michon et al., 2016). Display of heterologous proteins requires the use of a signal peptide and an anchor domain (Michon et al., 2016).

The single cell membrane and the thick, peptidoglycan-containing cell wall of Gram-positive bacteria, such as LAB, makes it possible to anchor proteins in many ways (Michon et al., 2016). When it comes to the *Lactobacillacea*e family, there are four main anchoring strategies that can be used to anchor proteins for cell surface display (Fig. 1.4):

1. Anchoring to the cell membrane using an N-terminal transmembrane anchor
2. Anchoring to the cell membrane using a lipoprotein anchor
3. Non-covalent anchoring through the association with a cell wall binding domain (e.g., a LysM domain)
4. LPxTG peptidoglycan anchors for covalent anchoring to the cell wall

The pSIP gene expression system, which is described in section 1.3, have been used to achieve protein secretion and protein anchoring using C-terminal cell wall anchors (Fredriksen et al., 2010; Mathiesen et al., 2008; Mathiesen et al., 2009). Later, the same system was further developed for N-terminal anchoring of proteins (Fredriksen et al., 2012).

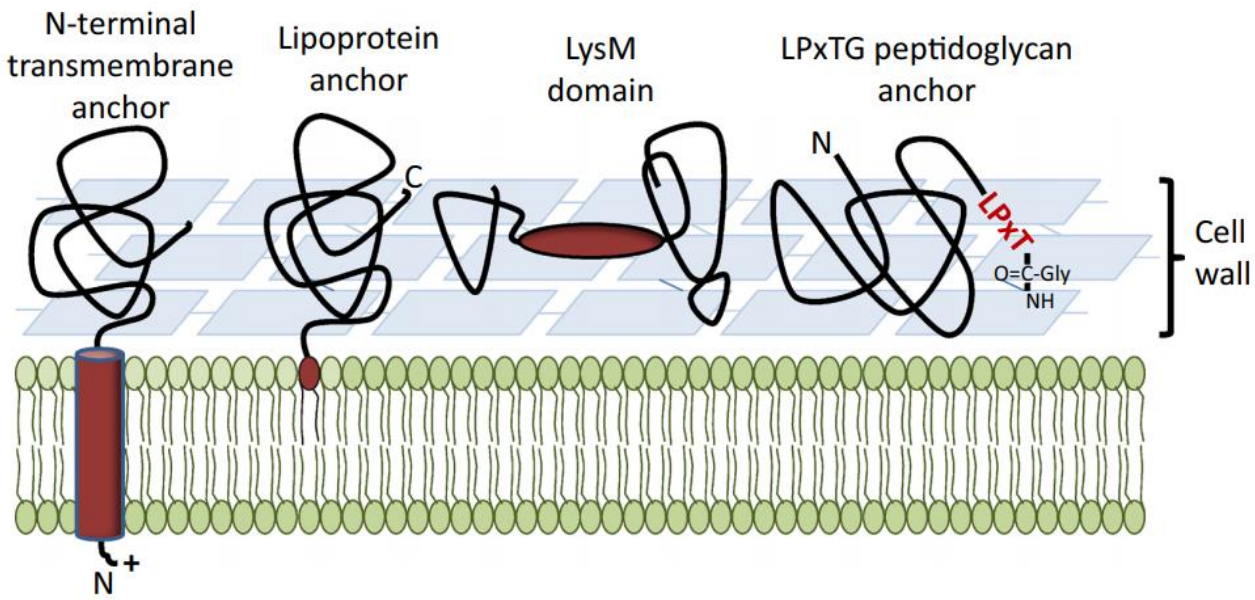


Figure 1.4. Methods for display of proteins in members of the *Lactobacillacea*e family. Illustration of the four most widely used strategies for protein display in lactobacilli. These are, from left to right: anchoring of the protein to the membrane using an N-terminal transmembrane anchor (1), covalent binding to a phospholipid in the membrane using a lipoprotein anchor (2), non-covalent anchoring to the cell wall through association with a cell wall binding domain (for example LysM) (3) or covalent anchoring to the cell wall using an LPxTG peptidoglycan anchor. The figure is taken from an article written by (Michon et al., 2016), which is distributed under the Creative Commons Attribution 4.0 International License (Creative Commons — Attribution 4.0 International — CC BY 4.0).

1.5.1 N-terminal transmembrane anchors

N-terminally anchored transmembrane proteins generally contain signal peptides (SPs), but their signal peptide sequence lack a cleavage site that can be recognized by SPase (Michon et al., 2016). After these proteins are translocated over the membrane, they will not be cleaved by any SPases and are therefore retained in the membrane (Fig 1.4). The hydrophobic H domain of the signal peptide sequence forms an alpha helix in the phospholipid bilayer that anchors the protein to the membrane (Michon et al., 2016). Hydrophobic residues in this domain generally contribute to helix formation, except for glycine and proline, which can destabilize the helical structure and are therefore known as “helix breakers” (Dong et al., 2012; Woolfson & Williams, 1990). The H domain of N-terminal transmembrane proteins is followed by an extracellular part which is often the largest part of the protein.

Like SPs of many other proteins, the SPs of N-terminal transmembrane proteins also contain an N domain that precedes the H domain. This domain is intracellular and will therefore be referred to here as the “intracellular tail” (iT). It is defined as the sequence from the start amino acid (methionine) to the last amino acid upstream of the predicted helix. This domain often has a positive net charge due to the presence of one or more of the positively charged residues lysine (K) and arginine (R) (Tjalsma et al., 2004). It is known arginine and lysine play an important role in determining the orientation of a helix. These residues are mostly found in non-transmembrane regions of the protein. This is known as “the positive-inside rule” (Sonnhammer et al., 1998).

The function of this intracellular tail is not known, but it has been suggested that one of the lysine or arginine residues in this region interacts with the translocation machinery and the membrane phospholipids, which are negatively charged (Tjalsma et al., 2004).

N-terminal transmembrane proteins can be utilized as anchors for protein display on the cell surface of LAB. To achieve this, the protein of interest is translationally fused to the anchor (Michon et al., 2016). The anchor sequence comprises the N-terminal part of the protein, containing the iT, the transmembrane helix and a part of the sequence that follows the helix on the extracellular side, known as the linker. The linker is the sequence between the transmembrane helix of the N-terminal transmembrane protein and the fused heterologous sequence. The anchor is, in most cases, a C-terminally truncated version of the complete N-terminal transmembrane protein. The anchor length can be varied by varying the proportion of the linker part of the protein that is used as the anchor sequence.

The whole N-terminal transmembrane protein can be used as an anchor if desired, or one can choose to only include a few residues after the signal peptide, which then will function as a linker (Michon et al., 2016). When constructing anchors for surface display of heterologous proteins, it is important to evaluate linker length. A very short anchor can make the displayed protein embedded in the cell wall, resulting in a low accessibility. On the other hand, if the anchor is very long, the protein can become completely exposed to the environment surrounding the cells, and hence make it more susceptible to proteolytic degradation.

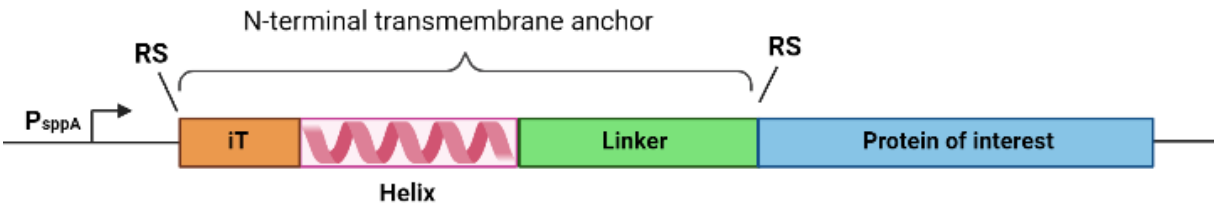


Figure 1.5. Schematic overview of an expression cassette for N-terminal transmembrane anchoring of a protein of interest. The illustration shows the composition of an N-terminal transmembrane anchor, and how a protein of interest can be fused to this anchor sequence for cell surface display. The anchor is composed of an intracellular, N-terminal region (iT), a transmembrane alpha helix, and an extracellular linker region that is located between the helix and the heterologous protein of interest. The N-terminal region of the protein of interest is anchored to the C-terminal region of the anchor sequence. The DNA sequence coding for the anchor can easily be cloned into a pSIP-expression vector, in the region directly upstream of the gene sequence encoding the protein of interest. This is achieved by introducing restriction sites (RS) in the beginning and the end of the anchor sequence that corresponds to the restriction sites that are used to digest the pSIP-expression vector. The figure is inspired by FIG 1 in the article written by (Fredriksen et al., 2012). Created with BioRender.com.

1.5.2 Lipoprotein anchors

Lipoproteins contain an N-terminal signal peptide and are secreted via the Sec pathway. The C-terminal part of their signal peptide contain a “lipobox motif” with the consensus sequence L – X – X – C. After the protein has been secreted via the Sec pathway, an enzyme called diacylglycerol transferase catalyzes a reaction that leads to covalent binding between the conserved cysteine in the lipobox motif and a phospholipid in the membrane. Subsequently, signal peptidase II recognizes the lipobox motif and cleaves off the signal peptide. Consequently, the cysteine that was part of the lipobox becomes the first residue of the mature version of the lipoprotein (Michon et al., 2016; Tjalsma et al., 2004). Lipoproteins can function as membrane anchors for surface display of proteins. This is achieved by fusing the protein of interest to the lipoprotein downstream of the lipobox (Michon et al., 2016).

1.5.3 Non-covalent anchoring to the cell wall through interaction with a cell wall binding domain

There are many different types of cell wall binding domains in Gram-positive bacteria, and these can be utilized to non-covalently attach proteins to the cell wall (Michon et al., 2016). Examples of cell wall binding domains are LysM domains, surface layer proteins, surface layer homology domains (SLPs and SLHDs), GW modules and WxL domains (Michon et al., 2016). LysM domains can bind to peptidoglycan and chitin. They vary in length from 44 to 65 residues, and single or multiple proteins of these domains can be present in one protein. Since these domains can bind to peptidoglycan, they have been widely used for surface display (Fig 1.4) (Visweswaran et al., 2014).

1.5.4 Covalent anchoring to the cell wall using LPxTG peptidoglycan anchors

Some surface proteins are covalently attached to the cell wall via an LPXTG anchor (Fig. 1.4). These proteins contain the sequence motif LPXTG that makes the protein become C-terminally anchored in the cell wall. This sequence motif is followed by a stretch of hydrophobic amino acids and a short region of positively charged amino acids (Michon et al., 2016) Like N-terminal transmembrane proteins and lipoproteins, proteins containing a LPXTG anchor have an N-terminal signal sequence that lead to membrane translocation via the Sec pathway. Upon secretion and SP cleavage, the enzyme sortase, which is a transpeptidase, cleaves between threonine and glycine in the LPXTG-motif. This leads to anchoring of the protein because the threonine residue becomes covalently attached to the peptidoglycan in the cell wall (Michon et al., 2016).

LPXTG anchors can be used to anchor a heterologous protein to the cell wall of LAB. The protein of interest is then C-terminally fused to a C-terminal region of the protein anchor that includes the LPXTG domain (Michon et al., 2016). Hence, LPXTG anchors lead to C-terminal anchoring of proteins, in contrast to N-terminal transmembrane anchors and lipoprotein anchors, which both result in N-terminal anchoring of proteins.

There are many studies where LPXTG anchors have resulted in successful anchoring and display of heterologous proteins on the cell surface of LAB (Cortes-Perez et al., 2005; Fredriksen et al., 2010; Kajikawa et al., 2011; Kuczkowska et al., 2016; Minic et al., 2015). For example, in a study by (Fredriksen et al., 2010), the 37-Kilodalton oncofetal antigen (OFA), which is a tumor immunogen, was successfully anchored to the cell wall of *L. plantarum* using

a LPXTG anchor. The study showed that the *L. plantarum* cells displaying OFA induced a specific immune response against OFA in mice (Fredriksen et al., 2010).

1.6 The immune system of the human intestine

The human intestine is a complex ecosystem containing an enormous amount of bacteria, and other microorganisms like archea, viruses and eukaryotic microorganisms (Shreiner et al., 2015). Together, they make up the intestinal microbiota. Some bacteria are regarded as natural inhabitants of the intestine (commensals) that are normally harmless. In some cases, pathogenic variants of commensals can arise because new genetic variants acquire virulence factors that makes them able to breach the intestinal epithelium and invade underlying tissues. This must be prevented by the mucosal immune system of the intestine (Parham, 2015).

The intestinal immune system is a complex system that involves immune cells from both the innate immune system and the adaptive immune system. It has an important role in preventing microorganisms in the intestine from causing infection. However, to maintain a healthy gut, inflammation must be avoided. This is because inflammation generally lead to disruptions and damaged tissue, and this makes it easier for intestinal microorganisms to invade the underlying tissues (lamina propria) (Parham, 2015). Dendritic cells (DCs) in the intestine have an important role as antigen-presenting cells (APC's) (Parham, 2015). DCs within the lamina propria can also extend their dendrites through the tight junctions between epithelial cells, and all the way out into the lumen, where they sample the environment by capturing antigens. Subsequently, DCs travel to one of the T-cell areas in the lymphoid tissues, where the processed antigens on their cell surface are presented to naive antigen-specific CD4 T-cells (Mohamadzadeh et al., 2009; Parham, 2015). This leads to activation and differentiation of the T-cells into helper CD4 T_{FH} cells. B-cells can also take up the antigens and present these on their cell surface. The activated helper CD4 T_{FH} cells activate B-cells that have taken up antigen, resulting in differentiation of the B-cells into plasma cells. The plasma cells secrete antigen-specific immunoglobulin A (IgA) into the lumen (Parham, 2015).

The intestinal tract is surrounded by secondary lymphoid tissues that are collectively called the **gut-associated lymphoid tissue** (GALT). The lamina propria and the Peyer's patches are important parts of the GALT. Efferent lymphatics arising in the lamina propria and the Peyer's

patches drain to **mesenteric lymph nodes**. These lymph nodes are not considered a part of the GALT, but they still have an essential role in defending the intestine (Parham, 2015).

An important part of the GALT is the connective tissue that underlies the intestinal epithelium, called the **lamina propria**. Many immune cells reside in this area, for example plasma cells, T-cells, macrophages, and mast cells (Parham, 2015). Macrophages located in the lamina propria are a "first line of defense" against invading microorganisms.

The part of the intestine that is covered with the most lymphoid tissue is the small intestine (Parham, 2015). The wall of the small intestine is made up of a single layer of epithelial cells. The epithelial wall is folded into characteristic projections called **villi**, and the apical membrane of the epithelial cells is further folded into **microvilli**.

The **Peyer's patches** are important secondary lymphoid organs of the small intestine. These patches lie directly underneath the epithelial cell wall. Regions of the epithelium that overlays the mucosal lymphoid follicles, such as the Peyer's patches, are called the follicular-associated epithelium (Wells & Mercenier, 2008). The Peyer's patches are populated with T-cells, B-cells, and dendritic cells. They are also associated with microfold cells (M-cells), which are highly specialized immune cells located in the follicular-associated epithelium (Parham, 2015).

M-cells have an important role in sampling the environment of the gut. On the apical membrane, these cells express many different adhesion molecules and cell-surface receptors that can recognize antigens in the lumen. Microorganisms and antigens that bind to receptors on the apical membrane of the M-cell, can be internalized in membrane vesicles (Parham, 2015). In a process called transcytosis, these membrane vesicles are transported across the M-cell and released on the basolateral side of the membrane. The antigens and microorganisms ends up in the Peyer's patch beneath the M-cell, where they can encounter dendritic cells and other immune cells (Parham, 2015). The dendritic cells and B-cells in the Peyer's patch can detect and capture the translocated antigens, and will further stimulate antigen-specific T-cells and B-cells to proliferate and differentiate (Parham, 2015). Importantly, M-cells are generally believed to be the main entry point for bacteria and particulate antigens (Wells & Mercenier, 2008).

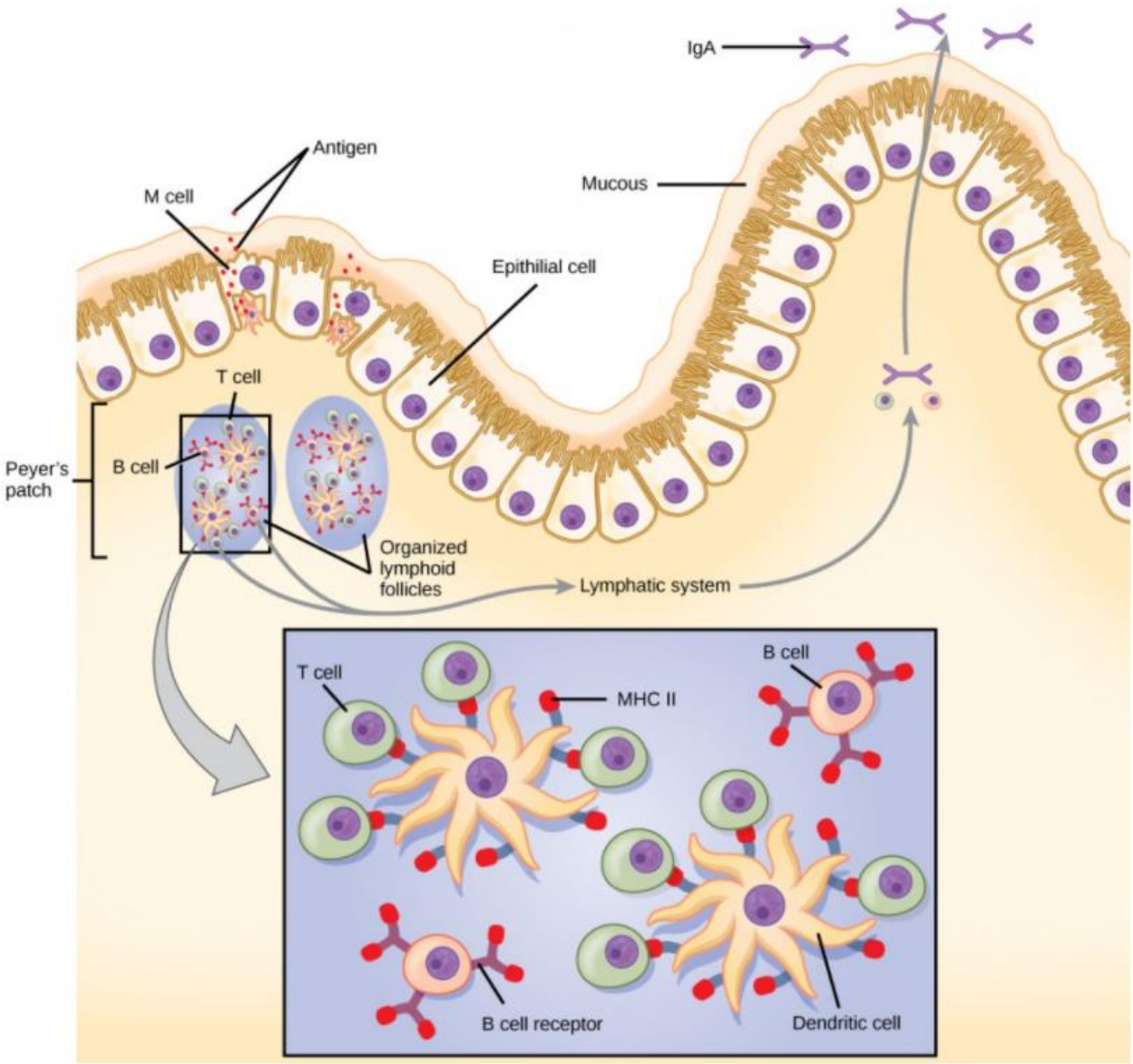


Figure 1.6. Illustration of the topology and function of the mucosa-associated lymphoid tissue (MALT) in the small intestine. Antigens, or whole bacteria, can be taken up by M-cells in the epithelium of the small intestine. The bacteria or antigens are transcytosed and excreted into an intraepithelial pocket in the M-cell. Antigen-presenting cells, such as dendritic cells, that reside in this pocket, phagocytose the antigens and present them on the cell surface by using MHC II molecules. The DCs then travel to underlying, organized lymphoid follicles called a Peyer's patch, where APC's and naive T-cells and B-cells are very abundant. The antigens presented by DCs can lead to activation of some of the naive T-cells and B-cells, which will then divide and differentiate into effector cells. Some of the antigen-expressing DCs travel via the lymphatic system to other lymph nodes, where they activate B-cells and T-cells. Activated immune cells return to MALT. The activated B-cells (plasma cells) secrete specific IgA into the lumen of the intestine (Gair, 2021). The figure is from the book "Concepts of Biology" by Charles Molnar and Jane Gair. This book is licensed under a Creative Commons Attribution 4.0 International License: Creative Commons — Attribution 4.0 International — CC BY 4.0.

1.7 Mucosal delivery of medically interesting proteins

There has been an increasing interest in developing LAB that can be used as mucosal delivery vehicles (Wells & Mercenier, 2008). This is because these delivery vehicles have many potential applications, for example delivery of vaccines and management of inflammatory bowel diseases and auto-immune diseases (Michon et al., 2016). However, LAB can be used to deliver a wide range of molecules with different applications, for example anti-infectives, therapies for gastrointestinal diseases, and therapies for allergic diseases (Wells & Mercenier, 2008). LAB have been used to produce many proteins that are medically interesting, like insulin (Chen et al., 2007), tetanus toxin (Norton et al., 1997; Robinson et al., 2004) and leptin (Bermúdez-Humarán et al., 2007).

In addition to LAB, other microorganisms have also been used as mucosal delivery vehicles (Wells & Mercenier, 2008). For example, several attenuated pathogens, like *Salmonella*, *Bordetella* and *Listeria* have been successfully used as vaccine vectors to deliver heterologous antigens. However, there is a concern that attenuated pathogens used as delivery vehicles can regain their virulence (Mohamadzadeh et al., 2009). LAB strains, such as *L. plantarum*, can be an attractive alternative to use as delivery vehicles for mucosal immunization (Grangette et al., 2001; Lee et al., 2006). Their GRAS-status, in combination with many other properties make LAB well-suited as delivery vehicles for medically interesting proteins, for example antigens. Lactobacilli have a high survival rate through the stomach (Mohamadzadeh et al., 2009), and are present in large numbers in the small intestine of humans (Wells & Mercenier, 2008). In addition, lactobacilli have immunomodulatory properties (Mohamadzadeh et al., 2009), meaning that they can skew the immune response in a direction that is beneficial to the host (Kuczkowska et al., 2017).

In studies where LAB has been exploited as delivery vehicles, the focus has mainly been to use these LAB in the development of mucosal vaccines (Wells & Mercenier, 2008). Most pathogens invade the body through the mucosal surfaces, and it is therefore important to develop efficient mucosal vaccines. An advantage with vaccines that are delivered via the mucosal routes, for example by nasal or oral administration, is that they have the potential to induce both mucosal and systemic immune responses. On the other hand, traditional vaccines that are delivered via the subcutaneous or intramuscular route, will in most cases lead to no, or very weak mucosal immunity (Lycke, 2012).

The antigens, or other medically interesting proteins, that are to be delivered to the immune system of the host, can be expressed using one of the three different expression strategies in LAB: cytoplasmic expression (CE), extracellular secretion (ES) or cell surface display (SD) (Oh et al., 2021). If the latter strategy is used, the protein of interest can be anchored on the bacterial cell surface by either of the anchors described in section 1.5.

Several vaccine candidates based on recombinant LAB have shown promising results in mice. For example, recombinant *L. plantarum* WCFS1 has led to partial protection against *Helicobacter felis* in mice (Corthésy et al., 2005). Also, a recombinant strain of *L. plantarum* NCIMB8826 that expressed tetanus toxin fragment C (TTFC) in the cytoplasm led to specific immune responses against this antigen in mice after intranasal administration (Grangette et al., 2001).

The ability to anchor and express antigens on the cell surface of *L. plantarum* WCFS1 makes it possible to develop vaccines against many different diseases. Dr. Geir Mathiesen and his research team has managed to anchor and display many different antigens on the surface of *L. plantarum* WCFS1, for example antigens from *Mycobacterium tuberculosis* (Kuczkowska et al., 2019). It was tested whether two *L. plantarum*-strains that displayed the tuberculosis antigens Ag85B and ESAT-6 on the cell surface were able to protect mice from infection by *M. tuberculosis*. The mice were vaccinated with the *L. plantarum*-based vaccines through intranasal immunization. The results showed that both strains of *L. plantarum* could be used as a “BCG boost” since they enhanced the protection against tuberculosis that was given by the BCG-vaccine (Kuczkowska et al., 2019).

Unfortunately, when non-pathogenic bacteria such as *L. plantarum* is used in immunization strategies, complete protection against the disease is usually not achieved. In some cases, vaccine adjuvants are needed to enhance the immune response generated by the bacterium that express the antigen (Fredriksen et al., 2012; Wells, 2011). For example, if adjuvants and antigens are coexpressed on the cell surface of *L. plantarum* used in a vaccine, it might result in a stronger immune response (Fredriksen et al., 2012). Early studies showed that coexpression of adjuvants with an antigen improved both systemic and mucosal immune responses (Fredriksen et al., 2012). Examples of adjuvants are the cytokines IL-12 and IL-6, and the outer membrane protein Invasin (Fredriksen et al., 2012).

There are many studies where LPXTG anchors have been successfully exploited for anchoring and display of proteins on the surface of LAB (Cortes-Perez et al., 2005; Fredriksen et al., 2010;

Kajikawa et al., 2011; Minic et al., 2015). On the other hand, the use of lipoprotein anchors for display of heterologous proteins have received little attention compared to other anchoring methods (Michon et al., 2016). The same also appears to be the case for N-terminal transmembrane anchors. However, there are some studies worth mentioning in which N-terminal transmembrane proteins have been used as anchors for surface display. For example, the poly-γ-glutamic acid synthetase protein A (PgsA) from *Bacillus subtilis* is a transmembrane protein that has been used as an anchor in several studies.

In a study published in 2006, the PgsA protein was used to anchor and display two segments of the spike (S) protein from SARS-associated coronavirus on the cell surface of *L. casei.* When the recombinant *L. casei* strains was given to mice through nasal and oral administration, it resulted in a high production of antibodies that had a neutralizing activity against the SARS-associated coronavirus. These results indicated that the recombinant *L. casei* could potentially be used as a mucosal vaccine against this virus (Lee et al., 2006).

In 2012, Fredriksen et al. managed to display the virulence factor Invasin on the cell surface of *L. plantarum* WCFS1 using four different anchors, including an N-terminal transmembrane anchor. This anchor was made by coupling Invasin to a C-terminally truncated version of the penicillin binding protein 2B (Lp_1568). Of the three other anchors used, two were lipoprotein anchors (Lp_1261 and Lp_1452) and the last one was a LysM anchor (Lp_3014). All these anchors led to surface display of Invasin. In addition, several of the strains harbouring these anchors also activated the transcription factor NF-κB in human monocytes. NF-κB regulates many different proinflammatory genes (Fredriksen et al., 2012). A more detailed description of Invasin is given below.

1.7.1 Invasin

Many pathogens cause infection by binding and entering eukaryotic cells. An example is the Gram-negative bacterium *Yersinia pseudotuberculosis*, which can cause gastroenteritis in humans (Hamburger et al., 1999). This bacterium expresses an outer membrane protein called Invasin, which is a virulence factor (Fredriksen et al., 2012). This protein has total length of 986 residues (Fredriksen et al., 2012; Hamburger et al., 1999). It consists of an intracellular region which is a transmembrane domain containing approximately 500 residues. The remaining 497 residues in the C-terminal part forms the extracellular region of Invasin. The crystal structure of this region was determined by (Hamburger et al., 1999) (Fig. 1.7). It was discovered that it had a rod-like structure and consisted of five different domains, which was

referred to as D1, D2, D3, D4 and D5. The first four domains (D1- D4) are mainly composed of β-sheets, while the D5 domain contains both α-helices and β-sheets.

For *Y. pseudotuberculosis*, M-cells are an important "portal" that it can use to invade the body. This is because the extracellular region of Invasin that is expressed on the cell surface of this bacterium promotes its uptake by M-cells. The mechanism behind this is that the apical membrane of the M-cells express β_1-integrin, which functions as a receptor for Invasin (Hamburger et al., 1999) (Fig 1.8). It has been shown that it is the D4 and D5 domains that is involved in the binding to β_1-integrin. Three of the residues that are believed to be required for integrin binding are indicated in figure 1.7 (D_{911}, D_{811} and R_{883}). These are an aspartate (D) in D5, and an aspartate (D) and an arginine (R) in D4 (Hamburger et al., 1999).

When Invasin that is expressed on the surface of *Y. pseudotuberculosis* binds to β_1-integrin, the bacterium is taken up by the M-cell (Fig. 1.8). The bacterium will then be transcytosed through the M-cell and released in the underlying Peyer's patch, where it can cause an infection. From the Peyer's patch, the bacteria can travel to the mesenteric lymph node and end up in other parts of the body, including the spleen and the liver (Clark et al., 1998).

The ability of Invasin to promote internalization into M-cells makes this protein interesting to use as an adjuvant in a mucosal vaccine. This can for example be done by co-expressing Invasin with an antigen on the cell surface of a bacterium that is used as a vaccine delivery vehicle. Importantly, when Invasin is used as an adjuvant, it does not in itself induce a stronger immune response, but it targets the antigen-expressing bacterium to M-cells.

As mentioned, (Fredriksen et al., 2012) managed to express Invasin on the cell surface of *L. plantarum*. This is a very promising result since it can make it possible to co-express vaccine antigens with Invasin on the surface of *L. plantarum* that is used as a vaccine delivery vehicle. Ideally, this may increase the ability of the recombinant *L. plantarum* to gain access to the Peyer's patches via M-cells. Once inside the Peyer's patches, the antigens

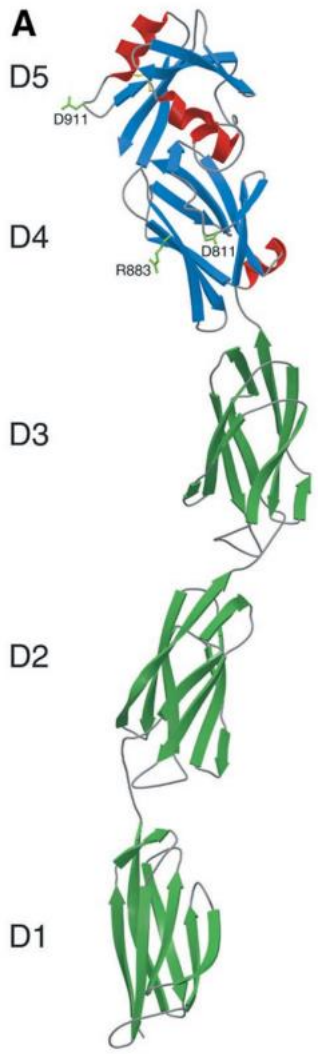


Figure 1.7. Ribbon diagram showing the structure of the extracellular region of invasin. This figure shows the crystal structure of the extracellular region of Invasin, determined by Hamburger et al. (1999). The figure is taken from (Hamburger et al., 1999). Reprinted with permission from AAAS (American Association for the Advancement of Science).

expressed on the surface of *L. plantarum* can be detected by immune cells that potentially can induce an immune response towards the antigen. In this way, Invasin can enhance the effect of the mucosal vaccine that uses *L. plantarum* as the delivery vehicle.

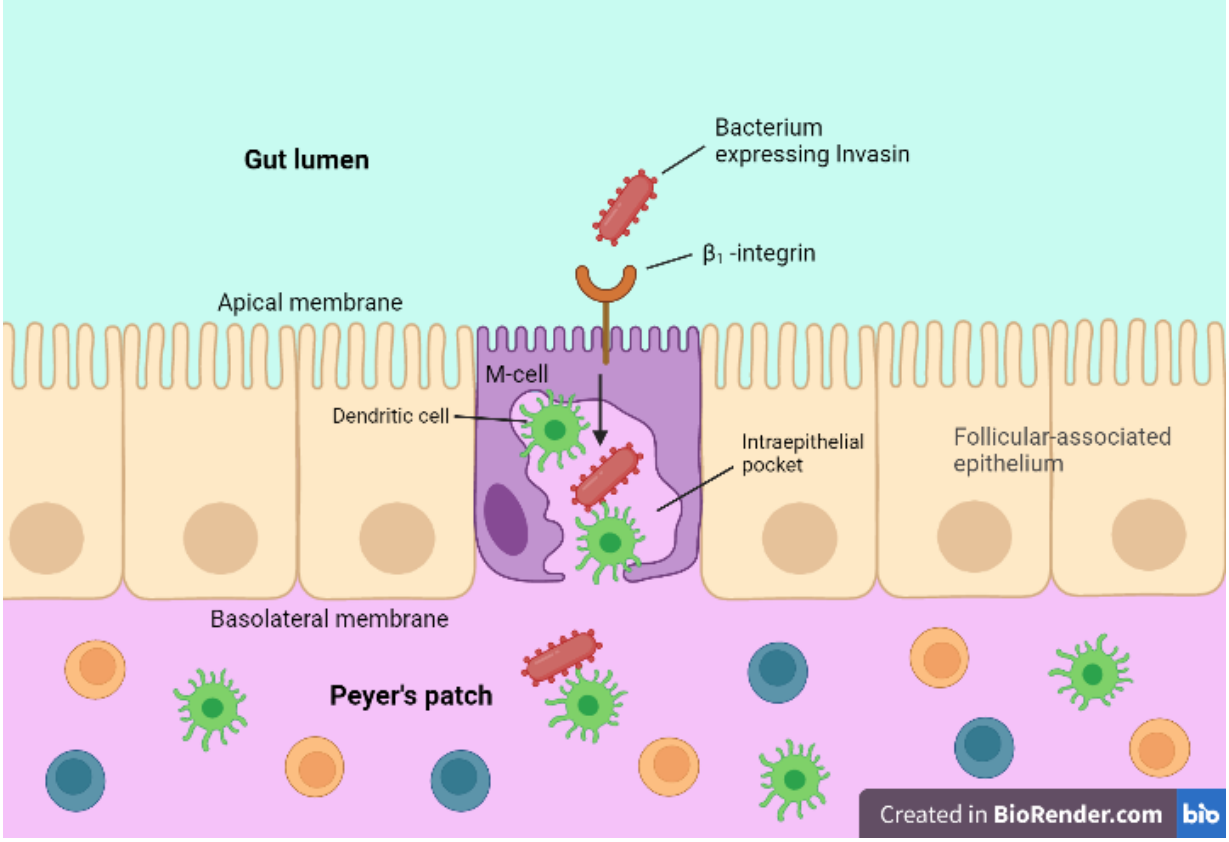


Figure 1.8. Transcytosis of an Invasin-expressing bacterium through an M-cell. An Invasin-expressing bacterium can be taken up by an M-cell in the follicular-associated epithelium in the small intestine. Invasin interacts with its receptor**, β_1-integrin**, on the apical membrane of the M-cell (Hamburger et al., 1999). This leads to a process called **transcytosis**, in which the bacterium expressing Invasin is taken up by the M-cell and transported in a membrane vesicle from the gut lumen and released in the intraepithelial pocket of the M-cell. The bacterium is detected by dendritic cells (DC's) that reside in the intraepithelial pocket and the Peyer's patch. The DC's can phagocytose the bacterium and present the antigens to naive T cells (orange) and B-cells (blue). The presentation of the antigen leads to activation of the naive lymphocytes, which divide and differentiate into effector cells(Parham, 2015). Created with BioRender.com.

The Caco-2 cell line is an immortalized line of human epithelial cancer cells that have become a valuable *in vitro* model of the intestinal epithelium. These cells can for example be used to study the absorption rate of drug candidates across the intestinal epithelial cell layer (Lakshmana Rao & Sankar, 2009). Non-polarized, undifferentiated Caco-2 cells express β_1-integrins, the receptors for Invasin. These cells can therefore be used as a model to study internalization of Invasin-expressing-bacteria (Critchley et al., 2004; Solberg, 2015). (Kernéis

et al., 1997) co-cultured polarized, differentiated Caco-2 cells with lymphocytes from Peyer's patches. This led to conversion of the Caco-2 cells into cells with M-cell characteristics. It was found that these "M-cell-like cells" had the ability to transcytose *Vibrio cholerae* from the apical surface to the basal compartment more efficiently than monocultures of Caco-2 cells. These M-cell like cells can therefore be used as a model to study in vitro transcytosis of bacteria across M-cells (Kernéis et al., 1997; Solberg, 2015).

1.8 Aim of this study

This thesis is part of a larger research project where the main goal is to develop a LAB-based mucosal vaccine for humans against tuberculosis. *Lactiplantibacillus plantarum* WCFS1, which is generally recognized as safe (GRAS), is used as the delivery vehicle of the vaccine. Invasin is a virulence factor from *Yersinia pseudotuberculosis* that can function as an adjuvant when used in a mucosal vaccine. This protein was therefore selected as a potential adjuvant in this study. The aim of this thesis was to anchor and display Invasin on the cell surface of *L. plantarum* by exploiting homologous N-terminal transmembrane proteins.

In-silico analyses of predicted N-terminal transmembrane proteins in *L. plantarum* WCFS1 were performed. Based on the results of these analyses, approximately five selected proteins were going to be used to construct anchors for cell surface display of Invasin in *L. plantarum*. The recombinant bacteria were characterized through growth curve analyses, Western blot, and flow cytometry. The results of these experiments were used to evaluate which of the anchors that were most well-suited for cell surface display of Invasin. The genes encoding Invasin, and the protein anchors were expressed using the pSIP-system for inducible gene expression.

The experimental work of this study was divided into the following main steps:

- *In silico* analysis of predicted N-terminal transmembrane proteins in the genome of *L. plantarum* to obtain knowledge of their properties and their functionality as anchors for Invasin. Selection of five candidate proteins to be used as anchors for Invasin
- Construction of vectors for inducible expression of N-terminally anchored Invasin using the pSIP-gene expression system.
- Transformation of the constructed expression plasmids into electrocompetent *L. plantarum*.
- Characterization of the recombinant *L. plantarum* strains through growth analysis in various media.
- Analysis of production of Invasin using Western blot analysis.
- Analysis of production and surface display of Invasin using flow cytometry.

2 Materials

2.1 Laboratory equipment, instruments, and software

Laboratory equipment	Supplier
CellStar® tubes, 15 and 50 mL	Greiner Bio-One
Cryovials, 1.5 mL	Sarstedt
Disposable cuvette, 1.5 mL	Brand
Electroporation cuvette, Gene Pulser®, 0.2 cm	Bio-Rad
Eppendorf tubes, 1.5 and 2.0 mL	Axygen
Borosilicate glassware	VWR
Falcon 2059 Polypropylene, round bottom tube, 14 mL	Fisher Scientific
FastPrep® tubes and lids	MP Biomedicals
Glass beads, acid-washed	Sigma
Microwell plate, 96 wells	Thermo Scientific
Parafilm	CURWOOD
PCR tubes, 0.2 mL	Axygen
Petri dishes	Heger
Plastic beakers, 50 to 1000 mL	VWR
Pipetboy comfort	Integra
Sealing tape	Thermo Scientific
Serological pipette 5, 10 and 25 mL	Sarstedt
Sterile filter, 0.22 µM pore size	Sarstedt
Syringes, 10-60 mL	Plastipak
Volumentric sylinders	VWR
Volumetric flasks	Duran
Waterbath	Fisher Scientific
Qubit assay tubes	Invitrogen

Instrument	Supplier
Centrifuge	
Allegra X-30R Centrifuge	Beckman Coulter
Eppendorf centrifuge 5418R	Eppendorf
Galaxy 14D Digital Microcentrifuge	VWR
Microcentrifuge Ministar Silverline	VWR
eBlot® L1 Protein Transfer System	GenScript
Electrophoresis Power Supply	
PowerPac Basic	Bio-Rad
PowerPac 300	Bio-Rad
Electroporation devices	
Bio-Rad Gene Pulser ® II	
Bio-Rad Pulse controller plus	
iBlot® Gel Transfer Device	Invitrogen
Incubators	
Innova® 44 Incubator Shaker Series	New Brunswick Scientific
Termaks Incubator	Termaks
Other instruments	
913 pH meter	Metrohm
CertoClav Sterilizer	CertoClav
FastPrep® -24 Tissue and Cell Homogenizer	MP Biomedicals
MacsQuant® Analyser	Miltenyi Biotec
Multiskan™ FC Microplate photometer	ThermoFisher
Varioskan Lux	ThermoFisher
PCR machines	
SimpliAmp Thermal Cycler	Applied Biosystems
SensoQuest Labcycler	SensoQuest
Photo- and optical equipment	
Azure c400	Azure biosystems
GelDoc EZ Imager	Bio-Rad
Transilluminator, UV-light source	VWR

Qubit™ Fluorometer	Invitrogen
Ultrospec® 10 Cell Density Meter	Biochrom US
Pipettes	
Finnpipette® F2	Thermo Scientific
Pipetboy comfort	Integra
SNAP i.d. 2.0 Protein Detection System	Millipore
Shaking or mixing devices	
IKA® MS3 Basic	IKA
IKA® RCT Classic	IKA
Water baths	
Isotemp® GPD 05	Fisher scientific
SBB Aqua 5 Plus	Grant

Software	Supplier
AzureSpot Analysis	Azure Biosystems
Biorender	https://biorender.com/
BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi
CLC Main DNA Workbench 7	Qiagen
FlowJo_v10.7.2	https://www.flowjo.com/
ImageLab	Bio-Rad
LocateP v2	(Zhou et al., 2008)
MACSQuantify™ software	Miltenyi biotec
Microbesonline	http://www.microbesonline.org/
NEBCloner®	https://nebcloner.neb.com/#!/redigest
pDraw32	https://www.acaclone.com/
Pfam	http://pfam.xfam.org/
Rstudio	https://www.rstudio.com/
SignalP-5.0	https://services.healthtech.dtu.dk/service.php?SignalP-5.0
SKANIT Software 2.5.1	Thermo Scientific
TMHMM v. 2.0	https://services.healthtech.dtu.dk/service.php?TMHMM-2.0

UniProt	https://www.uniprot.org/
WebLogo	https://weblogo.berkeley.edu/logo.cgi

2.2 Chemicals

Chemical	Supplier
Acetone, C_3H_6O	Merck
Agar powder	VWR
Brain-Heart Infusion (BHI)	Oxoid
De Man, Rogosa and Sharpe (MRS)	Oxoid
Dithiothreitol (DTT), $C_4H_{10}O_2S_2$	Sigma-Aldrich
DNA Gel Loading Dye 6X	NEB
Erythromycin, $C_{37}H_{67}NO_{13}$	Sigma-Aldrich
Ethanol, C_2H_5OH	Sigma-Aldrich
Glycerol, $C_3H_8O_3$	Merck
Glycine, $C_2H_5NO_2$	Duchefa Biochemie
Hydrochloric acid, HCl	Sigma
Polyethylene glycol, PEG_{1450}	Aldrich
Potassium chloride, KCl	Merck
Potassium phosphate monobasic, KH_2PO_4	VWR
SeaKem® LE Agarose	Lonza
Sodium chloride, NaCl	Merck
Sodium deoxycholate	Sigma-Aldrich
Sodium hydroxide, NaOH	Sigma
Sodium Monohydrogen Phosphate, Na_2HPO_4	Merck
Super Optimal Broth with Catabolite Repression (S.O.C.)	Invitrogen
Trichloroacetic acid (TCA), $C_2HCl_3O_2$	Sigma
Tris-base, $C_4H_{11}NO_3$	Sigma
Tween-20	Sigma-Aldrich

2.3 DNA, proteins, and enzymes

Components	Supplier
5X In-Fusion HD Enzyme Premix	Takara Bio
Antibodies	
Anti-Invasin, bleed # 3	ProSci Incorporated
Anti-Rabbit IgG (whole molecule)-FITC	Sigma
Goat Anti-Rabbit IgG (H+L), HRP	Invitrogen
The BenchMark™ Protein Ladder	Invitrogen
Bovine Serum Albumin (BSA)	Sigma
GeneRuler™ 1 kb DNA ladder	Fermentas
Inducer Peptide SppIP	CASLO
MagicMark® XP Western Protein Standard	Invitrogen
peqGREEN	peqlab
Restriction enzymes and buffers	
NdeI	NEB
SalI-HF®	NEB
NEB Buffer 2.1 (10X)	NEB
NEB Buffer 3.1 (10X)	NEB
NEB CutSmart (10X)	NEB
Polymerase Master Mix	
Q5® Hot Start High-Fidelity 2X Master Mix	NEB
VWR RedTaq DNA Polymerase Master Mix	VWR

2.4 Primers

Table 2.1: List of primers used in this study. Restriction sites are underlined.

Primer name	Sequence
Lp_1413F	GGAGTATGATTCATATGCAAAAAAATGGTTTTTGGGCCACGAT
Lp_1413R	CGGTGACGCTGTCGACAATTACCGCGTTCTGTAAGTTTTTT

Lp_1751F	GGAGTATGATTCATATGGCAGGTAACAATGAAC
Lp_1751R	CGGTGACGCTGTCGACGACGACGGCATTCTTAAGT
Lp_2341F	GGAGTATGATTCATATGAGTTGTCAAAACTGTGGT
Lp_2341R	CGGTGACGCTGTCGACTTGACCAGCCAACCCATT
Lp_0424F	GGAGTATGATTCATATGAGAAAAGACCTGCTCGAAA
Lp_0424R	CGGTGACGCTGTCGACAAAAACGTTAGTATAAACTAATAGCT
Lp_2132F	GGAGTATGATTCATATGACTAAGGGCCGGGAGTTGTTAAA
Lp_2132R	CGGTGACGCTGTCGACCACGTCATAGTGCGGATTCAATT
Lp_1576F	GGAGTATGATTCATATGGTTTTAGGTGCATTGAGC
Lp_1576R	CGGTGACGCTGTCGACGTTCATCCGAACTTTCAAACT
Lp_2341F (Short)	GGAGTATGATTCATATGAGGCCACAAACACTCCTT
Lp_1751F (Short)	GGAGTATGATTCATATGTGGTTTCGGCGGATTATTTTAT
Lp_1576F (Short)	GGAGTATGATTCATATGTTGAGCGGCATGGCGTAT
SekF	GGCTTTTATAATATGAGATAATGCCGAC

Table 2.2: Description of primers used in this study

Name	Description
Lp_1413F	Forward primer used for amplification of a fragment of the Lp_1413 gene from *L. plantarum* WCFS1.
Lp_1413R	Reverse primer used for amplification of a fragment of the Lp_1413 gene from *L. plantarum* WCFS1.
Lp_1751F	Forward primer used for amplification of a fragment of the Lp_1751 gene from *L. plantarum* WCFS1.
Lp_1751R	Reverse primer used for amplification of two fragments of the Lp_1751 gene from *L. plantarum* WCFS1.
Lp_2341F	Forward primer used for amplification of a fragment of the Lp_2341 gene from *L. plantarum* WCFS1.
Lp_2341R	Reverse primer used for amplification of two fragments of the Lp_2341 gene from *L. plantarum* WCFS1.
Lp_0424F	Forward primer used for amplification of a fragment of the Lp_0424 gene from *L. plantarum* WCFS1.
Lp_0424R	Reverse primer used for amplification of a fragment of the Lp_0424 gene from *L. plantarum* WCFS1.

Lp_2132F	Forward primer used for amplification of a fragment of the Lp_2132 gene from *L. plantarum* WCFS1.
Lp_2132R	Reverse primer used for amplification of a fragment of the Lp_2132 gene from *L. plantarum* WCFS1.
Lp_1576F	Forward primer used for amplification of a fragment of the Lp_2132 gene from *L. plantarum* WCFS1.
Lp_1576R	Reverse primer used for amplification of two fragments of the Lp_1576 gene from *L. plantarum* WCFS1.
Lp_2341F (Short)	Forward primer used for amplification of a fragment of the Lp_2341 gene from *L. plantarum* WCFS1.
Lp_1751F (Short)	Forward primer used for amplification of a fragment of the Lp_1751 gene from *L. plantarum* WCFS1.
Lp_1576F (Short)	Forward primer used for amplification of a fragment of the Lp_1576 gene from *L. plantarum* WCFS1.
SeqF	Forward primer for sequencing of all pSIP derivates.

2.5 Bacterial strains and plasmids

Table 2.3. Bacterial strains used in this study

Strain	Source
***Escherichia coli* TOP10**	Invitrogen
***Escherichia coli* Stellar**	Takara Bio
***Lactiplantibacillus plantarum* WCFS1**	(Kleerebezem et al., 2003)

Table 2.4. Plasmids used in this study

Plasmid name	Description	Source
pEV	pSIP401 derivative (Sørvig et al., 2003) without any target genes, called “empty vector”. Used as a negative control.	(Fredriksen et al., 2012)
pLp_1261Inv	pSIP401 derivative where the sequence corresponding to the lipoanchor of Lp_1261 is fused to the extracellular domain of the Invasin gene. Used in this study as a vector for the development of new plasmids containing N-terminal transmembrane anchors of Invasin.	(Fredriksen et al., 2012)
pLp_1452Inv	pSIP401 derivative where the sequence corresponding to the lipoanchor of Lp_1452 is fused to the extracellular domain of the Invasin gene. Used as a control in this study.	(Fredriksen et al., 2012)
pLp_1568InvS	pSIP401 derivative where the N-terminal transmembrane anchor Lp_1568 is fused to the extracellular domains D4 and D5 of Invasin. Used as a control in this study.	(Fredriksen et al., 2012)

pLp_1413Inv	pSIP401 derivative for cell surface expression of Invasin. The N-terminal transmembrane anchor Lp_1413 is fused to the extracellular domain of the Invasin gene.	This study
pLp_2341Inv	pSIP401 derivative for cell surface expression of Invasin. The N-terminal transmembrane anchor Lp_2341 is fused to the extracellular domain of the Invasin gene.	This study
pLp_1751 Inv	pSIP401 derivative for cell surface expression of Invasin. the N-terminal transmembrane anchor Lp_1751 is fused to the extracellular domain of the Invasin gene.	This study
pLp_0424 Inv	pSIP401 derivative for cell surface expression of Invasin. The N-terminal transmembrane anchor Lp_0424 is fused to the extracellular domain of the Invasin gene.	This study
pLp_2132 Inv	pSIP401 derivative for cell surface expression of Invasin. The N-terminal transmembrane anchor Lp_2132 is fused to the extracellular domain of the Invasin gene.	This study
pLp_1576 Inv	pSIP401 derivative for cell surface expression of Invasin. The N-terminal transmembrane anchor Lp_2341 is fused to the extracellular domain of the Invasin gene.	This study
pLp_2341Short Inv	pSIP401 derivative for cell surface expression of Invasin. A shortened version of the N-terminal transmembrane anchor Lp_2341, lacking the intracellular tail, is attached to the extracellular domain of the Invasin gene.	This study
pLp_1751Short Inv	pSIP401 derivative for cell surface expression of Invasin. A shortened version of the N-terminal transmembrane anchor Lp_1751, lacking the intracellular tail, is attached to the extracellular domain of the Invasin gene.	This study
pLp_1576Short Inv	pSIP401 derivative for cell surface expression of Invasin. A shortened version of the N-terminal transmembrane anchor Lp_1576, lacking the intracellular tail, is attached to the extracellular domain of the Invasin gene.	This study

2.6 Kits

Kits	Supplier
eBlot® L1 Protein Transfer system	GenScript
eBlot® L1 Transfer Sponge	
eBlot® L1 PVDF Equilibration Buffer 10X	
PVDF membrane	
Transfer cassette	
Blotting roller	
In-Fusion® HD Cloning Kit	Clontech
5X In-Fusion HD Enzyme Premix	
Novex® NuPAGE® SDS-PAGE Gel System	Invitrogen
NuPAGE® Novex Bis-Tris gels 8 cm x 8 cm x 1 mm, 10 and 15 wells	
NuPAGE® LDS Sample Buffer (4X)	
NuPAGE® Reducing Agent (10X)	
NucleoSpin® Gel and PCR Clean-Up	Macherey-Nagel
NucleoSpin® Gel and PCR Clean-Up columns	
Collection tubes, 2 ml	
Binding buffer NTI	
Wash buffer NT3	
Elution buffer NE	
NucleoSpin® Microbial DNA	Macherey-Nagel
Lysis buffer MG	
Wash Buffer BW	
Wash Buffer B5	
Elution Buffer BE	
Liquid Proteinase K	
NucleoSpin® Bead Tubes Type B	

NucleoSpin® Microbial DNA Columns (light green rings)	
Collection Tubes, 2 mL	
NucleoSpin® Plasmid	Macherey-Nagel
Buffer A1, A2, A3 and A4	
Buffer AW	
Elution Buffer AE	
NucleoSpin ® Plasmid Columns	
Collection Tubes, 2 mL	
SNAP i.d. ® 2.0 Protein Detection System	Millipore
SNAP i.d. ® 2.0 Mini Blot Holder	
SNAP i.d. ® Spacer	
SNAP i.d. ® Blot roller	
Filter paper	
SuperSignal™ West Pico PLUS Chemiluminescent Substrate	Thermo Scientific
Luminol/enhancer	
Stable Peroxide Buffer	
Q5 Hot Start High-Fidelity 2X Master Mix	NEB
Qubit™ dsDNA BS Assay Kit	Invitrogen
Qubit™ Assay Tubes	
Qubit™ dsDNA BR Buffer	
Qubit™ dsDNA BR Reagent	
Qubit™ dsDNA BR Standard 1 and 2	

2.7 Agars and media

All the components and suppliers are listed in section 2.2.

Brain-Heart Infusion (BHI) liquid medium:

18.5 g BHI was dissolved in 500 mL dH2O. Sterilized in CertoClav at 115 °C for 15 minutes.

BHI agar medium:
BHI liquid medium supplemented with 1.5 % (w/v) agar. After sterilization in a CertoClav, the medium was cooled down to ~ 60 °C before addition of the appropriate antibiotic. The medium was poured into sterile dishes and cooled. When the medium had solidified, the agar plates were stored at 4 °C.

De Man, Rogosa and Sharpe (MRS) liquid medium:

26 g MRS was dissolved in 500 mL dH2O. Sterilized in CertoClav at 115 °C for 15 minutes.

MRS agar medium:
MRS liquid medium supplemented with 1.5 % (w/v) agar. After sterilization in a CertoClav, the medium was cooled down to ~ 60 °C before addition of the appropriate antibiotic. The medium was poured into sterile dishes and cooled. When the medium had solidified, the agar plates were stored at 4 °C.

MRS + sucrose + magnesium chloride (MRSSM)
Premade laboratory stock

RPMI 1640 Medium (1X)
Prremade from manufacturer (Gibco™)

Super Optimal Broth with Catabolite Repression (S.O.C.)
Premade from manufacturer (Invitrogen)

2.8 Buffers and solutions

Phosphate-buffered saline (PBS)

The following salts were dissolved in 800 ml distilled water:

8 g NaCl

0.2 g KCl

1.44 g Na_2HPO_4

0.24 g KH_2PO_4

The pH was adjusted to 7.4 with HCl, and H_2O was added to 1 liter. The solution was sterilized at 121 °C for 15 min in a Certoclav.

Tris-HCl

121.1 g/l Tris base

121.1 g Tris base was added to 800 ml H_2O. pH was adjusted to 8.0 and H_2O was added to 1 liter. The solution was sterilized at 121 °C for 15 min in a CertoClav.

Tris Buffered Saline (TBS)

150 mM NaCl

10 mM Tris-HCl, pH 8

Sterile-filtrated and stored at room temperature

TTBS

TBS

0.1 % (w/v) Tween-20

Tris-acetate/EDTA (TAE) 50X

Pre-made laboratory stock

Tris-glycine-SDS (TGS) (10X)

Ready-made from BIO-RAD

3 Methods

3.1 Cultivation of bacteria

E.coli were grown in liquid Brain Heart Infusion (BHI) medium or on solid BHI-agar plates containing 200 µg/ml erythromycin. Overnight cultures were incubated at 37 °C with shaking, while agar plates were incubated on 37 °C without shaking.

L. plantarum WCFS1 were grown in liquid de Man, Rogosa and Sharpe medium (MRS) or on solid MRS containing 10 µg/ml erythromycin. Liquid overnight cultures and agar plates were incubated on 37 °C (or 30 °C) without shaking.

3.2 Long-term storage of bacteria

For long-term storage of bacteria, a glycerol stock was made. Glycerol prevents disruption of the bacterial cells at low temperatures.

Bacteria of relevant strains were grown overnight in 10 mL of the appropriate growth medium and supplemented by antibiotics. A glycerol stock was made by transferring 1 ml overnight culture and 300 µl glycerol to a 1.5 mL cryovial. The vial was inverted a 3-4 times to mix the content, and immediately placed in a -80 °C freezer for long-term storage.

When needed, overnight cultures of bacteria were made from glycerol stocks. A small amount of the frozen culture was picked from the glycerol stock using a sterile toothpick. The toothpick was then transferred to a culture tube containing the appropriate growth medium and antibiotics.

3.3 Isolation of genomic DNA

For amplification of specific genes from *L. plantarum* WCFS1, genomic DNA from this strain was isolated. The bacterial cells were lyzed using the FastPrep®-24 Tissue and Cell Homogenizer. The instrument was set to 6.5 m/s for 30 seconds. This was performed three times, with a 5-minute rest between each run where the samples were placed on ice.

The genomic DNA from the disrupted cells was isolated using the NucleoSpin® Microbial DNA kit from Macherey-Nagel. The manufacturers standard protocol for Gram-positive bacteria was followed.

3.4 Plasmid isolation from bacteria

The NucleoSpin® Plasmid Kit (Macherey-Nagel) was used to isolate plasmid DNA from overnight cultures of bacteria. The NucleoSpin® Plasmid Kit protocol 5.1 for high-copy plasmids was always followed.

3.5 Plasmid digestion with restriction enzymes

Restriction enzymes are enzymes that digest DNA at specific sites. Each type of restriction enzyme recognizes one specific sequence of DNA, often a palindromic sequence. If the two cuts are made on the same position in the DNA, a blunt end is created. The two cuts can also be made with a few base pairs of distance in between, giving rise to ends with overhang. Enzyme digestion is often used to prepare DNA fragments for ligation. In these cases, it is common to use two restriction enzymes to cut DNA at two different sites, known as double digestion. The enzymes are used to digest both the insert and the vector, generating complementary ends that enables ligation of the fragments.

Materials

Plasmid DNA

Restriction enzymes

Compatible buffer

dH_2O

Procedure

1. The components for restriction enzyme digestion were added in the order shown in the table 3.1.

Table 3.1. Example of double digestion using restriction enzymes and buffers from New England Biolabs (NEB):

Component	Volume (µL)
dH_2O	to 50
DNA	<1 µg (variable)
10 X NEB buffer*	5
NdeI	1.5
Sal-HF	1.5

*Buffers that were compatible with the restriction enzymes used, were found at NEBCloner®: https://nebcloner.neb.com/#!/redigest

2. The reaction was gently mixed before it was incubated at 37 °C for approximately 2 hours.
3. The mixture was loaded onto an agarose gel for separation of the DNA fragments (Section 3.8).

3.6 Determination of DNA concentration

After isolation of plasmid DNA from *E. coli*, the DNA concentration was determined using the Qubit™ Fluorometer and the Qubit® dsDNA BR Assay Kit (Invitrogen). The Qubit™ Fluorometer was also used to determine DNA concentration after purification of both genomic DNA and plasmid DNA from agarose gels.

Materials

Qubit™ dsDNA BS Assay Kit (Invitrogen)

Qubit™ Fluorometer

Procedure

1. A master mix was made by diluting Qubit™ dsDNA BR Reagent in the Qubit™ dsDNA BR Buffer according to manufacturer's protocol. The samples to be measured were kept on ice.
2. 1-3 µl of each DNA sample was mixed with 197-199 µl of the master mix in a Qubit™ Assay Tube. The DNA-concentration was determined using the Qubit™ Fluorometer.

3.7 Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR) is used for *in vitro* amplification of DNA fragments of interest. The first step is to increase the temperature to 94-98 °C to separate the two strands of double-stranded DNA. The single-stranded DNA fragments serve as template DNA for a thermostable DNA polymerase, which synthesize a new, complementary DNA strand from

each single-stranded DNA fragment. This gives rise to double-stranded DNA that is an exact copy of the original DNA fragment that was initially added to the reaction mixture. To begin the synthesis of a new strand, DNA polymerase requires an oligonucleotide (primer) that is complementary and can hybridize to a stretch of nucleotides on the template DNA. This gives DNA polymerase a starting point for extension of a new strand by incorporation of deoxynucleotides (dNTPs) from the solution.

The three steps of the PCR cycle are typically repeated 25-35 times. This leads to an exponential increase in the number of copies of the specific DNA fragment.

3.7.1 PCR using Q5® High-Fidelity 2X Master Mix

The DNA fragments to be ligated into a linearized vector using In-Fusion cloning (Section 3.10) were first amplified from the genomic DNA of *L. plantarum* WCFS1. To amplify fragments of genomic DNA, the Q5™ Hot Start High-Fidelity 2X Master Mix was used. The master mix contains all necessary components for the PCR reaction, apart from the template DNA and primers. PCR reaction was performed according to the descriptions from the manufacturer (Table 3.2).

Procedure

1. The components were added to a sterile PCR tube (0.2 ml) as shown in table 3.2. The tube was placed on ice until the PCR reaction was ready to be started.

Table 3.2. The components of a Q5 PCR reaction

Component	Volume (µl)	Final concentration
Q5™ Hot Start High-Fidelity 2X Master Mix	25	1X
10 µM Forward Primer	2.5	0.5 µM
10 µM Reverse Primer	2.5	0.5 µM
Template DNA	variable	<1 µg
Nuclease-free water/dH_2O	to 50	

2. The reaction was gently mixed, before the PCR tube was placed in the PCR instrument. .
3. The correct program and settings were chosen before starting the PCR reaction. Adjust the settings to the program shown in table 3.3.

Table 3.3. The Q5 Cycling Program

Step	Time	Temperature
Initial denaturation	30 seconds	98 °C
Denaturation	5-10 seconds	98 °C
Annealing	10-30 seconds	*50-72 °C
Extension/Elongation	20-30 seconds/ kb**	72 °C
final extension/elongation	2 minutes	72 °C
Hold		4-10 °C

*The annealing temperature depends on the melting temperatures (T_m) of the primers.

**The elongation time depends on the length of the DNA fragment.

3.7.2 Colony PCR using VWR RedTaq DNA Polymerase Master Mix

After transformation, colonies were analysed with PCR to determine whether the transformation procedure had been successful. The VWR RedTaq DNA Polymerase Master Mix was used for colony PCR, and the PCR was performed according to the protocol from the manufacturer (VWR). For colony-PCR of *L. plantarum*, which is Gram-positive bacteria, the PCR tube was heated in a microwave oven for 1 minute on full effect to break down the peptidoglycan cell wall. This was done after addition of the colony to the tube, but prior to addition of the remaining PCR ingredients.

<u>Procedure</u>

1. The colony was transferred from the agar plate to a sterile PCR-tube using a sterile toothpick.
2. The tube was placed on ice and the remaining reactants for the PCR reaction was added in the order shown in table 3.3.

Table 3.3. The components of the Red Taq PCR

Components	Volume (µl)	Final concentration
Taq 2X Master Mix	25	1X
Forward primer 10 µM	1	0.2 µM
Reverse primer 10 µM	1	0.2 µM
dH_2O	23	
Template DNA	Variable	-

3. The instrument was adjusted to the settings shown in table 3.4, and the tube was placed in the PCR instrument.

Table 3.4. The Red Taq Polymerase PCR program

Step	Cycles	Time	Temperature
Initial denaturation	1	2 minutes	95 °C
Denaturation	25-35	20-30 seconds	95 °C
Annealing	25-35	20-40 seconds	50-65 °C*
Elongation	25-35	30 seconds	72 °C
Final elongation	1	5 minutes	72 °C

*The annealing temperature depends on the melting temperatures (T_m) of the primers.

3.8 Agarose gel electrophoresis

The principle of agarose gel electrophoresis is that DNA fragments are separated according to their size. The agarose gel is built up of a matrix of polysaccharides that limits the movement of nucleic acids. DNA is negatively charged and will therefore migrate from the negative pole to the positive pole when an electric field is applied. All DNA molecules have an equal amount of negative charge per mass. Therefore, the separation of DNA molecules on gel electrophoresis is only dependent on their size. The largest DNA fragments will migrate the shortest distances in the gel because they are stopped by the matrix of polysaccharides.

Gels with 1.2 % agarose and was used to separate DNA fragments larger than 200 bp. A DNA ladder with fragments of known sizes was added to the gels to determine the fragments sizes in the samples.

Materials

SeaKem® LE Agarose

1 x TAE buffer

peqGREEN

Loading dye

1 kb DNA ladder

Procedure

1. 6 g of agarose was dissolved in 500 ml 1x TAE buffer to make a 1.2% agarose stock solution.
2. The solution was sterilized for 15 minutes at 115 °C in a CertoClav. The stock solution was stored at 60 C until use.
3. One agarose gel was prepared by pouring 55-60 ml of the 500 mL stock solution into a beaker and adding 2.5 µl peqGREEN. The solution was mixed and transferred to a moulding tray with combs.
4. After solidifying of the gel, the combs were removed and transferred to the electrophoresis chamber.
5. Prior to electrophoresis, loading dye was added to the relevant samples. The samples were loaded into wells in the gel. In addition, 10-15 µl of a 1 kb ladder was added to one of the wells.
6. The samples were run on 90 V. The time of the run depended on the sizes of the DNA fragments.
7. After completion of electrophoresis, the gels were photographed using GelDoc EZ Imager.

3.9 Extraction and purification of DNA from agarose gels

DNA fragments detected on gel electrophoresis could, if desired, be excised from the gel and purified for further analyses. The DNA in the gel pieces was extracted and purified using the kit NucleoSpin® Gel and PCR Clean-Up (Macherey-Nagel), following the manufacturers standard protocol for DNA extraction from agarose gels.

Both the enzymatically linearized vector and the PCR-amplified insert were purified after gel electrophoresis. The insert could then be cloned into the vector using In-Fusion cloning (Section 3.10).

3.10 DNA ligation using In-Fusion cloning

When both the linearized vector and the insert had been purified and extracted from agarose gels (Section 3.9) the insert could be cloned into the vector. This was achieved using the In-Fusion HD Cloning Kit. This kit contains enzymes that recognize the 15-bp overhangs at each end of the DNA-fragment to be inserted into the vector. These 15-bp overhangs are generated in the PCR reaction by using primers that have a 15-bp tail that is complementary to the linearized vector. The PCR-amplified insert and the linearized vector are added to the same reaction mix, which is incubated at 50 °C for 15 minutes. This cloning reaction can subsequently be used in a transformation procedure (Section 3.12).

Materials

5X In-Fusion HD Enzyme Premix

Linearized vector

Purified PCR product

dH_2O

Procedure

1. The components for the In-Fusion cloning reaction were added as shown in table 3.5 below, with a 2:1 molar ratio between the insert and vector. To determine the optimal amount of linearized vector and purified PCR product in each reaction, the In-Fusion® Molar Ratio Calculator was used: https://www.takarabio.com/learning-centers/cloning/primer-design-and-other-tools/in-fusion-molar-ratio-calculator

Table 3.5. Ingredients used for the In-Fusion cloning reaction

Ingredient	Volume (µl)
5X In-Fusion HD Enzyme Premix	2
Linearized vector	X
Purified PCR fragment (insert)	X
dH_2O (as needed)	X
Total volume	10

2. The reaction was incubated for 15 minutes at 50 °C, then placed on ice. The cloning reaction is now ready to be used in the transformation procedure (Section 3.12). Alternatively, the reaction was stored at -20 °C for transformation later.

3.11 DNA sequencing of plasmids and PCR products

In the cases where colony-PCR of transformed cells showed bands corresponding to the newly inserted gene on the plasmid, plasmid DNA was sent to Eurofins for Sanger sequencing to confirm that the sequence of the inserted fragment was correct.

Materials

DNA template (approximately 400-500 ng of plasmid DNA, or 100-400 ng of a PCR fragment)

25 pmol of the primer

dH_2O to a total volume of 11 µl

Procedure

The ingredients were added to an Eppendorf tube. The tube was labeled with a sticker containing a unique barcode and an ID-number for identification. After sequencing, the samples were analysed with CLC DNA Main Workbench 7.

3.12 Transformation of *E. coli*

In-Fusion cloning reactions were transformed into chemically competent *E. coli* cells. Two different strains of competent cells were used: OneShot® Top10 (Invitrogen) Chemically competent cells and Stellar Chemically (Takara Bio) competent cells.

For both types of competent cells, transformation was performed according to the protocol from the manufacturer .

3.13 Preparation of electrocompetent *Lactiplantibacillus plantarum* WCFS1

Constructed plasmids were transformed to electrocompetent *L. plantarum*. The cells to be made electrocompetent were grown in MRS containing glycine. During growth, glycine will gradually replace alanine in the bacterial cell wall. This makes the cell wall more permeable, which facilitates the uptake of foreign DNA through electroporation. Transformation was performed according to the protocol made by (Aukrust et al., 1995).

<u>Materials</u>

MRS

MRS + 1% glycine

20% glycine

30% PEG_{1450} (needs to be made fresh)

MRSSM (MRS + 0.5M sucrose + 0.1M $MgCl_2$)

Dry ice

<u>Procedure</u>

1. *L. plantarum* WCFS1 was cultured overnight in 10 ml MRS at 37 °C without shaking.
2. The next day (day 2), 1 ml of the overnight culture was used to make a serial dilution (10^{-1}-10^{-10}) in MRS + 1% glycine. The cultures were further incubated overnight at 37 °C.

3. On day 3, OD_{600} was measured for the cultures in the serial dilution to find a culture with an appropriate OD_{600} (2.5±0.5). The culture with appropriate OD_{600} was further diluted in pre-warmed 20 ml MRS + 1% glycine. The culture was then grown on 37 °C until it reached the logarithmic phase (OD_{600} of 0,7±0,07) (approximately 3 hours), and then placed on ice for 10 min.
4. The culture was centrifuged at 5 000 x g for 10 minutes at 4 °C, and the supernatant was discarded.
5. The pellet was resuspended in 5 ml ice cold 30% PEG_{1450}. Additional 20 ml of 30 % PEG_{1450} was added. The tube was gently inverted and placed on ice for 10 minutes.
6. The cells were collected by centrifugation for 5 000 x g for 10 minutes at 4 °C, and the supernatant was discarded.
7. The pellet was resuspended in 400 µl 30 % PEG_{1450} (2% of the harvested volume). Portions of 40 µl were pipetted into Eppendorf tubes, and the tubes were immediately placed at dry ice and further frozen at -80 °C.

3.14 Transformation of electrocompetent *Lactiplantibacillus plantarum* WCFS1

Electroporation of electrocompetent *L. plantarum* WCFS1 was performed according to the protocol described by (Aukrust et al., 1995). The principle of electroporation is that an electric pulse is used to disrupt the cell membrane of electrocompetent bacteria, enabling uptake of DNA, for example plasmids.

<u>Materials</u>

Bio-Rad Gene Pulser ® II

Bio-Rad Pulse controller plus

Electrocompetent *Lactiplantibacillus plantarum* WCFS1

Plasmid DNA

Gene Pulser® Electroporation cuvette 0.2 cm (chilled on ice)

MRSSM medium

MRS plates with 10 µg/ml erythromycin

Procedure

1. Prior to electroporation, the electroporation parameters were adjusted to the following settings:
    - Capacitance: 25 µF
    - Volt: 1,5 kV
    - Resistance: 400 Ω
2. For each transformation reaction, one tube with 40 µl electrocompetent cells was thawed on ice. 5 µl of plasmid DNA was added to the tube of electrocompetent cells, and the mixture was transferred to a chilled electroporation cuvette. The cuvette was carefully tapped a few times to mix the solution and remove any air bubbles.
3. The cuvette was placed in the electroporation instrument and given the electric pulse.
4. 450 µl of room-tempered MRSSM was added to the cuvette, and the cell suspension was transferred to a sterile Eppendorf tube.
5. The tube containing the transformed cells was incubated at 37 °C for 2-4 hours without shaking.
6. 120 µl of the cell suspension was spread out on an MRS agar plate containing 10 µg/ml erythromycin and incubated on 37 °C overnight, or for up to 48 hours.

3.15 Cultivation and harvesting of *Lactiplantibacillus plantarum* WCFS1

After electrocompetent *L. plantarum* WCFS1 cells have been successfully transformed, cell cultures of these bacterial cells can be cultivated and induced to initiate expression of the genes on the inserted plasmid. This is achieved by adding inducer peptide (SppIP) to the cultures. The induced cells were harvested for Western blot analysis (Section 3.19) and flow cytometry analysis (Section 3.21).

Materials

MRS

Spectrophotometer

Cuvettes

Inducer peptide, SppIP

PBS

Procedure

1. Recombinant *L. plantarum* WCFS1 were grown at 37 °C overnight in MRS medium containing 10 µg/ml erythromycin.
2. The cultures were diluted in 50 ml pre-warmed (37 °C) MRS to an OD_{600} of ~0.15.
3. The diluted cultures were incubated at 37 °C until OD_{600} was 0.28-0.33. At this point, the cultures were induced using 25 ng/ml SppIP. Straight after the addition of SppIP, the tubes were turned upside down once to mix the culture.
4. The induced cultures were incubated at 37 °C for three hours, then placed on ice to inhibit further growth.
5. Prior to centrifugation of the cell cultures, a small amount of cell culture of each recombinant strain was transferred to Eppendorf tubes for flow cytometry analysis (Section 3.21).
6. The remaining cultures were centrifuged at 5000 x g at 4 °C for 5 minutes and placed on ice again.
7. Approximately 1.5 ml of supernatant from each centrifugated cell culture was transferred to an Eppendorf tube and stored at -20 °C for western blot analysis (Section 3.19). Subsequently, the supernatant was decanted off before the cell pellet was washed 1-2 times with 5 ml cold PBS.
8. The suspension was centrifuged at 5 000 x g and 4 °C for 5 minutes.
9. The cell pellet was stored at -20 °C until further analysis.

3.16 Growth curve analysis of recombinant strains of *L. plantarum* WCFS1

The growth of recombinant strains could be analysed with a growth curve analysis. This was performed the same day as the procedure for cultivation and harvesting of *L. plantarum* WCFS1. A small amount of cell culture was transferred to a 96-well microtiter plate before and after induction. The growth was measured by a plate reader, either Multiskan™ FC Microplate photometer or Varioskan Lux.

Materials

96-well microtiter plate

Plastic to cover the microtiter plate

Plate reader

MRS

Procedure

1. Prior to induction, 200 µl of each cell culture were added to wells of the 96-well plate, in triplicates.
2. When the cell cultures reached an OD_{600} of 0.28-0.33, they were induced as described in section 3.15.
3. Straight after induction, 200 µl of each induced cell culture were added to the well in the 96-well plate in triplicates.
4. The 96-well plate was sealed with plastic and placed in the Multiskan™ FC Microplate photometer or Varioskan Lux. The plate reader was set to measure OD_{600} every 5 minutes for 24 hours.

3.17 Preparation of cell lysate using glass beads

Materials

FastPrep tubes

Glass beads

FastPrep® - 24 Tissue and Cell Homogenizer

PBS

Procedure

1. Bacterial cells were cultivated and harvested according to the description in section 3.15.
2. For each tube containing harvested bacteria, the cell pellet was resuspended in 1 ml PBS. The cell suspension was transferred to a FastPrep-tube containing approximately 0.4 g of glass beads.

3. The FastPrep-tube was placed in a FastPrep®-24 Tissue and Cell Homogenizer and shaken at 6,5 m/s for 45 seconds. This step was repeated 2-3 times to properly lyse the cells.
4. The tubes were placed on ice for 5 minutes between each run.
5. After the last run in the FastPrep®-24 Tissue and Cell Homogenizer, the Fast-Prep tubes were centrifuged at 16 000xg at 4 °C for 1 min. 800-900 µl of the supernatant was transferred to a new Eppendorf tube.
6. The Eppendorf tube was also centrifuged at 16 000 x g at 4 °C to remove remaining glass beads.
7. The protein extract was transferred to a new Eppendorf tube and stored at -20 °C.

3.18 Gel electrophoresis of proteins (SDS-PAGE)

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is a rapid and widely used method to separate proteins in a sample. The gel used in this method contains polyacrylamide, which forms a porous matrix. Prior to SDS-PAGE, a sample buffer, containing lithium dodecyl sulphate (LDS), and a reducing agent (Dithiothreitol, DTT), are added to the samples, and they are boiled for 10 minutes to denaturate the proteins. The sample buffer disrupts non-covalent interactions between different parts of the protein chain, while the reducing agent (DTT) breaks covalent disulfide bonds.

Sodium dodecyl sulfate (SDS) is an anionic detergent which makes the proteins denature, unfold, and gives them an evenly distributed, negative charge. The negative charge is proportional to its molecular mass and that masks the original charge of the protein. When running the gel, the proteins in the samples, which are now negatively charged due to SDS, will move through the pores in the gel matrix. The proteins migrate from the negative electrode to the positive electrode, and separate based on their molecular weight. SDS-PAGE is therefore an excellent method for determining the molecular weight of proteins, the degree of protein expression and sample purity (Nowakowski et al., 2014).

<u>Materials</u>

NuPAGE® Novex Bis-Tris gels 8 cm x 8 cm x 1 mm, 10 or 15 wells

NuPAGE® LDS Sample Buffer (4X)

NuPAGE® Reducing Agent (10X)

TGS-buffer

MagicMark® XP Western Protein Standard

Procedure

1. 7.5 µl LDS Sample Buffer, 3 µl Reducing Agent (Dithiothreitol (DTT)) and 20 µl protein sample were combined for a total volume of 30.5 µl.
2. The samples were incubated in a water bath with boiling water for 10 minutes.
3. The pre-made SDS-PAGE gel was placed in the electrophoresis chamber, and TGS-buffer was added.
4. The boiled samples, and the MagicMark® XP Western Protein Standard, were applied to the gel. The gel was run at 260 V for 20 minutes.
5. When the electrophoresis was finished, the gel was placed in dH_2O for a short time before further analyses. Alternatively, if blotting was going to be performed using the eBlot Protein Transfer system (Section 3.19.1), the gel was placed in 10% ethanol for 5-10 minutes for calibration before further analysis.

3.19 Western blot

Western blot is a procedure used to detect the presence of a protein of interest in a sample. An advantage with this method is that it can detect many different types of proteins, also proteins that are present in low concentrations (Kurien & Scofield, 2015).

The proteins are first separated on an SDS-PAGE gel. In the next step, called blotting, the proteins are transferred from the gel to a membrane, with the use of an electric current. Different types of membranes can be used, such as a nitrocellulose membrane or a PVDF membrane. After the protein transfer is completed, a *blocking solution* of bovine serum albumin (BSA) is added to the membrane to prevent unspecific interactions between the membrane and the antibody. First, the membrane is incubated with a primary antibody that hybridizes specifically to the protein of interest. Subsequently, the membrane is washed to remove excess primary antibody. This is followed by a new incubation with a secondary antibody that hybridizes specifically to the protein-primary antibody complex. The secondary antibody is linked to an enzyme that, in combination with a chemiluminescent reagent, emits

a signal if the protein of interest is present (Fig. 3.1). In this study, the enzyme conjugated to the secondary antibody was Horse Radish Peroxidase (HRP) and the chemiluminescent reagent was luminol. HRP catalyzes the oxidation of luminol by hydrogen peroxide. This oxidation releases detectable light, making the protein of interest visible if it is present (Fig. 3.1).

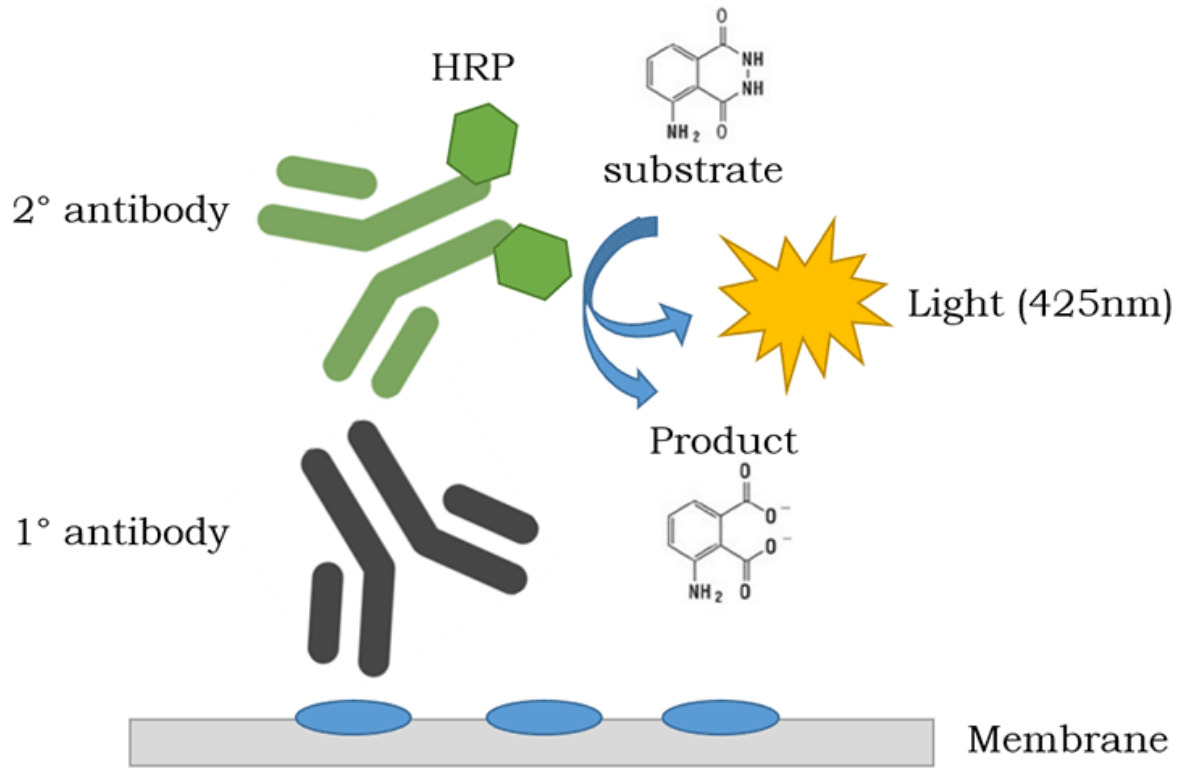


Figure 3.1. Chemiluminescent detection of proteins in Western blot. The primary antibody binds to its specific protein on the membrane (PVDF or nitrocellulose) and a secondary antibody conjugated to HRP binds to the primary antibody. HRP catalyzes the oxidation of luminol by hydrogen peroxide, resulting in the release of light. This enables visualization of the protein of interest. The figure is taken from ThermoFisher: https://www.thermofisher.com/no/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/chemiluminescent-western-blotting.html

3.19.1 Blotting with the eBlot® L1 Protein Transfer system

The eBlot® L1 Protein Transfer system is a rapid wet blotting system that was used to transfer proteins from an SDS-PAGE gel to a PVDF membrane using an electric current.

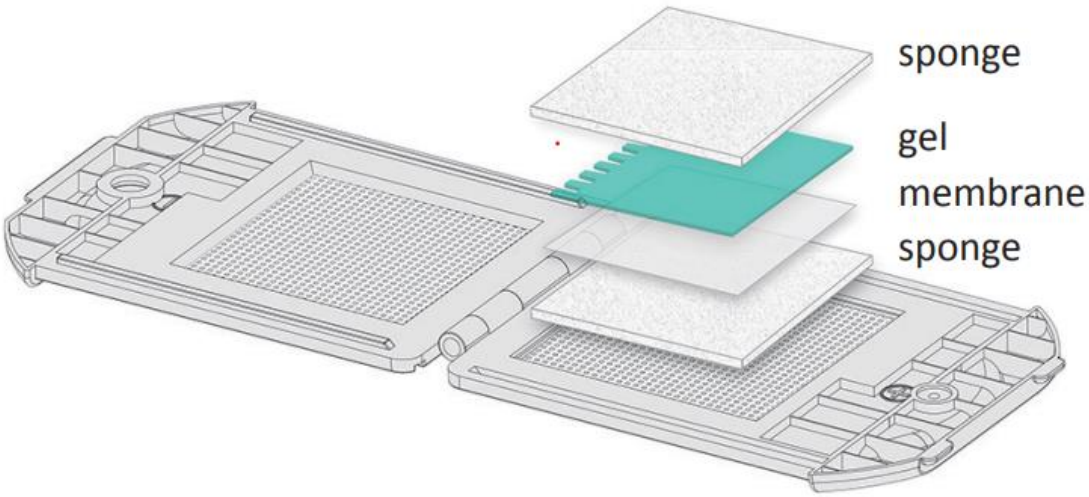


Figure 3.2. Assembly of the components in the eBlot® L1 Protein Transfer system. A transfer cassette was opened, and a piece of sponge was placed within the metal frame on the anode side of the cassette (marked with "+"). The calibrated membrane was placed on top of the sponge, and the gel was placed on top of the membrane. Another piece of sponge was placed on top of the gel before the cassette was closed and placed in a channel in the eBlot® Protein Transfer Instrument.

Materials

eBlot® L1 Protein Transfer System

100% ethanol

10 % ethanol

eBlot® L1 PVDF Membrane Equilibration buffer

dH_2O

eBlot® L1 Transfer sponges

Procedure

1. The gel from SDS-PAGE (Section 3.18) was equilibrated with 10% ethanol for 10 minutes before it was placed in distilled water.
2. A PVDF membrane was soaked in 100% ethanol to activate the membrane before it was soaked in approximately 10 ml of eBlot L1 PVDF Membrane Equilibration Buffer.
3. One transfer cassette was placed on a table. One piece of sponge was placed within the metal frame on the anode side of the transfer cassette (the side marked with "+").

4. The calibrated membrane was placed on top of the sponge, and the gel was placed on top of the membrane. A blotting roller was used to carefully remove air bubbles between the gel and the membrane. Another piece of sponge was placed on top of the gel.
5. The transfer cassette was closed and placed in one of the two channels (A or B) in the eBlot® Protein Transfer Instrument. The start button for the chosen channel was clicked to start the protein transfer program.
6. When the program had ended, the gel and the sponges were discarded.
7. The membrane was rinsed in distilled water before proceeding with SNAP i.d.® immunodetection (3.19.2). Alternatively, immunodetection was performed through incubation overnight.

3.19.2 Immunodetection using the SNAP i.d. 2.0 Protein Detection System

The SNAP i.d. 2.0 Protein Detection System (Millipore) is a system for hybridization of antibodies to proteins on the membrane. Unlike other methods where the membrane must be incubated for longer periods, the SNAP i.d. 2.0 Protein Detection System utilizes vacuum to drive reagents through the membrane.

Figure 3.3. The Snap i.d.® 2.0 Protein Detection System. The base unit of the SNAP i.d. 2.0 Protein Detection System has room for two blot holder frames, enabling simultaneous immunodetection of two membranes. The system is connected to a vacuum pump. When vacuum is applied, the solutions used in immunodetection are driven through the membrane.

Materials

Snap i.d.® 2.0 Protein Detection System

Tween-TBS (TTBS) (TBS+ 0,1 % v/v tween-20)

TTBS/ 3% BSA (blocking solution)

TTBS/ 2% BSA (blocking solution containing primary or secondary antibody)

Primary antibody Anti-Invasin (bleed # 3)

Secondary Antibody Goat Anti-Rabbit IgG (H+L), HRP

Procedure

1. A blot holder was soaked in dH_2O. The membrane was placed in the holder with the protein side facing down.
2. A filter paper was soaked in dH_2O and placed on top of the membrane. The blot holder was closed, and a blot roller was used to remove any air bubbles. The blot holder was placed in a cassette in the SNAP i.d.® 2.0 Protein Detection System (Fig. 3.3) in such a way that the protein side of the membrane was facing upwards.
3. 30 ml blocking solution (TTBS/ 3% BSA) was gradually poured over the membrane, adding approximately 10 ml of the solution at a time. Vacuum was applied until the membrane was dry.
4. 10 µl Anti-Invasin (primary antibody) was added to 5 ml TTBS/ 2% BSA and mixed. The solution was then added to the membrane and incubated for 10 minutes at room temperature.
5. After the incubation step, vacuum was applied again. When the membrane appeared dry, 30 ml of TTBS (wash buffer) was added to the membrane. This step was repeated 2 times. After last wash the vacuum was turned off.
6. 0.35 µl HRP Goat Anti-Rabbit IgG (secondary antibody) was added to 5 ml TTBS/ 2% BSA. The solution was mixed and added to the membrane and incubated for 10 minutes at room temperature.
7. The washing step 5 was repeated.
8. The membrane was removed from the blot holder and was ready for detection of proteins through chemiluminescence (3.19.3).

3.19.3 Detection of proteins through chemiluminescence

Materials

SuperSignal™ West Pico PLUS Chemiluminescent Substrate

Luminol/enhancer

Stable Peroxide Buffer

Aluminum foil

Procedure

1. The membrane (from step 3.19.2) was placed in a plastic container.
2. The substrate solution was made by mixing 5 ml Luminol/Enhancer and 5 ml Stable Peroxide Buffer.
3. The substrate solution was added to the membrane, the container was closed and covered in aluminum foil to avoid light exposure. The membrane was incubated in the substrate solution for 5 minutes.
4. The instrument Azure c400 was used for visualization of the proteins.

3.20 Protein TCA precipitation to increase protein concentration prior to Western blot

Proteins in the supernatant was concentrated by TCA precipitation prior to SDS-PAGE and blotting. The supernatant (SN) from harvested cells were thawed on ice.

1. The pH of the SN was adjusted to 7, because of sodium deoxycholate.
2. 1.6 ml of the pH adjusted SN was transferred to a 2 ml Eppendorf tube.
3. Sodium deoxycholate was added to a final concentration of 0.2 mg/ml and incubated for 30 minutes on ice.
4. TCA was added to a final concentration of 16% and precipitated for 20-60 minutes on ice, or overnight at 4 °C.
5. The samples were centrifuged at 16000 x g and 4 °C for 25 minutes.
6. The protein pellet was washed twice with ice-cold acetone. It is important to be careful at this step to prevent the pellet from coming loose of the wall. The samples were centrifuged at 16000 x g and 4 °C for 20 minutes.

7. The pellets were dissolved in SDS sample buffer (total volume 30 µl) for SDS-PAGE
8. 25-30 µl of the samples were applied to the SDS-PAGE gel and run at 260 V for 20 minutes.
9. After SDS-PAGE, the gel could be analysed with western blotting (Section 3.19).

3.21 Flow cytometry

Flow cytometry is a technique that enables characterization of individual cells from a heterogeneous cell sample. The information about each individual cell makes it possible to distinguish between different cell populations or cell types in the sample. In addition, flow cytometry can be used to detect antigens or other types of proteins that are displayed on the surface of cells.

A small volume of a cell-containing sample is taken up in the flow cytometer, and the cells are led into a chamber containing a laser beam. As the cells pass through the laser one by one, the laser light will be scattered. Each individual cell scatters the light in a distinct pattern that reveals important information about the cell. The scattering is divided into two main types: forward scatter (FSC) and side scatter (SSC), which give information about cell size and cell granularity, respectively. This data for all the cells in the sample can be visualized in a dot plot. In this plot, each dot represents a single cell, and different cell types can be observed as clusters in the plot.

Flow cytometry can also be used to detect the presence proteins displayed on the surface of bacterial cells. This can be achieved using a primary antibody that binds specifically to the protein of interest, and a secondary antibody conjugated to a fluorochrome that binds specifically to the primary antibody. When the laser beam hits the fluorochrome, it becomes excited and emits a fluorescence signal. This signal is detected by a detector in the flow cytometer. Fluorescence data can be visualized for example in a histogram. In this study, flow cytometry was used to detect Invasin on the bacterial cell surface. The secondary antibody was conjugated to the fluorochrome Fluorescein Isothiocyanate (FITC).

<u>Materials</u>

PBS

PBS/ 2% BSA

Anti-Invasin, bleed # 3 (primary antibody)

Anti-Rabbit IgG (whole molecule)-FITC (secondary antibody)

Aluminum foil

Procedure

1. Cell cultures of recombinant strains of *L. plantarum* WCFS1 were induced and cultivated according to the description in section 3.15. After the cultures had been placed on ice, a small volume of each cell culture was transferred to an Eppendorf tube determined by the formula $\frac{500}{OD_{600}}$.
2. The cell culture was centrifuged at 8000 x g for 3 minutes, and the supernatant was decanted off.
3. The cells were washed with 1 ml PBS and centrifuged on 8100 x g for 3 minutes. The supernatant was pipetted carefully off. The cell pellet could be stored at 4 °C for one day.
4. In an Eppendorf tube, 50 µl PBS/ 2% BSA, and 6.7 µl of the primary antibody Anti-Invasin were combined as in accordance with (Fredriksen et al., 2012).
5. The samples were incubated for 30 minutes in room temperature.
6. The samples were centrifugated on 8100 x g for 3 minutes. The supernatant was carefully pipetted off.
7. The cells were washed three times with 600 µl PBS/ 2% BSA and centrifuged on 8100xg for 3 minutes between each wash. The supernatant was decanted off after each centrifugation.
8. In an Eppendorf tube, 50 µl PBS/ 2% BSA, and 0.3 µl Anti-Rabbit IgG (whole molecule)-FITC were added to the sample.
9. The samples were protected from light and incubated for 30 minutes in room temperature.
10. The samples were centrifuged on 8100 x g for 3 minutes. The supernatant was carefully pipetted off.
11. Step 8 was repeated.
12. The cells were resuspended in 1 ml PBS.
13. 100 µl of the cell suspension was diluted in 900 µl PBS and analysed with flow cytometry in the MacsQuant® Analyser and MacsQuantify™ software.

3.22 Cultivation of selected strains in various media

Two selected strains were grown in three different types of media: Complete RPMI, RPMI/10% MRS and MRS. The cells were induced, and it was performed a growth curve analysis (Section 3.16). Five hours after induction, cell culture was harvested from the cultures and analysed with flow cytometry the next day to determine whether Invasin was displayed on the cell surface. The aim of this experiment was to determine the ability of the recombinant *L. plantarum* to grow in RPMI. RPMI is a common growth medium adapted for eukaryotic cells, for example Caco-2 cells (Section 1.7).

Materials

MRS

RPMI/10 % MRS

RPMI

Spectrophotometer

Cuvettes

SppIP (inducer peptide)

PBS

Procedure

1. The recombinant strains of *L. plantarum* WCFS1 were grown at 37 °C overnight in MRS medium containing 10 µg/ml erythromycin.
2. The next day, 1.5 ml of each of the overnight cultures were transferred to Eppendorf tubes.
3. The Eppendorf tubes with cell culture were centrifuged at 6000 x g for 30 seconds. The supernatant was pipetted off before each cell pellet were resuspended in 500 µl – 750 µl of the appropriate growth medium.
4. The resuspended cell cultures were added to pre-warmed growth medium to a final OD_{600} of 0.1-0.15.
5. The diluted cultures were incubated at 37 °C, and OD was measured at least once every hour, and more often as the cultures were close to achieving the correct OD_{600}. For growth curve analysis, 200 µl of the cell cultures were transferred to a 96-well plate before and after induction, as described in section 3.16. The cultures were

incubated until OD_{600} was 0.28-0.33. At this point, the cultures were induced using 25 ng/ml SppIP. Straight after the addition of SppIP, the tubes were turned upside down once to mix the culture.

6. After induction, the cell cultures were incubated for 5 hours before a small amount of each cell culture was transferred to an Eppendorf tube for analysis with flow cytometry (Section 3.21).

4 Results

4.1 *In silico* analysis of proteins in *L. plantarum*

This study is part of a larger project where the long-term goal is to develop a mucosal vaccine against tuberculosis. *Lactiplantibacillus plantarum* WCFS1 is used as a delivery vehicle of the vaccine. The aim of this thesis was to construct and characterize new N-terminal transmembrane anchors for cell surface display of Invasin.

Antigens, adjuvants, and other medically interesting proteins can be anchored to the surface of *L. plantarum* WCFS1 using either of the four main anchoring strategies that is described in the introduction of this thesis. The strategy that exploits N-terminal transmembrane proteins as anchors seems to have received little attention. It would therefore be interesting to gain a better insight of N-terminal transmembrane proteins in *L. plantarum*, and their functionality as anchors for medically interesting proteins.

Invasin is a virulence factor from *Yersinia pseudotuberculosis.* By interacting with a type of receptors, called β_1-integrins, on the cell surface of the M-cells, Invasin promotes internalization of *Y. pseudotubeculosis* into M-cells (Hamburger et al., 1999).

It has been suggested that co-expression of antigens and the adjuvant Invasin on the surface of *L. plantarum* can give this bacterium an increased ability to encounter immune cells in the gut-associated lymphoid tissue (GALT). This is because surface expression of Invasin by *L. plantarum* might increase the uptake of this bacterium in M-cells.

The starting point for the thesis was a list generated by LocateP2 (Zhou et al., 2008) containing the amino acid sequences of predicted proteins in the *L. plantarum* WCFS1 proteome. In total, the list included 3013 proteins predicted to be intracellular proteins, extracellular proteins, proteins linked to the cell membrane or the cell wall, secreted proteins, hypothetical proteins, and prophage proteins. From the proteome, LocateP2 predicted that 171 proteins were N-terminally anchored. These proteins were the starting point for more in depth analyses. A set of criteria for selecting proteins to anchor Invasin was established to narrow down the 171 N-terminal transmembrane proteins.

<u>The protein to be used as an N-terminal transmembrane anchor should:</u>

1. Give rise to an anchor of approximately 100 amino acids
2. Be predicted to have a single N-terminal transmembrane helix.
3. Not be a hypothetical protein or a prophage protein.
4. Not have a predicated cleavage site, as this might lead to secretion of the protein.
5. Not have a domain which to a large degree overlapped with the sequence of the protein that were going to make up the anchor.

A new and shortened list of exactly 100 proteins was made that only included the proteins that fulfilled all the criteria. The proteins varied greatly when it came to their predicted function. However, a high number of these proteins were predicted to have enzymatic activity. For example, several proteins were annotated as cell surface hydrolases, cell wall hydrolases or proteases. Other proteins had functions related to DNA. Some proteins were related to transport, for example ABC-transporters. Many proteins in the list were only annotated as extracellular, membrane-bound proteins, indicating that their function is still unknown.

The 100 proteins predicted to be N-terminally anchored, were further analysed using the bioinformatic programs SignalP-5.0 and TMHMM v 2.0.

TMHMM (Transmembrane Hidden Markov Model) was used to predict the presence of transmembrane helices in each of the 100 proteins. TMHMM is a program based on a Hidden Markov Model (Krogh et al., 2001; Sonnhammer et al., 1998). The program uses the amino acid sequence of a protein to predict the presence of transmembrane helices in the protein. The program can be used to predict the number of alpha helices in a membrane-spanning protein, their location in the protein and their orientation, meaning whether the helices run inwards or outwards (Sonnhammer et al., 1998). Information about the orientation of transmembrane helices is essential for determining the overall topology of a protein (Sonnhammer et al., 1998). One of the main advantages with TMHMM is that it can also be used to predict helix length (Krogh et al., 2001).

Furthermore, the proteins were analysed with SignalP-5.0. SignalP-5.0 can be used to predict the presence of signal peptides in proteins and the location of their cleavage sites (Almagro Armenteros et al., 2019). The candidate proteins could not contain a cleavage site as this lead to secretion of the protein into the surrounding medium.

All 100 proteins were analysed using TMHMM to predict the presence of a transmembrane helix, and further analysed in SignalP to predict the probability that the protein contained a signal peptide. The protein sequences were converted to FASTA format for rapid analysis with TMHMM and SignalP (Table A-1 and A-2 in Appendix). This was achieved using a script shown in figure A-1.

Importantly, the analyses in TMHMM and SignalP revealed that of the 100 proteins, many were not very suitable for use as N-terminal transmembrane anchors. For example, many of the proteins that had not been predicted to have a cleavage site by LocateP2, had a high probability of containing a cleavage site according to the analyses with SignalP (Tab. A-2). The number of proteins with over 30 % probability of containing a Sec signal peptide, was 46, while the number of proteins with over 50 % probability of containing a Sec signal peptide, was 32 (Tab. A-2). Also, according to the analyses with TMHMM, twenty of the proteins were not predicted to have any transmembrane helix (Tab. A-1). For the eighty remaining proteins that were predicted to have a transmembrane helix, there was little variation in helix length. All proteins had a helix length between 18 and 23 amino acids, and the majority (68%) had a helix length of 23 amino acids.

Subsequently, a high number of the 100 proteins were analysed in Pfam (Mistry et al., 2020) to detect the presence of domains. Pfam gives information about the number of domains of a protein, and their name and localization in the protein sequence. Pfam was used to find out whether a protein contained a domain that overlapped with the sequence that potentially would be used as an anchor. This is because the presence of a complete enzymatic domain, or parts of such a domain, in the anchor sequence may lead to unwanted enzymatic activity.

Weblogos were made to determine whether the proteins contained any consensus sequences. Two weblogos were made based on multiple alignments of sequences of 20 of the 100 predicted N-terminal transmembrane proteins (Data not shown). All proteins included in the alignment were predicted to contain a transmembrane helix in TMHMM and had a low probability of containing a cleavage site according to SignalP. Both weblogos were made based on multiple alignments of parts of the iT of the proteins (Fig. 1.5). The weblogos showed no consensus sequences in these regions. However, based on the sequences of each of these 20 proteins, it was found that lysine (K) and arginine (R) were very common in the iT. Arginine was particularly often the last residue, or second last residue, before the helix.

Based on the results of the *in-silico* analyses, **six** proteins were selected as potential anchors for Invasin (Table 4.2). In this thesis, these proteins will be referred to using the gene number. The composition of the six selected proteins, including their domains, which is predicted by Pfam, is shown in figure 4.2. Table 4.1 shows the probabilities for cleavage sites in the six selected proteins, as predicted by SignalP. This information was used to select out the six proteins. More information about the properties of the selected proteins is given in table 4.2.

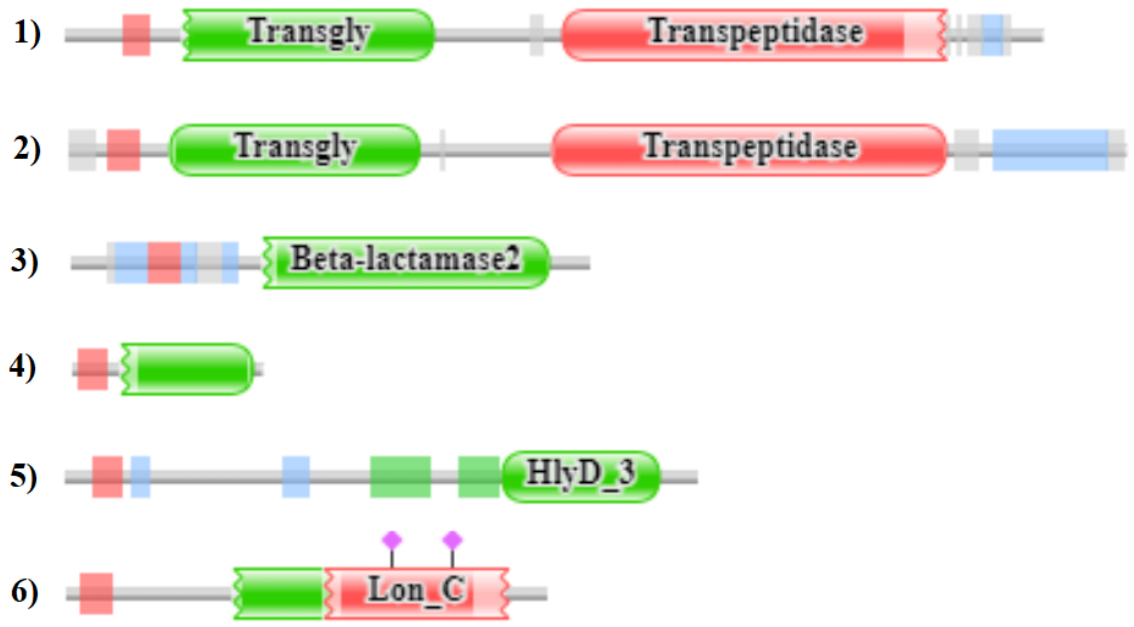


Figure 4.1. Predicted domains in the six selected proteins. Protein 1, Lp_1413, and protein 2, Lp_1751, both contain a transglycosylase domain followed by a transpeptidase domain. The third protein that was chosen out, Lp_2341, contains a beta-lactamase2 domain. Protein 4, Lp_1576, contains a "rhodanese-like" domain (shown in green). Protein 5, Lp_0424, contains a domain named "HLyD family secretion protein". Protein 6, Lp_2132, contains a PDZ domain (shown in green) followed by a Lon protease (S16) C-terminal proteolytic domain (shown in red). The transparent regions show predicted transmembrane helices (red), low complexity-regions (blue), regions with disorder (grey) and regions containing coiled coils (green).

Table 4.1. The results from SignalP (probability of cleavage site) for the six proteins that were selected out to be used as anchors for Invasin.

Gene number	Protein name	Signal peptide (Sec/SPI)	TAT Signal peptide	Lipoprotein signal sequence (Sec/SPII)
Lp_1413	transpeptidase-transglycosylase (penicillin binding protein 2A)	0.0064	0.0002	0.0022
Lp_1751	transpeptidase-transglycosylase (penicillin binding protein 1A)	0.0303	0.0016	0.0087
Lp_2341	beta-lactamase, class A	0.0376	0.0068	0.0252
Lp_0424	bacteriocin ABC transporter, accessory factor PlnH	0.0478	0.0014	0.0084
Lp_1576	rhodanese family protein	0.0699	0.0099	0.0037
Lp_2132	extracellular protease Lon, peptidase S16 family, membrane-bound	0.2101	0.0043	0.0061

Table 4.2. Information about the six proteins that were selected out to be used as anchors for Invasin.

Gene number	Gene name	Protein name	Total protein length (in amino acids)	Position of first domain (analysis from Pfam)	Anchor length (in amino acids)
Lp_1413	*pbp2A*	transpeptidase-transglycosylase (penicillin binding protein 2A)	709	86-267	115
Lp_1751	*pbp1A*	transpeptidase-transglycosylase (penicillin binding protein 1A)	767	75-255	104
Lp_2341	*bla2*	beta-lactamase, class A	376	141-348	133
Lp_0424	*plnH*	bacteriocin ABC transporter, accessory factor PlnH	458	224-265	94
Lp_1576	*Lp_1576*	rhodanese family protein	137	36-131	85
Lp_2132	*Lon*	extracellular protease Lon, peptidase S16 family, membrane-bound	348	122-199	86

The genes *pbp2A* (Lp_1413) and *pbp1A* (Lp_1751) encode penicillin-binding proteins (PBPs). PBPs are a group of enzymes that have an essential role in the synthesis of peptidoglycan in the bacterial cell wall. According to Pfam, both proteins have one transglycosylase domain followed by one transpeptidase domain. Penicillin-binding proteins are the targets of β-lactam antibiotics, for example penicillin. (Sainsbury et al., 2011).

According to Pfam, the gene Bla2 (Lp_2341) encodes a protein with a beta-lactamase domain. Beta-lactamases are enzymes that inactivate penicillins and other β-lactam antibiotics by hydrolyzing the β-lactam ring (Bush & Bradford, 2016). These enzymes can therefore contribute to antibiotic resistance.

The gene *plnH* (Lp_0424) encodes bacteriocin ABC transporter, accessory factor PlnH (Lp_0424).

The gene Lp_1576 encodes rhodanese family protein. According to Pfam, this protein has a rhodanese-like domain. Rhodanese, also called thiosulfate sulfurtransferase, detoxifies cyanide (CN^-) by converting it to thiocyanate (SCN^-).

Lon (Lp_2132) is a gene encoding extracellular protease Lon, peptidase S16 family, membrane-bound. This protein is annotated in Uniprot as “Endopeptidase La, EC”.

Only a fragment of each selected N-terminal transmembrane protein was used to anchor Invasin to the cell membrane. The composition of an N-terminal transmembrane anchor is shown in figure 1.5 in section 1.5. The full-length anchors consisted of the intracellular, N-terminal part of the protein (the iT), the transmembrane helix, and a part of the protein sequence following the helix, which would function as a linker between the helix and Invasin. It was decided beforehand that the length of the linker region should be 53 amino acids in all the anchors. The aim of this was to make the anchors easier to compare to each other.

4.2 Construction of plasmids for N-terminal transmembrane anchoring of Invasin

All selected N-terminal transmembrane anchors were cloned into the pSIP expression vector pLp1261Inv for inducible gene expression. pLp1261Inv (Fredriksen et al., 2012) harbours the Invasin coding gene under control of the inducible SppA promoter (Table 2.4) (Figure 4.1). pLp_1261Inv served as the template for construction of pLp_1751Inv (Figure 4.2) and all the other new plasmids (Table 2.4).

All plasmids constructed in this thesis were made analogous to pLp_1751Inv (Fig. 4.2), but with the use of other pairs of forward and reverse primers (Table 2.1). The plasmids used in this study are listed in table 2.4. The strategy used for constructing the plasmid pLp_1751Inv, containing the anchor sequence from penicillin-binding protein 1A (Lp_1751) fused to Invasin is shown in figure 4.2.

The N-terminal end of Lp_1261 contains the restriction site for NdeI, while the C-terminal end of this anchor contains the restriction site for SalI (Fig. 4.2), therefore gene specific primers containing these restriction sites were designed (Table 2.1). Each primer consisted of the same, 15-bp sequence complementary to the vector that was necessary for In-Fusion cloning, and a sequence that is complementary, and specifically binds to the anchor sequence (reverse primer) or promoter sequence (forward primer). A description of the primers used in this study is given in table 2.2. Genomic DNA was isolated from an overnight culture of a wild-type strain of *L. plantarum* WCFS1 and used as the template to amplify the gene fragments coding for all the anchor sequences. Lp_1751 was amplified using the primer pair Lp_1751F and Lp_1751R (Table 2.1). Through In-Fusion cloning (Section 3.10), the PCR-amplified gene fragment of Lp_1751 were inserted into the NdeI/SalI-digested expression vector pLp_1261Inv (Fig. 4.2).

For the proteins Lp_2341, Lp_1751 and Lp_1576, both long and short anchor versions were made. The aim was to gain insight into the function of the intracellular tail (iT), for example whether this region is necessary for anchoring of Invasin to the cell membrane. It was suggested that a lack of the iT could lead to secretion of Invasin into the surrounding medium. Since these three anchors were made also as a short version, it was constructed in total nine plasmids.

The short versions of these anchors, which were named pLp_2341Short Inv, pLp_1751Short Inv and pLp_1576Short Inv, were made by excluding almost the whole gene sequence that encoded the iT (Fig. 1.5). In these constructs, the kept sequence consists only of the first five amino acids preceding the helix, including the start amino acid methionine. For amplification of shorter sequences of the genes, new forward primers were made, but the reverse primers were the same as those used to amplify the long versions of the anchors (Table 2.1).

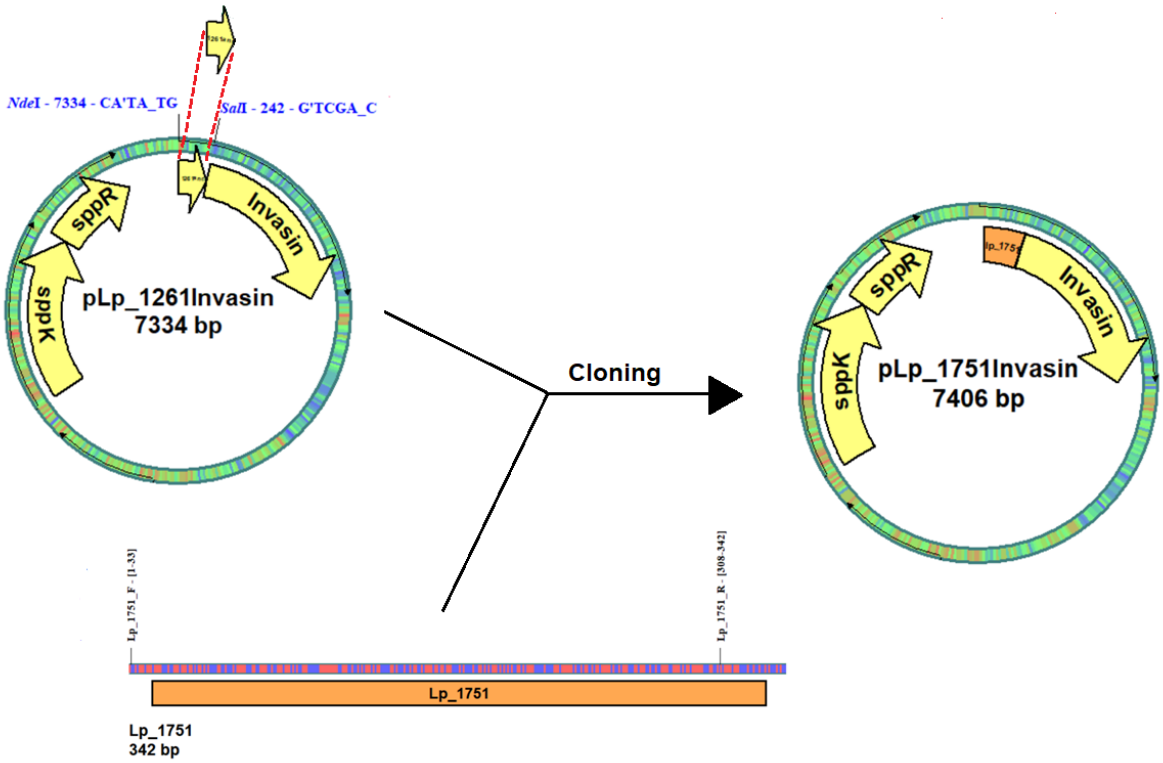


Figure 4.2. Strategy for construction of the plasmid pLp_1751Inv. The gene fragment encoding the N-terminal transmembrane anchor Lp_1751 was PCR-amplified from genomic DNA of *L. plantarum* WCFS1. The pLp_1261Inv plasmid was digested with the restriction enzymes NdeI and SalI-HF to remove the Lp_1261 anchor, as indicated by the red, stapled lines. Through In-Fusion cloning, the Lp_1751-amplified fragment was ligated into the enzymatically digested pLp_1261Inv vector, resulting in pLp_1751Inv. All other constructed plasmids were made analogous to pLp_1751Inv, using other pairs of forward and reverse primers (table 2.1).

After the transformation to competent *E. coli*, colony-PCR was performed on colonies to determine whether the plasmid DNA with the new anchor sequences was successfully obtained (Section 3.7.2). The PCR products were run on a gel electrophoresis to confirm correct plasmid. Figure 4.3 shows the result of colony PCR after transformation of pLp_1751Inv and pLp_2341Inv. The observed bands have the correct size. Sanger sequencing confirmed that six out of the nine plasmid constructs had the correct sequence. The six correct plasmids were further transformed into electrocompetent *L. plantarum* WCFS1 (Section 3.14). The three plasmids that were not successfully transformed into *E. coli*, were pLp_1413Inv, pLp_0424Inv and pLp_1576Inv (all contained the long version of the anchor). Sequence analysis showed that these three plasmids contained mutations in the anchor sequences.

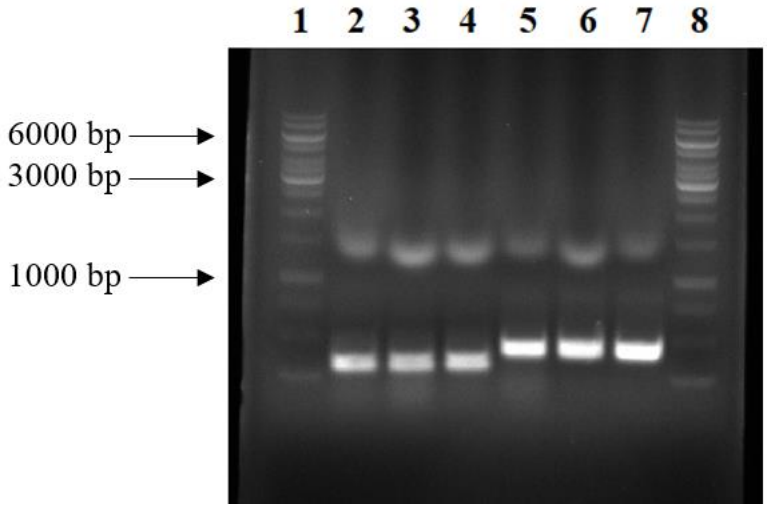


Figure 4.3. Image of gel electrophoresis after colony PCR of pLp_1751Inv and pLp_2341Inv from *E. coli*. The result of colony PCR of pLp_1751Inv (wells number 2-4) and pLp_2341Inv (wells number 5-7) after transformation of these two plasmid constructs into *E. coli*. The length of the nucleotide sequence encoding the Lp_1751 anchor is 312 bp, while the nucleotide sequence encoding the Lp_2341 anchor is 399 bp. This corresponds with the length of the bands observed on the gel. Well number 1 and 8 contain the GeneRuler™ 1 kb DNA ladder.

4.3 Growth curve analysis of *L. plantarum* harbouring plasmids for Invasin production

Growth curve analyses were performed for the six recombinant *L. plantarum* strains harbouring different N-terminal transmembrane anchors translationally fused to Invasin. The aim was to investigate whether the recombinant plasmids harboured by the strains affected the growth rates. The growth of these *L. plantarum* strains made in this study were compared to the *L. plantarum* strains harbouring pEV (empty vector), which do not contain any target genes.

In addition, the growth of the strains was compared with two previously constructed recombinant *L. plantarum* strains, which harbour plasmids for N-terminal transmembrane anchoring (Lp_1568InvS) and lipoprotein anchoring (Lp_1452Inv), respectively. Surface display of Invasin has previously been confirmed for these strains (Fredriksen et al., 2012). pLp_1568InvS express the N-terminal transmembrane anchor Lp_1568 which is fused to the extracellular domains D4 and D5 of Invasin (InvS). The Lp_1568 anchor has a length of anchor 671 amino acids and is thus considerably longer than the anchors constructed in this study. The reason that this strain was included in the growth curve analysis was that a version of this plasmid containing the complete extracellular domain of Invasin (D1-D5) was not obtained

(Fredriksen et al., 2012). pLp_1452Inv express a lipoanchor that was made using a fragment of Lp_1452, which encodes the lipoprotein peptidylprolyl isomerase PrsA (Michon et al., 2016).

The recombinant *L. plantarum* strains were cultivated and induced at OD_{600} ~0.3 as described in section 3.15. After induction, the OD_{600} was measured every fifth minute for 24 hours using a plate reader. OD_{600} was measured both for non-induced and induced cell cultures of each strain (Section 3.16).

Growth curve analyses were repeated three times. The growth of the strains generally showed the same trend in these experiments (Fig. 4.4). Figure 4.4 shows the results from one of the experiments were the growth of all the six recombinant strains were analysed.

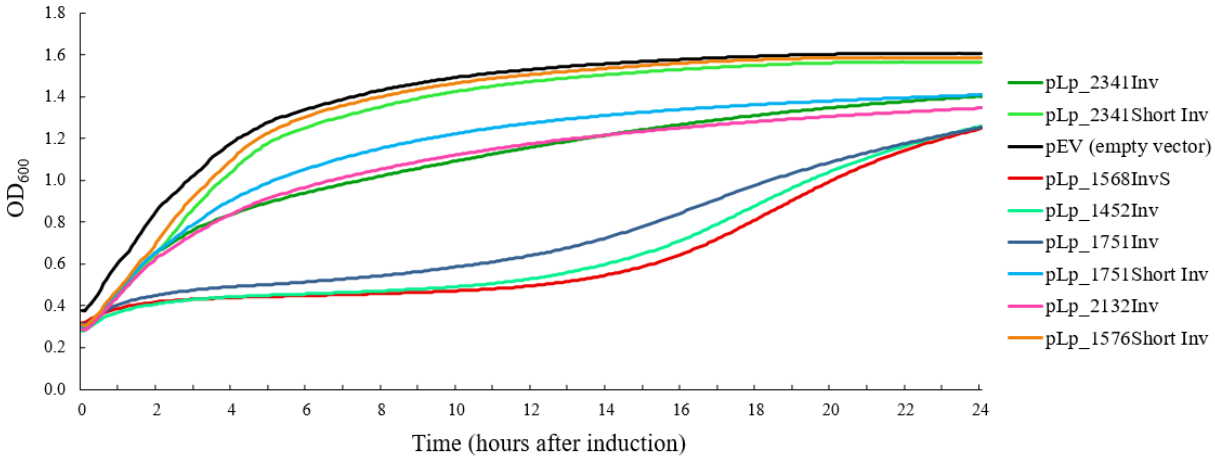


Figure 4.4. Growth curves of recombinant strains harbouring plasmids for production of Invasin. The results shown are from one representative analysis of the three performed growth curve analyses. The bacterial cultures were induced with 25ng/mL SppIP at time point 0. The result for each strain is the average of three technical replicates.

All the non-induced cultures showed similar growth rates as pEV in Fig. 4.4 (Data not shown). After induction, all cultures harbouring N-terminal transmembrane anchors showed reduced growth rate compared to pEV, except for pLp_1576Short Inv and pLp_2341Short Inv, which had a growth rate that was almost identical to pEV even after induction (Fig. 4.4). Interestingly, figure 4.4 shows that the growth rate of pLp_2341Inv was clearly lower than the growth rate of pLp_2341Short Inv. However, both anchor versions of Lp_2341 showed markedly higher growth rates compared to pLp_1568InvS and pLp_1452Inv. The growth rate of pLp_2132Inv was very similar to that of pLp_2341Inv (Fig. 4.4).

pLp_1751Inv showed a considerably lower growth rate after induction. The growth rate of this strain was very similar to pLp_1568InvS and pLp_1452Inv. The growth of pLp_1751Short Inv was also clearly reduced but showed a growth pattern that was more similar to that of pEV.

4.4 Western blot analysis of intracellular Invasin production

After cultivation and induction of the recombinant strains, cells were harvested for Western blot analysis. Western blot was performed to determine whether the recombinant strains of *L. plantarum* produced Invasin. To detect total Invasin production, samples of the protein extracts was analysed for all strains. In addition, the supernatant samples of the recombinant strains were analysed to determine whether the lack of an iT could lead to secretion of Invasin into the growth medium. Prior to Western blot, the harvested cells were lysed using glass beads and applied to SDS-PAGE gels (Section 3.18) to separate the proteins in the cell-free protein extracts (Fig. 4.5). When Invasin is fused to one of the constructed anchors made in this study, the molecular mass is between 60 and 70 kDa. The SDS-PAGE gels did not show any bands of this size that were not also present in the pEV sample (Fig. 4.5). It was therefore necessary to proceed with Western blot to have a possibility of detecting Invasin.

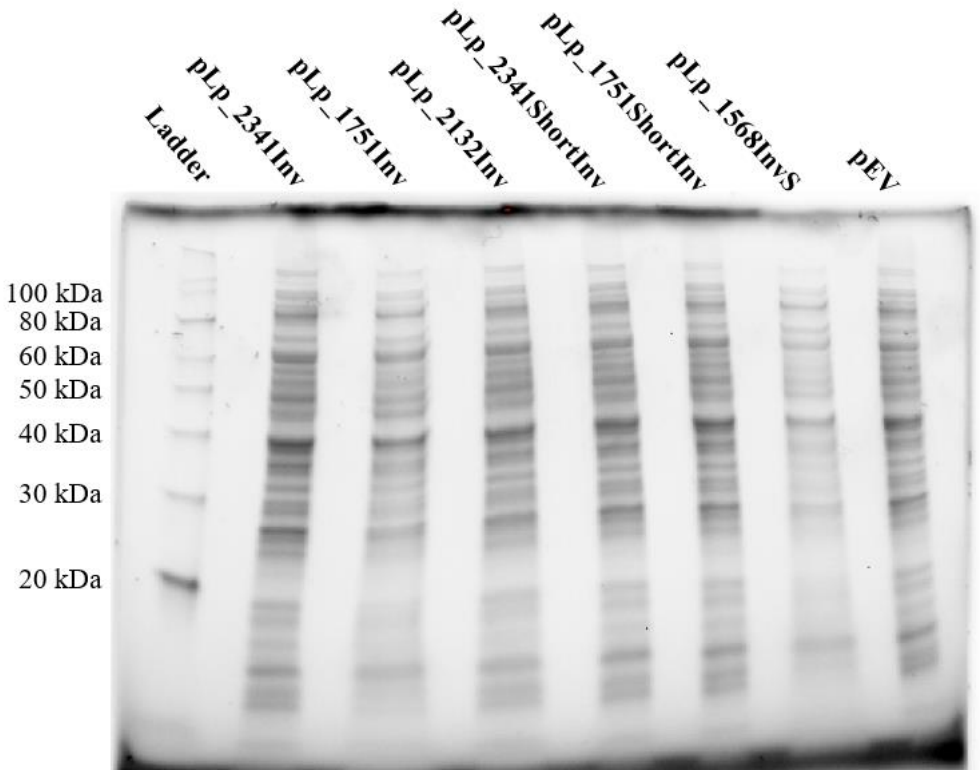


Figure 4.5. A representative SDS-PAGE gel of the protein extract samples of five of the anchors. The gel shows the protein extract samples of pLp_2341Inv, pLp_1751Inv, pLp_2132Inv, pLp_2341Short Inv, pLp_1751Short Inv, pLp_1568InvS and pEV. There was not detected any bands of the size corresponding to Invasin that were not also present in the sample of pEV.

The proteins were transferred from the SDS-PAGE gel to a PVDF membrane as described in section 3.19.1. The SNAP i.d. 2.0 Protein Detection System was used for immunodetection, as described in section 3.19.2. Figure 4.6 and 4.7 show the results of the first Western blot analysis.

On the membrane containing the protein extract samples of all the recombinant strains, it was observed specific binding of antibody in the wells containing pLp_1568InvS, pLp_1452Inv and pLp_1751Inv, showing that Invasin is produced by these strains (Fig. 4.6). The bands in the wells containing pLp_1452Invs and pLp_1751Inv have a size of between 60 and 70 kDa. This corresponds to the molecular mass of the anchors Lp_1452 and Lp_1751 when they are translationally fused to Invasin. The dark area in the well containing pLp_1568InvS, which goes from 220 kDa to 50 kDa, shows that this strain produced Invasin. When the Lp_1568-anchor is fused to InvS, it has a molecular mass of ~94 kDa. Correct sizes of the bands are indicated with arrows (Fig. 4.6).

No specific bands were observed for the other strains analysed, indicating that Invasin was not produced by these strains, or that it was only produced in a low amount. The light grey and white bands observed on the membrane indicate overloading of samples in the wells (Fig. 4.6).

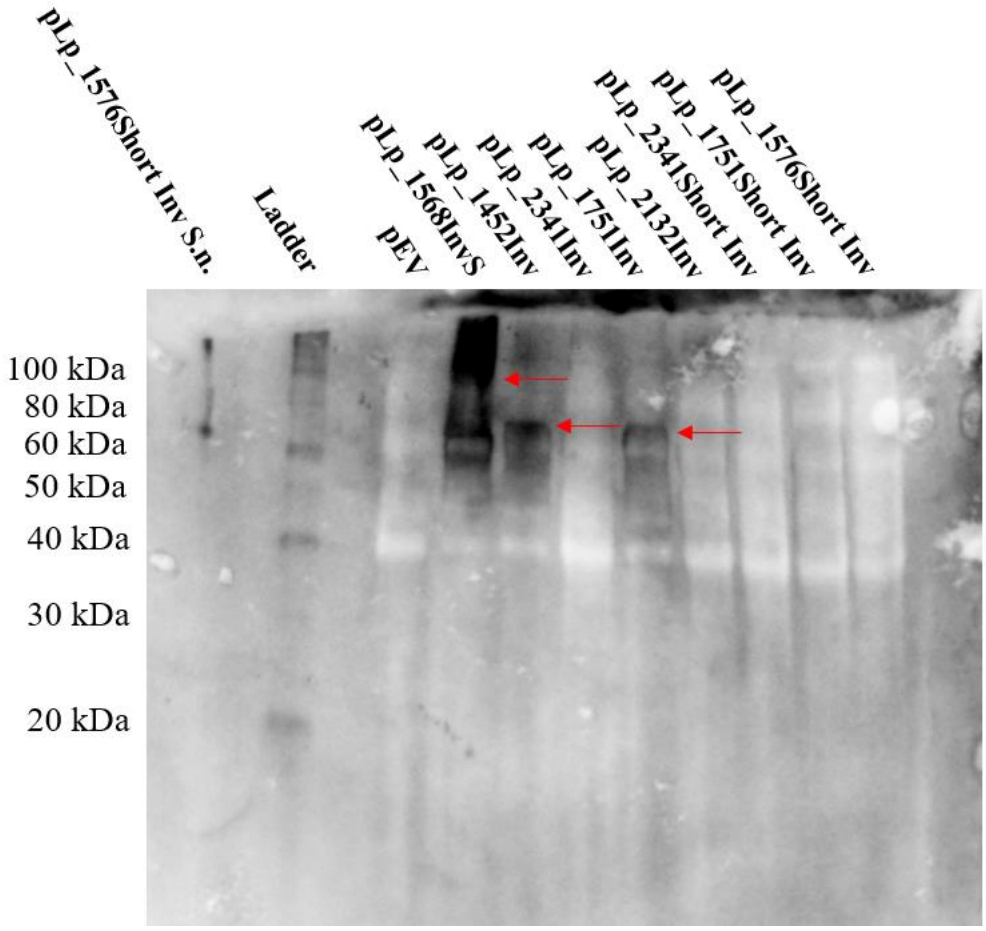


Figure 4.6. Western blot analysis of intracellular production of Invasin. The picture shows the result of a Western blot analysis of cell-free protein extracts from recombinant *L. plantarum* strains harbouring various plasmids for Invasin production. After induction of Invasin production, the strains were grown at 37° C for three hours before harvesting the cells. Lane 1 contains the supernatant sample of pLp_1576Short Inv, which has a predicted molecular mass of ~63 kDa. Lane 3 contains the Magic Mark^{TM} XP Western Standard. The other lanes contain protein extracts from *L. plantarum*-strains harbouring the following plasmids (the predicted molecular masses of each anchor fused to Invasin is written in the parentheses):

5) pEV (empty vector, no signal expected), 6) pLp_1568InvS (~94 kDa), 7) pLp_1452Inv (~68 kDa), 8) pLp_2341Inv (~68 kDa), 9) pLp_1751Inv (~65 kDa), 10) pLp_2132Inv (~63 kDa), 11) pLp_2341Short Inv (~62 kDa), 12) pLp_1751Short Inv (63 kDa), 13) pLp_1576Short Inv (~63 kDa).

The observed specific binding in the wells containing pLp_1568InvS, pLp_1452InvS and pLp_1751Inv show that Invasin is produced by these strains. The expected sizes of the bands in the lanes containing these strains are indicated with arrows.

Western blot analysis was also performed for samples of supernatant from the *L. plantarum* strains harbouring plasmids for Invasin production. The aim was to investigate whether any of the strains harboured anchors that resulted in secretion of Invasin into the surrounding medium (Fig. 4.7). It was previously suggested that the iT of the N-terminal transmembrane anchors

could be important for anchoring Invasin to the membrane, and that the lack of this region would lead to secretion of Invasin into the surrounding medium.

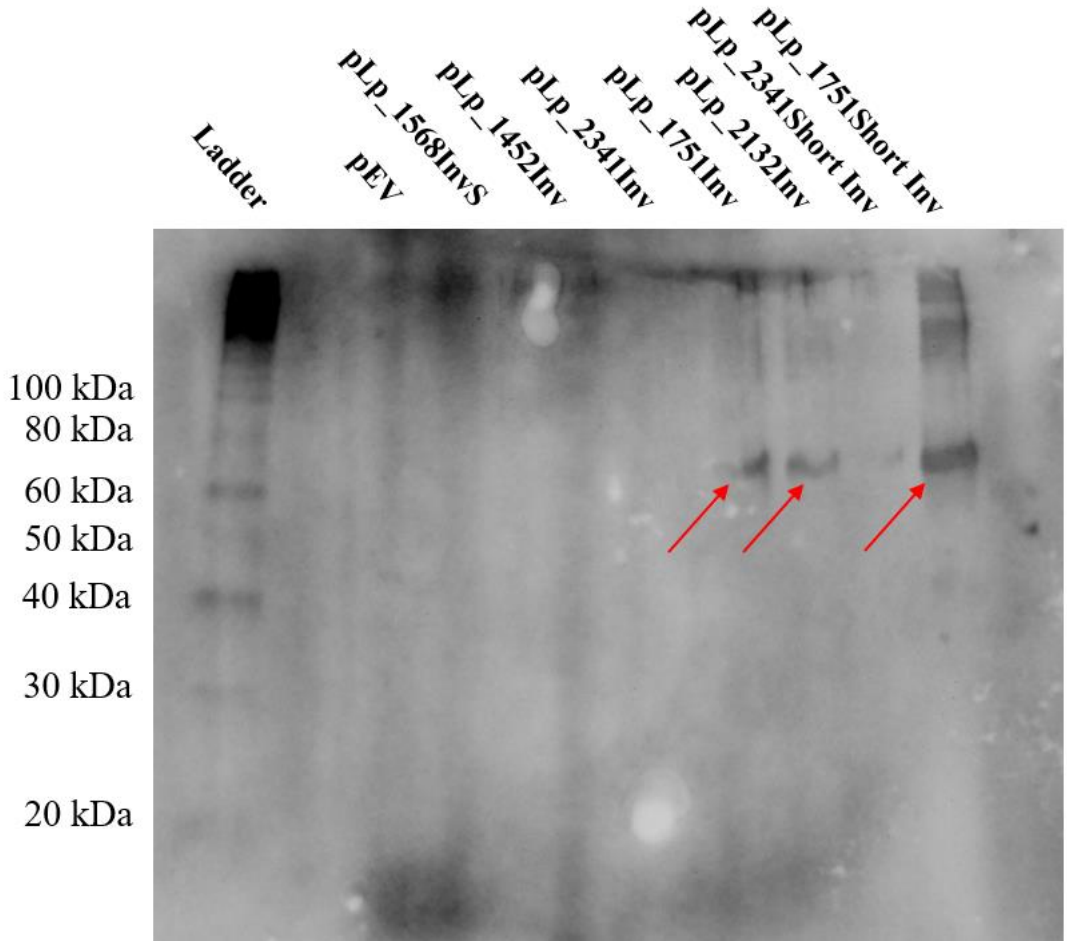


Figure 4.7. Western blot analysis of supernatant samples from recombinant *L. plantarum* strains harbouring plasmids for Invasin production. The picture shows the result of a Western blot analysis of supernatant samples from recombinant *L. plantarum* strains harbouring various plasmids for Invasin production. The supernatant sample of pLp_1576Short Inv is located on the membrane with the cell free protein extract samples (Fig. 4.5). Lane 1 contains the Magic Mark^{TM} XP Western Standard. The other lanes contain protein extracts from *L. plantarum* strains harbouring the following plasmids (the predicted molecular masses of each anchor fused to Invasin is written in the parentheses):

3) pEV (empty vector, no signal expected), 4) pLp_1568InvS (~94 kDa), 5) pLp_1452Inv (~68 kDa), 6) pLp_2341Inv (~68 kDa), 7) pLp_1751Inv (~65 kDa), 8) pLp_2132Inv (~63 kDa), 9) pLp_2341Short Inv (~62 kDa), 10) pLp_1751Short Inv (63 kDa).

Specific bands were observed in the lanes containing supernatant samples from the strains harbouring the anchors Lp_1751Inv, Lp_2132Inv and Lp_1751Short Inv. These bands are indicated with arrows.

Interestingly, some specific bands could be observed with the expected sizes of approximately 60-65 kDa on the membrane with the supernatant samples (Fig. 4.7). These bands were present in the lanes containing supernatant samples from the strains pLp_1751Inv, pLp_2132Inv and

pLp_1751Short Inv. From the position of the bands observed on the membrane, it seems possible that their molecular mass can be a little higher than 60 kDa. This indicates that the anchors harboured by pLp_1751Inv, pLp_2132Inv and pLp_1751Short Inv lead to secretion of Invasin into the medium.

In most of the Western blot analyses, no specific bands were observed on the membranes. Several attempts were made on optimizing the Western blot procedure to better detect intracellular production of Invasin. The BSA-concentration in the blocking solution was varied, using either 1%, 2% or 3% BSA. The first Western blot attempts involved the use of PVDF membranes (Section 3.19.1) (Fig. 4.6-4.7). Since it was difficult to detect Invasin production with this method, blotting was instead performed using the iBlot™ Dry Blotting System (Invitrogen). This procedure involved protein transfer from the SDS-PAGE gel to a nitrocellulose membrane. In another Western blot analysis, immunodetection were performed through incubation overnight. Unfortunately, neither of these changes in the procedure led to detection of intracellular Invasin production. The membranes from these analyses often showed a high number of unspecific bands, or no bands at all (Data not shown).

Another optimization strategy was to vary the concentration of the antibodies that was added in the immunodetection procedure (3.19.2). The concentration of anti-Invasin was increased 10-fold, from 1 µl to 10 µl. The volume of the secondary antibody, Goat Anti-Rabbit IgG (H+L), HRP, was increased from 0.3 µl to 0.35 µl. These concentrations of antibodies were used by a previous master student who achieved successful detection of intracellular Invasin production on Western blot (Nygaard, 2011) However, these changes in antibody concentrations did not improve the Western blot results in this study.

Additionally, TCA precipitation of proteins was also performed once prior to SDS-PAGE and Western blot increase the protein concentration in the samples (Section 3.19). However, this procedure resulted in a “smear” of proteins in the wells in the SDS-PAGE gel, and no specific bands were detected on Western blot after the precipitation of proteins (Data not shown).

4.5 Detection of Invasin on the surface of *L. plantarum* using flow cytometry

Flow cytometry was performed to determine whether Invasin was successfully anchored and displayed on the cell surface of the recombinant strains. The bacterial cells were stained with anti-Invasin. Invasin molecules that are expressed on the cell surface can be bound by anti-Invasin. Subsequently, the cells are stained with the secondary antibody, Anti-Rabbit IgG (whole molecule), conjugated to the fluorochrome FITC. Therefore, cells that display Invasin on the surface will have a fluorescent shift to the right compared to the negative control strain (pEV, empty vector). The fluorescence signal of the recombinant strains made in this study were also compared to pLp_1568InvS and pLp_1452Inv, made by (Fredriksen et al., 2012). In a previous flow cytometry analysis, both pLp_1568InvS and pLp_1452Inv showed strong fluorescence signals compared to the control strain pEV, confirming that Invasin is produced and displayed on the cell surface of these strains (Fredriksen et al., 2012).

The recombinant strains to be analysed with flow cytometry, were cultivated, and induced as described in section 3.15. The samples were stored at 4°C and analysed with flow cytometry the following day. All strains were analysed on day 1 and day 2, except for pLp_1576Short Inv, which was only analysed on day 1.

Since this study is part of a research project where the long-term goal is to use recombinant *L. plantarum* strains in a mucosal vaccine, it is important that these bacteria remain stable during storage. Therefore, two of the strains, pLp_1751Inv and pLp_2132Inv, were also analysed on day 6. The bacteria analysed on day 2 was harvested from a different culture than bacteria analysed on day 1 and day 6. The flow cytometry results for five of the six anchors, analysed on day 1 and day 2, are shown in figure 4.10.

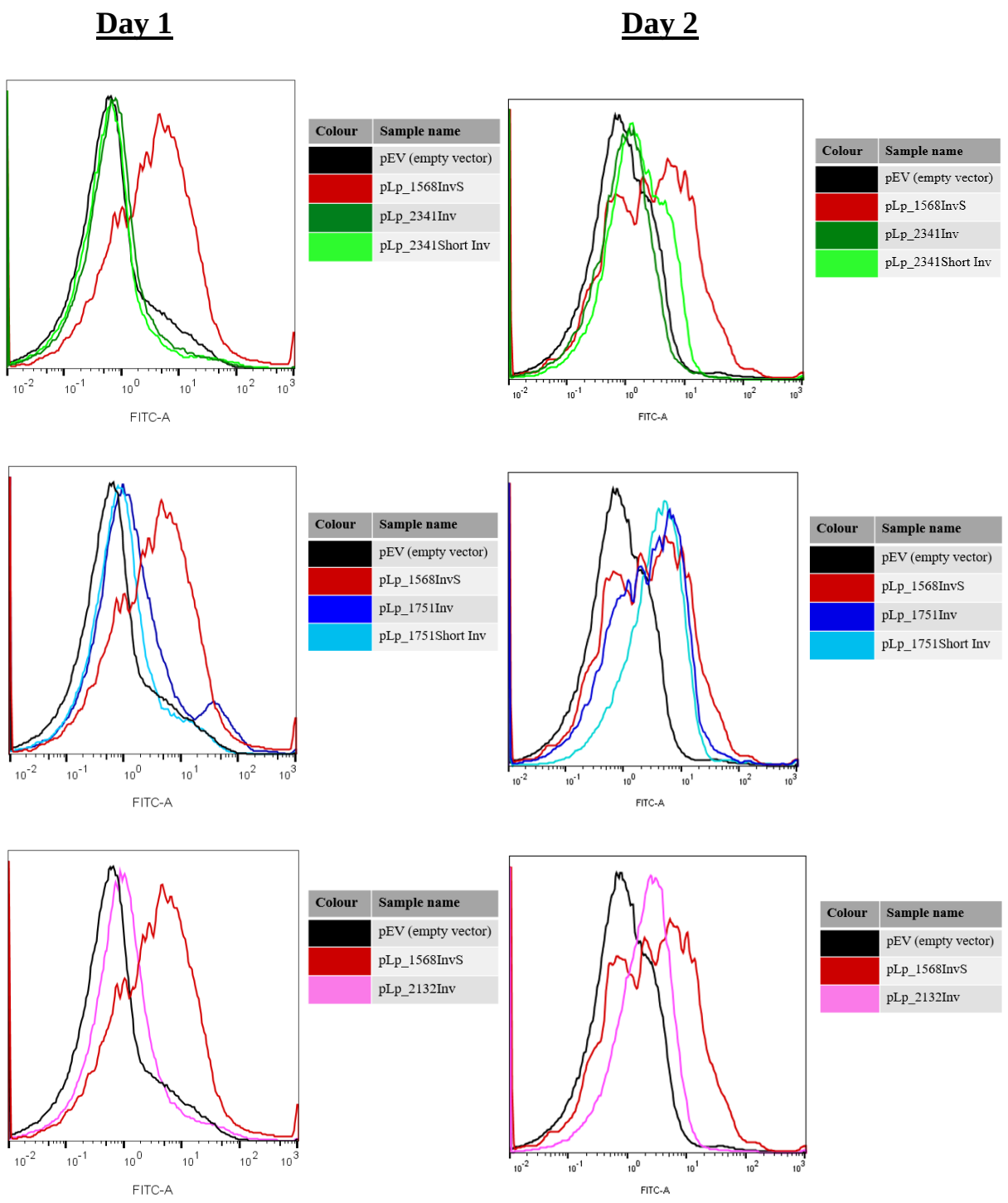


Figure 4.10. Flow cytometry histograms of five of the six recombinant *L. plantarum* strains, from analyses performed on day 1 and day 2 after harvest. Dark green curves: pLp_2341Inv, light green curves: pLp_2341Short Inv, dark blue curves: pLp_1751Inv, light blue curves: pLp_1751Short Inv. Pink curves: pLp_2132Inv. Red curves: pLp_1568InvS. Black curves: pEV (empty vector).

The results for pEV and the positive control pLp_1568InvS were as expected (Fig. 4.10). There was not detected any fluorescence signal for the negative control pEV. The strongest fluorescence signal was detected for the positive control strain, pLp_1568InvS, which showed a strong fluorescent shift compared to pEV. Of the strains harbouring the plasmid constructs made in this study, the strongest fluorescence signals were detected for the strains harbouring the full-length versions of pLp_1751Inv and pLp_2132Inv (Fig. 4.10). This strongly indicated that Invasin is anchored to the cell surface of these strains.

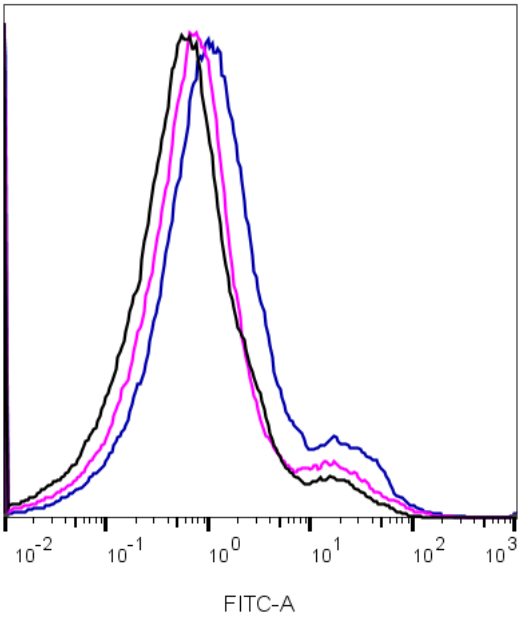
The fluorescence signals of the recombinant strains harbouring anchors for Invasin was generally stronger on day 2 than on day 1.

The strain harbouring the full-length anchor, pLp_1751Inv, and the strain harbouring the short version of this anchor, pLp_1751Short Inv, both showed a weak, but clearly visible fluorescence signal on day 1. Interestingly, on day 1, a “shoulder” could be observed on the curve for pLp_1751Inv (Fig. 4.10). This indicates that cells of this strain formed two distinct cell populations with different degrees of fluorescence. pLp_2132Inv showed a weak fluorescence signal on day 1 and a stronger fluorescence signal on day 2.

On day 1, negligible fluorescence signals were detected for the full-length version and the short version of the Lp_2341 anchor compared to pEV (empty vector). However, slightly stronger fluorescence signals were detected for these two strains on day 2.

pLp_1576Short Inv was only analysed on day 1. No fluorescence signal was observed for this strain on flow cytometry (Data not shown).

The two strains that showed the strongest fluorescence signal, pLp_1751Inv and pLp_2132Inv, were also analysed on day 6 (Fig. 4.11). A weak fluorescence signal was detected for pLp_2132Inv, and a slightly stronger fluorescence signal was detected for pLp_1751Inv. In accordance with the results on day 1, a “shoulder” was observed on the histogram for pLp_1751Inv. The flow cytometry results on day 6 for these two strains were similar to the results observed on day 1.



Colour	Sample name
	pEV (empty vector)
	pLp_2132Inv
	pLp_1751Inv

Figure 4.11. Flow cytometry histograms of pLp_1751Inv and pLp_2132Inv, from an analysis performed on day 6 after harvest. A slightly stronger fluorescence signal was detected for pLp_1751Inv (blue curve) than for pLp_2132Inv (pink curve). pLp_1751Inv also exhibited a small "shoulder", indicating two distinct cell populations.

4.6 Growth curve analysis and flow cytometry of two recombinant *L. plantarum* strains cultivated in various media

Two selected strains, pLp_1751Inv and pLp_2132Inv, were cultivated in three different types of media: RPMI, RPMI with 10% MRS and normal MRS medium. These two strains were selected because they exhibited the strongest fluorescence signals on flow cytometry, showing that Invasin is most likely expressed on the cell surface (Section 4.5). The aim of this experiment was to determine whether these recombinant LAB strains had the ability to survive, and possibly grow, in RPMI medium, which is a medium suitable for culturing of mammalian cells, for example Caco-2 cells (Section 1.7).

The strains were cultivated and induced with 25ng/mL SppIP as described in section 3.22. A growth curve analysis was performed to investigate the growth of the strains in the three different media. Figure 4.12 shows the growth of the induced strains in the first 15 hours of the experiment.

No growth was detected in complete RPMI-medium for either of the strains. When cultivated in this medium, all three strains had a low OD_{600} that remained stable between 0.16 and 0.19. The growth in MRS showed the same trend as that observed in the previous growth curve

analyses (Section 4.3). Interestingly, it was observed that the strains could to some extent grown in RPMI containing 10% MRS (Fig. 4.12). The strain harbouring pEV had the highest growth rate in this medium and the growth reached a stationary phase after around six hours at a OD_{600} of around 0.72-0.73. The growth rate of pLp_2132Inv in RPMI/ 10% MRS was similar to that of pEV. pLp_2132Inv reached a stationary phase after approximately seven hours, having reached an OD_{600} of 0.62-0.63. On the other hand, pLp_1751Inv grew very slowly throughout the experiment and reached a final OD_{600} of 0.44.

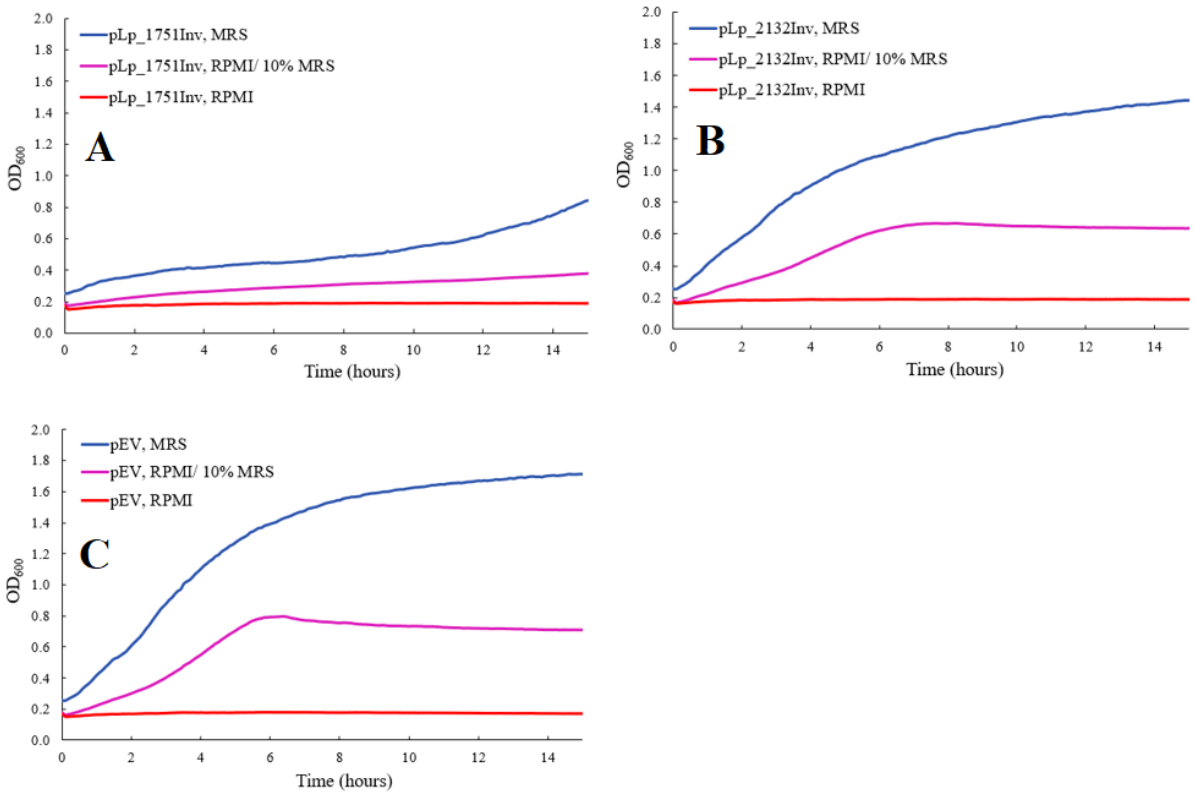
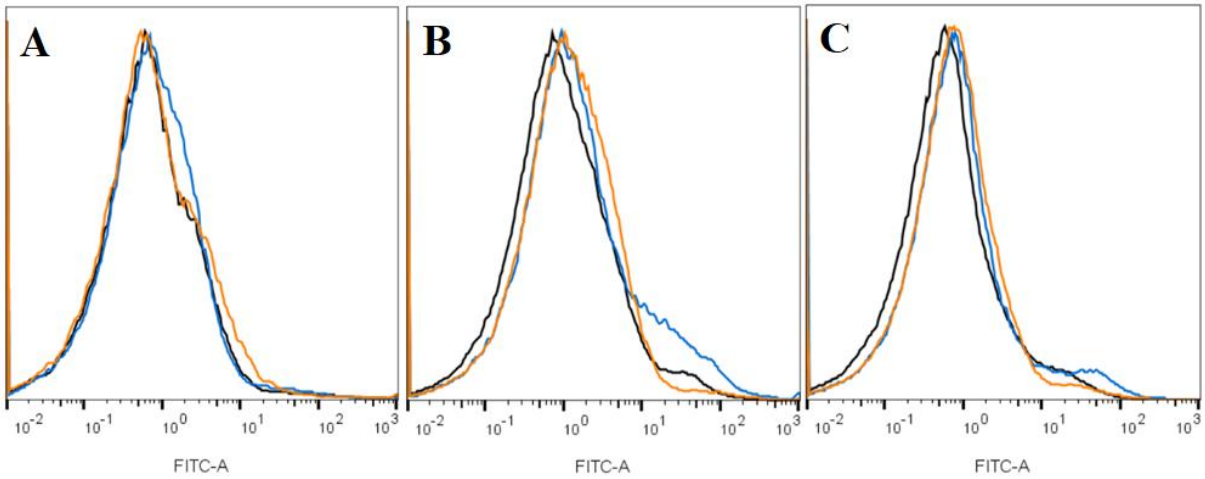


Figure 4.12. Growth of pLp_1751Inv, pLp_2132Inv and pEV (empty vector) after cultivation in three different types of media. The graphs show the growth of induced cultures of pLp_1751Inv (shown in A), pLp2132Inv (shown in B) and pEV (empty vector) (shown in C). The induced cell cultures were cultivated in three different types of media: RPMI, RPMI with 10% MRS. The bacterial cultures were induced with 25ng/mL SppIP at time point 0. For each strain, the blue curves show the growth in MRS, the purple curves show the growth in RPMI/10% MRS and the red curves show the growth in RPMI. The results for each strain are the average of three biological replicates.

Five hours after induction of the cell cultures, samples were prepared for flow cytometry as described in section 3.21. The results of the flow cytometry analyses are shown in figure 4.13. The histograms show the fluorescence of the two strains pLp_1751Inv and pLp_2132Inv in the three different media: RPMI (A), RPMI/10% MRS (B) and MRS (C). The fluorescence signals were generally weak for both strains. While there was no detectable fluorescence signal when

the strains were cultivated in RPMI, weak fluorescence signals were observed for the two strains when cultivated in RPMI/10% MRS or MRS.



4.13. Flow cytometry analysis of pLp_1751Inv, pLp_2132Inv and pEV after induction and cultivation of the strains in three different types of media. The histograms show the fluorescence of pLp_1751Inv (blue), pLp_2132Inv (orange) and pEV (black) when grown in three different types of media: RPMI (A), RPMI/ 10% MRS (B) and MRS (C).

5 Discussion

5.1 *In silico* analyses of proteins in *L. plantarum*

This thesis involved a comprehensive, *in silico* analysis of predicted N-terminal transmembrane proteins from the proteome of *L. plantarum* WCFS1. The goal was to select five candidate proteins to be used as anchors for surface display of Invasin on the cell surface of *L. plantarum*. The starting point of this study was the 3013 predicated proteins by LocateP2 in the *L. plantarum* WCFS1 proteome. Of these 3013 proteins, LocateP2 predicted that 171 proteins were N-terminally anchored. As described in section 4.1, criteria were established to narrow down the list of potential candidates. It was made a new, shorter list that only included the proteins that, according to LocateP2, fulfilled all the criteria (Table A-1 and A-2). This list contained exactly 100 proteins with a great variety of predicted functions. These proteins were analysed using the bioinformatic programs TMHMM and SignalP (Table A-1 and A-2).

It is important to keep in mind that the information about the proteins from these bioinformatic programs, and from LocateP2, are predictions, and there is always a chance that predictions are not correct. The information from these programs should therefore not be regarded as a “final answer” to which proteins that are suitable as N-terminal transmembrane anchors. Rather, the information is indicative of which of the proteins that may be suitable for this purpose.

The results of the analyses from TMHMM and SignalP indicated that many of the proteins in the list were not suitable as anchors, because they were not predicted to have a transmembrane helix, or because they had a high predicted probability of containing a cleavage site. Importantly, it was shown that many of the proteins that did not contain a cleavage site according to LocateP2, were predicted to contain a cleavage site by SignalP. As mentioned, almost half of the proteins, 46, had over 30 % probability of containing a cleavage site.

The TMHMM analyses of the proteins gave more detailed information about the proteins than LocateP2. One of the most important findings was that 20 of the 100 proteins that had been predicted to be N-terminally anchored in SignalP, were not predicted to have a transmembrane helix by TMHMM. It was surprising that the predictions in SignalP and TMHMM differed so much from the predictions in LocateP2. It demonstrates why it was important to analyse the proteins in SignalP and TMHMM, in addition to LocateP2.

According to (Krogh et al., 2001), TMHMM predicts 97-98% of transmembrane helices correctly. Therefore, it seems likely that the predictions of the presence of transmembrane helices were more inaccurate in LocateP2 than in TMHMM.

The output from TMHMM for each protein also showed the orientation of the helix, meaning whether the helix runs outwards or inwards. The vast majority of the proteins have a helix that run outwards, meaning that the N-terminus of the protein is intracellular and the C-terminal part of the protein, following the helix, is extracellular. This can also be called a N_{in} – C_{out} orientation. 6 of the 100 proteins had the opposite orientation, meaning that their N-terminal part was extracellular, and their C-terminal part was intracellular (N_{out} – C_{in}) (Table A-1 and A-2 in Appendix). These proteins are not suitable as anchors in this study. This is because Invasin was translationally fused to the C-terminus of the N-terminal transmembrane anchors, and proteins with a N_{out} – C_{in} orientation would lead to intracellular localization of Invasin. It was found that one of the proteins with a N_{out} – C_{in} orientation was encoded by the gene Lp_1576 and belonged to the rhodanese family proteins. This protein was selected as a candidate protein and was used to make one of the short anchors, pLp_1576Short Inv. In the initial analyses of the proteins, the N_{out} – C_{in} orientation of the protein encoded by Lp_1576 was overlooked, and it was not discovered before the experiments had been completed. This property of Lp_1576 can explain why no fluorescent shift was detected for pLp_1576Short Inv in the flow cytometry analysis (Data not shown).

It was also discovered that some of the proteins predicted to be N-terminally anchored in LocateP2 were likely other types of proteins. Notably, the protein with the gene number Lp_3014 (Protein nr. 80 in table A-1 and A-2 in Appendix) were in LocateP2 annotated as “extracellular transglycosylase with LysM peptidoglycan binding domain”. However, LocateP2 also predicted that this protein was N-terminally anchored and that it did not contain a cleavage site. On the other hand, SignalP predicted that the protein had a 99% probability of containing a cleavage site that leads to secretion via the Sec translocase pathway (Tab. A-2). It is known from previous research that the protein encoded by the Lp_3014-gene is a transglycosylase, containing an N-terminal LysM domain that binds non-covalently to peptidoglycan in the cell wall (Fredriksen et al., 2012; Pham et al., 2019). This protein was used to develop a LysM-type anchor for surface display of Invasin in *L. plantarum* (Fredriksen et al., 2012; Kuczkowska et al., 2015; Mathiesen et al., 2020; Michon et al., 2015; Pham et al., 2019).

Like Lp_3014, Lp_1884 (Protein nr. 49 in table A-1 and A-2 in Appendix) was also annotated as an "extracellular protein with LysM peptidoglycan binding domain. However, this protein was predicted to have a low probability of containing a cleavage site by SignalP, and similarly to Lp_1576, it had a N_{out} – C_{in} orientation.

Based on multiple sequence alignments of twenty of the predicted N-terminal transmembrane proteins, Weblogos were made to investigate whether the intracellular tail of the proteins contained any consensus sequence (Data not shown). No consensus sequence was found. However, it was observed the positively charged residues lysine and arginine were often present in the N-terminal intracellular region of the proteins. This is in accordance with the previously mentioned "positive-inside rule", meaning the fact that lysine and arginine plays a role in determining the orientation of the helix (Sonnhammer et al., 1998). It has also been suggested that lysine and arginine interact with the negatively charged membrane phospholipids (Tjalsma et al., 2004; van Roosmalen et al., 2004). Perhaps this interaction is important to achieve stable anchoring of the N-terminal transmembrane protein in the membrane.

It could be interesting to perform additional analyses of the N-terminal transmembrane proteins. For example, the program protpi (https://www.protpi.ch/Calculator/ProteinTool) can be used to find the net charge of the iT of these proteins. It could be interesting to investigate whether the net charge of the iT varies with the length of this region.

5.2 Construction of vectors for N-terminal transmembrane anchoring of Invasin

Based on the results of the *in-silico* analyses, six proteins were chosen out to be used as anchors for Invasin (Table 4.2). In addition, short anchors, with deleted iT, were made for three of the selected anchors. Thus, the total number of plasmids constructs were nine. Six were successfully transformed into competent *E. coli* cells, and further transformed into electrocompetent *L. plantarum* WCFS1 cells. The three plasmid constructs that were not successfully transformed in *E. coli*, were pLp_1413Inv, pLp_0424Inv and pLp_1576Short Inv. The reason was that mutations had arisen in the anchor sequences, which in most cases were indel mutations. This may indicate that these plasmid constructs were toxic to the *E. coli* cells.

5.3 Growth curve analysis of *L. plantarum* harbouring plasmids for Invasin production

For the six recombinant *L. plantarum* strains constructed in this study, the growth was measured to investigate whether the plasmids containing anchored Invasin affected their growth rates after induction of gene expression (Fig. 4.4). If a plasmid containing a foreign protein is transformed into bacterial cells, it can affect the growth rate of those bacteria because increased metabolic energy is required for plasmid function and maintenance (Fakruddin et al., 2012). Previous studies have shown that production of heterologous proteins in lactobacilli can reduce the growth rate of the producer strain (Fredriksen et al., 2012; Mathiesen et al., 2020).

In the growth curve analysis performed in this study, the most substantial decreases in growth rates after induction were observed for the strains pLp_1568InvS and pLp_1452Inv (Fig. 4.4). This is in accordance with previous analyses, which showed that the *L. plantarum*-strains harbouring anchors derived from the proteins Lp_1452 or Lp_1568 showed a clearly reduced growth after induction (Fredriksen et al., 2012). Surface display of Invasin has previously been confirmed for these strains through flow cytometry and immunofluorescence microscopy (Fredriksen et al., 2012). The reduced growth rates of the strains pLp_1568InvS and pLp_1452Inv indicate that overproduction of the anchors and Invasin resulted in a significant amount of stress (Fredriksen et al., 2012).

Of the six recombinant *L. plantarum* strains made in this study, the largest reduction in growth rate after induction was observed for the strain harbouring pLp_1751Inv, indicating Invasin production (Fig. 4.4). Interestingly, the growth pattern of this strain resembled that of pLp_1568InvS and pLp_1452Inv. pLp_1751Inv, and the other recombinant *L. plantarum* strains developed in this study harbour plasmids for production of a recombinant protein comprising Invasin and an N-terminal transmembrane anchor. These recombinant proteins must therefore be secreted via the sec translocase system to anchor Invasin to the cell membrane. This can lead to overloading of the translocation machinery which may lead to stress responses (Mathiesen et al., 2020). This can be a possible explanation for the reduced growth rates that were observed for some of the strains after induction (Fig. 4.4). The stress might also be related to anchoring of the recombinant proteins in the membrane.

The five other recombinant *L. plantarum*-strains, pLp_2341Inv, pLp_2341Short Inv, pLp_2132Inv, pLp_1751Short Inv and pLp1576Short Inv, showed only a low or moderate reduction in growth rate after induction (Fig. 4.4). This indicates that Invasin was not produced

by these strains. Another possible explanation is that these strains did produce Invasin, but that the anchors did not lead to successful anchoring of Invasin in the cell membrane, resulting in secretion of Invasin into the surrounding medium. It might be that secretion of Invasin is less stressful for the bacteria than anchoring of Invasin to the cell membrane. If this is the case, it can explain why only a low or moderate reduction in growth rate were observed for these strains.

Notably, the recombinant strains harbouring the short versions of the anchors had a higher growth rate than the strains harbouring the long versions of the anchors. The induced strains harbouring pLp_2341Inv or pLp_1576Short Inv only showed a slight reduction in growth rate compared to the non-induced strains and the strains harbouring empty vector (pEV) (Fig. 4.4). The reason might be that the anchors lacking the intracellular tail are not successfully anchored in the membrane, but instead secreted into the surrounding medium. When it comes to pLp_1576Short Inv, the anchor harboured by this strain was made from Lp_1576 which had a N_{out}-C_{in} orientation, leading to intracellular production of Invasin. It might be that it is less stressful for the bacteria to express Invasin intracellularly, compared to translocation and surface expression of Invasin.

5.4 Western blot analysis of intracellular Invasin production

Cell free protein extracts of the strains were analysed by Western blot to detect total Invasin production. In addition, supernatant samples were analysed to investigate whether anchors without the iT would result in secretion to the supernatant of Invasin.

On the first Western blot analysis of the cell-free protein extracts, it was observed dark areas or bands caused by specific binding of anti-Invasin (Fig. 4.6). This was observed in the lanes containing pLp_1568InvS, pLp_1452Inv and pLp_1751Inv, confirming that Invasin was produced by these strains. This result is in accordance with the growth curve analysis, which showed a substantially reduced growth rate for these three strains (Fig. 4.4). The strains pLp_1751Inv and pLp_1568InvS were also analysed with flow cytometry, and this analysis showed clear fluorescent shifts for these two strains compared to pEV (Fig. 4.10). this was also in consistence with the specific binding of anti-Invasin that was observed for these two strains (Fig. 4.6). Unfortunately, strong specific bands like the ones on the membrane in figure 4.6 were not observed in later analyses of the cell-free protein extracts.

Interestingly, in one of the Western blot analyses of the supernatant samples, specific bands were observed for pLp_1751Inv, pLp_2132Inv and pLp_1751Short Inv (Fig. 4.7). The sizes of these bands appeared to be approximately 60-65 kDa, in accordance with the predicted mass of these anchors when they are fused to Invasin. This result indicates that these strains to some extent secrete Invasin into the surrounding medium. However, this result was not reproduced, and the experiment should be repeated before drawing strong conclusions.

When performing Western blot, many factors can influence the result. It is important to consider which concentration of BSA to use in the blocking solution, and the concentration of the primary and secondary antibodies added in the immunodetection procedure. In addition, the exposure time during development of the signal from the membrane can influence the intensity of specific bands. The exposure time used in the Western blot analyses should perhaps have been increased to better detect specific bands that potentially were present.

As described in section 4.4, many attempts were made to optimize the Western blot procedure to better detect Invasin production. The concentration of BSA in the blocking solution was varied, and in the final Western blot analyses, the concentrations of the antibodies were increased. This did not lead to detection of specific bands on the membrane.

The low detection of Invasin on SDS-PAGE and Western blot may indicate that the intracellular production of Invasin is low in the recombinant strains harbouring the N-terminal transmembrane anchors. Another possibility is that the antibodies used have become old and therefore partially degraded.

In most of the Western blot analyses, the volume of cell-free protein extract from each of the strains was not adjusted according to the OD_{600}-value at the time of harvest. This should have been done to ensure that the samples of protein extracts applied on the SDS-PAGE gel are from approximately the same number of cells. When the samples are not adjusted according to the OD_{600}-value at the time of harvest, it can make the results more uncertain and difficult to interpret. For example, if bands with a size that corresponds to Invasin are observed in two lanes, but the bands have different intensity, this can be caused by differences in the degree of Invasin production. However, the reason can also be that the cell-free protein extracts from the different strains are from a different number of cells.

5.5 Detection of Invasin on the surface of *L. plantarum* using flow cytometry

Flow cytometry analyses were performed to investigate whether Invasin was anchored and displayed on the cell surface of the induced, recombinant *L. plantarum* strains. Figure 4.10 shows the flow cytometry histograms for five of the six recombinant *L. plantarum* strains constructed in this study, from analyses performed on day 1 and day 2 after harvest.

On both days, the positive control strain pLp_1568InvS showed the strongest fluorescence signal compared to the negative control (pEV) (Fig. 4.10). Among the novel plasmids, pLp_1751Inv and pLp_2132Inv showed the strongest fluorescence signals compared to pEV. This result shows that Invasin was most likely successfully anchored and displayed on the cell surface of these strains. The strain harbouring the short version of the Lp_1751 anchor, pLp_1751Short Inv, also showed a clear fluorescent shift, indicating that Invasin is also displayed on the cell surface of this strain (Fig. 4.10).

At day 1, the histogram of pLp_1751Inv showed a “shoulder” to the right of the main curve. This is called bimodality and indicates that the recombinant strain formed two distinct cell populations (Mathiesen et al., 2020).

On day 1, the fluorescence signals of the five recombinant *L. plantarum* strains were clearly lower than the fluorescence signal for the strain pLp_1568InvS. However, the fluorescence signals detected on day 2 were stronger for all five strains, and for pLp_1751Inv and pLp_1751Short Inv, it was detected fluorescence signals with similar signal intensities to that of pLp_1568InvS (Fig. 4.10). One possible explanation for the increased fluorescence signals at day 2 can be that the cell wall of the bacteria had begun to break open after two days of storage at 4 °C, which may have resulted in more exposed Invasin.

One possible reason for the low fluorescence signals observed for the recombinant strains developed in this study compared to pLp_1568InvS on day 1, might be the result of the length of the anchor sequences. The length of the Lp_1568 anchor is 671 amino acids. This corresponds to the complete protein, but with a C-terminal truncation of seven residues (Fredriksen et al., 2012). The weak fluorescence signals observed for recombinant *L. plantarum* strains harbouring the anchors made in this study, can indicate that Invasin was anchored, but that it was partly embedded in the cell wall. The Lp_1568 anchor is considerably longer, and it is thus likely that InvS became more exposed to the surrounding environment, resulting in a

stronger fluorescence signal for pLp_1568InvS compared to the fluorescence signals detected for the strains developed in this study.

Based on these results, it could be interesting to construct longer versions of the anchors that were constructed in this study. These anchors could be transformed into *L. plantarum cells*, which subsequently could be analysed with flow cytometry to investigate whether these longer anchor versions give a stronger fluorescence signal compared to the anchor versions made in this study.

Alternatively, cells of the recombinant strains could be treated with lysozyme prior to flow cytometry. Lysozyme leads to partial degradation of the cell wall and can therefore reveal proteins, such as Invasin, that are embedded in the cell wall. Treatment with lysozyme might enable more accessible Invasin. (Øverland, 2013) constructed vectors for secretion and anchoring of the *Mycobacterium tuberculosis*-antigen Ag85B-ESAT6 fused to a dendritic-cell binding peptide on the surface of *L. plantarum*. Fluorescence signals were detected for all strains harbouring antigen. Still, the fluorescent shift was weak for a construct that harboured an antigen anchored to the membrane using a lipoprotein anchor. However, after a lysozyme treatment of the cells prior to flow cytometry, all strains harbouring antigen showed stronger fluorescence signals. This indicates that the lysozyme treatment had partially broken down the cell wall and revealed antigens that were previously embedded in the cell wall.

Another possible explanation for the lower fluorescence signals for the recombinant strains developed in this study compared to the strain pLp_1568InvS, is that the expression of Invasin on the cell surface of these strains is low compared to pLp_1568InvS. The Invasin anchors made in this study is expressed using the same, inducible promoter (P_{sppA}) as the one used to express pLp_1568InvS. Therefore, if Invasin is produced by some of the strains developed in this study, it seems likely that the intracellular production of Invasin in these strains is at the same level as for pLp_1568InvS. However, it might be that some of the anchors led to lower surface expression of Invasin due to problems in the translocation process.

The strains pLp_1751Inv and pLp_2132Inv were also analysed with flow cytometry on day 6 (Fig. 4.11) to check whether Invasin was still displayed on the cell surface. For both strains, the fluorescence signals detected were weaker on day 6 than on day 2. This may indicate that some of the Invasin proteins displayed on the cell surface of these strains have become degraded or released from the cell membrane after six days.

Since some of the recombinant *L. plantarum* strains showed fluorescence signals on flow cytometry, it could also have been interesting to analyse the strains with immunofluorescence microscopy. This type of microscopy can be used to detect proteins that are displayed on the cell surface of bacteria, and how these proteins are distributed on the cell surface.

Importantly, the clear fluorescence signals detected for pLp_1751Inv, pLp_1751Short Inv and pLp_2132Inv strongly indicate that Invasin is anchored, and surface displayed on the cell surface of the bacterial cells. It could be interesting to repeat the experiment but vary some factors as an attempt to optimize the method. For example, it might be that the antibodies are degraded more rapidly in room temperature than at 4° C. It could therefore be interesting to incubate the samples at 4° C instead of room temperature to investigate whether this results in stronger fluorescence signals.

5.6 Growth curve analysis and flow cytometry of two recombinant *L. plantarum* strains cultivated in various media

The recombinant strains pLp_1751Inv and pLp_2132Inv were cultivated in three different types of media: RPMI, RPMI with 10% MRS and normal MRS medium. To aim was to test whether the strains were able to grow in RPMI-medium. RPMI-medium is used to cultivate mammalian cells, for example Caco-2 cells. Caco-2 cells can be converted into M-cell like cells that can be used to study in vitro transcytosis of bacteria across M-cells (Kernéis et al., 1997; Solberg, 2015).

Figure 4.12 shows that none of the tested strains grew in complete RPMI-medium. This does not mean that RPMI is harmful to the cells. Rather, it indicates that this medium lacks some of the nutrients that are essential for the growth of the *L. plantarum* cells.

It was observed that the strains to some extent could grow in RPMI/ 10% MRS (Fig. 4.12). This indicates that the bacterial cells can grow with only a small amount of MRS in the medium. pLp_1751Inv grew slowly throughout the experiment. However, the growth of pLp_2132Inv and pEV was faster, and these stains reached a stationary phase after seven and six hours, respectively. It is likely that the MRS in the medium had been completely consumed at this time points, inhibiting further growth of these strains. In the flow cytometry analysis of the cells incubated in RPMI/ 10% MRS, it was detected a slight fluorescent shift for both pLp_1751Inv and pLp_2132Inv, indicating that Invasin is expressed on the surface of these bacterial cells (Fig. 4.13 B).

The fact that the strains were unable to grow in complete RPMI-medium can be an advantage. This is because it makes it easier to have control on the number of bacterial cells during analyses of internalization of the bacteria by Caco-2 cells. When bacterial cells are going to be co-incubated with Caco-2 cells, there needs to be an optimal ratio between the number of bacterial cells and the number of Caco-2 cells. It is therefore important to know the number of bacterial cells that are internalized compared to the total number of bacterial cells that were added in the beginning of the experiment.

An alternative way to investigate internalization of LAB by Caco-2 cells is to first cultivate the LAB-strains in another medium than RPMI, preferably MRS, to obtain a high number of cells. This medium is subsequently washed off, and RPMI-medium is added to the bacterial cells to co-incubate them with Caco-2 cells.

5.7 Concluding remarks and future perspectives

This thesis describes an attempt to anchor and display Invasin on the cell surface of *L. plantarum* using homologous N-terminal transmembrane proteins. Invasin production was investigated using Western blot and flow cytometry. The flow cytometry analyses showed clear fluorescence signals for three of the *L. plantarum* strains: pLp_1751Inv, pLp_1751Short Inv and pLp_2132Inv. This is a strong indication that Invasin is anchored and displayed on the cell surface of these three strains. In future studies, it could therefore be interesting to investigate whether the anchoring motifs harboured by these strains can be used for anchoring other proteins relevant in vaccine development. This can for example be other adjuvants than Invasin, or antigens.

Notably, other types of molecules than Invasin can also be used to target LAB to M-cells (Ma et al., 2018). An alternative strategy to enhance the effect of a LAB-based mucosal vaccine is to express dendritic-cell targeting peptides on the surface of LAB (Ma et al., 2018; Mohamadzadeh et al., 2009). Interestingly, (Ma et al., 2018) developed an oral vaccine against porcine epidemic diarrhea virus (PEDV) using recombinant *Lactobacillus casei* 393 as a delivery vehicle. It was developed four recombinant strains of *L. casei* 393 that all expressed the core neutralizing epitope (COE) of the PEDV spike protein on the cell surface. One of the strains expressed the COE antigen only, while two of the other strains expressed the COE-antigen fused with either an M-cell targeting peptide (Col) or a DC-targeting peptide (DCpep), respectively. The last strain expressed the COE-antigen fused to both the M-cell targeting peptide and the DC-targeting peptide. The recombinant strains were orally administered in mice

to investigate their immunogenicity. It was shown that the three strains expressing the M-cell targeting peptide and/or the DC-targeting peptide induced faster and stronger immune responses both at the mucosal and systemic levels compared to the strain that only expressed the COE-antigen (Ma et al., 2018). The result of this study indicates that the M-cell targeting peptides expressed on *L. casei* most likely led to increased uptake of this bacterium in M-cells, and that the expression of dendritic cell-targeting peptides resulted in more frequent capturing of *L. casei* by DC's residing in the epithelial cell wall of the small intestine.

The covid-19 pandemic has demonstrated the importance of efficient production and administration of vaccines. Compared to systemic injection of vaccines, an advantage with oral mucosal vaccination is that it enables more efficient mass vaccination during pandemics, and that mucosal vaccines can be easier to produce and administer (Ma et al., 2018; Owen et al., 2013). For these reasons, further development of mucosal vaccines is important to get control of infectious agents that spreads disease via the mucosal route and that can give rise to epidemics, or even global pandemics.

One of the most important advantages with the *L. plantarum*-based delivery system for mucosal vaccines that was exploited in this study is that it can be used to develop vaccines against a very high number of different diseases, for example tuberculosis, chlamydia, HIV-1 and cancer (Fredriksen et al., 2010; Kuczkowska et al., 2017; Kuczkowska et al., 2019; Mobergslien et al., 2015). This can be done by developing new recombinant *L. plantarum* strains that express the relevant antigens.

Anchoring and display of molecules on the surface of LAB can have many other biotechnological applications, such as immobilization of enzymes on the surface of *L. plantarum* (Nguyen et al., 2019; Pham et al., 2019; Pham et al., 2020). This can for example be enzymes involved in the breakdown of nutrients, for example polysaccharides. The end products of some enzymatic reactions can be very interesting for the food industry. An example is β-mannanase which catalyses the breakdown of several types of mannans into manno-oligosaccharides (MOS) (Nguyen et al., 2019). Earlier studies have shown that MOS have health-promoting effects in livestock and humans (Nguyen et al., 2019). Studies on immobilization of enzymes on the surface of LAB, for example as a strategy to convert by-products into more valuable products, can also be interesting in the future.

Based on the results from this study, of the recombinant *L. plantarum* strains that were made, pLp_1751Inv seems to be the most promising strain for anchoring and cell surface display of

Invasin. It could therefore be interesting to include this strain in future studies on N-terminal transmembrane anchoring of Invasin or other medically interesting proteins.

From three of the six candidate N-terminal transmembrane proteins, both long and short anchor versions were made. In the short anchor versions, the intracellular tail (iT) of the proteins were not included. The aim was to gain insight into the function of the iT. The obtained results from this study can not be used to draw any conclusion about this. Further studies are therefore necessary to potentially get a better understanding of the function of the iT of the N-terminal transmembrane proteins, and whether this region is needed for successful anchoring of these proteins to the membrane.

This thesis can be regarded as a small step towards a future goal, which is the ability to co-express antigens and Invasin on the cell surface of *L. plantarum* that is used as a delivery vehicle for a mucosal vaccine. The *in-silico* analyses conducted as a part of this thesis provide an overview of the predicted N-terminal transmembrane proteins in *L. plantarum* WCFS1. This can be useful in potential future studies of these proteins.

6 References

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G. & Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology*, 37 (4): 420-423. doi: 10.1038/s41587-019-0036-z.

Aukrust, T. W., Brurberg, M. B. & Nes, I. F. (1995). Transformation of Lactobacillus by electroporation. *Methods Mol Biol*, 47: 201-8. doi: 10.1385/0-89603-310-4:201.

Axelsson, L. (2004). *Lactic Acid Bacteria, Microbiological and Functional Aspects*. . 3. ed.: Marcel Dekker, Inc.

Bermúdez-Humarán, L. G., Nouaille, S., Zilberfarb, V., Corthier, G., Gruss, A., Langella, P. & Issad, T. (2007). Effects of intranasal administration of a leptin-secreting Lactococcus lactis recombinant on food intake, body weight, and immune response of mice. *Applied and environmental microbiology*, 73 (16): 5300-5307. doi: 10.1128/AEM.00295-07.

Bush, K. & Bradford, P. A. (2016). β-Lactams and β-Lactamase Inhibitors: An Overview. *Cold Spring Harbor perspectives in medicine*, 6 (8): a025247. doi: 10.1101/cshperspect.a025247.

CDC. (2020). *Adjuvants and Vaccines*: Centers for Disease Control and Prevention (CDC). Available at: https://www.cdc.gov/vaccinesafety/concerns/adjuvants.html (accessed: 08.09.2021).

Chen, S., Zhong, J. & Huan, L. (2007). [Expression of human insulin in lactic acid bacteria and its oral administration in non-obese diabetic mice]. *Wei sheng wu xue bao = Acta microbiologica Sinica*, 47 6: 987-91.

Chevallier, B., Hubert, J.-C. & Kammerer, B. (1994). Determination of chromosome size and number of rrn loci in Lactobacillus plantarum by pulsed-field gel electrophoresis. *FEMS Microbiology Letters*, 120 (1-2): 51-56. doi: 10.1111/j.1574-6968.1994.tb07006.x.

Clark, M. A., Hirst, B. H. & Jepson, M. A. (1998). M-cell surface beta1 integrin expression and invasin-mediated targeting of Yersinia pseudotuberculosis to mouse Peyer's patch M cells. *Infection and immunity*, 66 (3): 1237-1243. doi: 10.1128/IAI.66.3.1237-1243.1998.

Cortes-Perez, N. G., Azevedo, V., Alcocer-González, J. M., Rodriguez-Padilla, C., Tamez-Guerra, R. S., Corthier, G., Gruss, A., Langella, P. & Bermúdez-Humarán, L. G. (2005). Cell-surface display of E7 antigen from human papillomavirus type-16 in Lactococcus lactis and in Lactobacillus plantarum using a new cell-wall anchor from lactobacilli. *J Drug Target*, 13 (2): 89-98. doi: 10.1080/10611860400024219.

Corthésy, B., Boris, S., Isler, P., Grangette, C. & Mercenier, A. (2005). Oral Immunization of Mice with Lactic Acid Bacteria Producing Helicobacter pylori Urease B Subunit Partially Protects against Challenge with Helicobacter felis. *The Journal of Infectious Diseases*, 192 (8): 1441-1449. doi: 10.1086/444425.

Critchley, R. J., Jezzard, S., Radford, K. J., Goussard, S., Lemoine, N. R., Grillot-Courvalin, C. & Vassaux, G. (2004). Potential therapeutic applications of recombinant, invasive E. coli. *Gene Ther*, 11 (15): 1224-33. doi: 10.1038/sj.gt.3302281.

Daniel, C., Roussel, Y., Kleerebezem, M. & Pot, B. (2011). Recombinant lactic acid bacteria as mucosal biotherapeutic agents. *Trends in biotechnology*, 29 (10): 499-508. doi: 10.1016/j.tibtech.2011.05.002.

de Vos, W. M. (2011). Systems solutions by lactic acid bacteria: from paradigms to practice. *Microb Cell Fact*, 10 Suppl 1 (Suppl 1): S2. doi: 10.1186/1475-2859-10-s1-s2.

Diep, D. B., Mathiesen, G., Eijsink, V. G. & Nes, I. F. (2009). Use of lactobacilli and their pheromone-based regulatory mechanism in gene expression and drug delivery. *Curr Pharm Biotechnol*, 10 (1): 62-73. doi: 10.2174/138920109787048571.

Dong, H., Sharma, M., Zhou, H.-X. & Cross, T. A. (2012). Glycines: role in α-helical membrane protein structures and a potential indicator of native conformation. *Biochemistry*, 51 (24): 4779-4789. doi: 10.1021/bi300090x.

Driessen, A. J. & Nouwen, N. (2008). Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem*, 77: 643-67. doi: 10.1146/annurev.biochem.77.061606.160747.

Eijsink, V. G., Axelsson, L., Diep, D. B., Håvarstein, L. S., Holo, H. & Nes, I. F. (2002). Production of class II bacteriocins by lactic acid bacteria; an example of biological warfare and communication. *Antonie Van Leeuwenhoek*, 81 (1-4): 639-54. doi: 10.1023/a:1020582211262.

Fakruddin, M., Mohammad Mazumdar, R., Bin Mannan, K. S., Chowdhury, A. & Hossain, M. N. (2012). Critical Factors Affecting the Success of Cloning, Expression, and Mass Production of Enzymes by Recombinant E. coli. *ISRN biotechnology*, 2013: 590587-590587. doi: 10.5402/2013/590587.

FAO/WHO. (2001). *Probiotics in food. Health and nutritional properties and guidelines for evaluation*

Fisher, K. & Phillips, C. (2009). The ecology, epidemiology and virulence of Enterococcus. *Microbiology*, 155 (6): 1749-1757. doi: https://doi.org/10.1099/mic.0.026385-0.

Fredriksen, L., Mathiesen, G., Sioud, M. & Eijsink, V. G. H. (2010). Cell wall anchoring of the 37-kilodalton oncofetal antigen by Lactobacillus plantarum for mucosal cancer vaccine delivery. *Applied and environmental microbiology*, 76 (21): 7359-7362. doi: 10.1128/AEM.01031-10.

Fredriksen, L., Kleiveland, C. R., Hult, L. T. O., Lea, T., Nygaard, C. S., Eijsink, V. G. H. & Mathiesen, G. (2012). Surface display of N-terminally anchored invasin by Lactobacillus plantarum activates NF-κB in monocytes. *Applied and environmental microbiology*, 78 (16): 5864-5871. doi: 10.1128/AEM.01227-12.

Gair, M. (2021). *Concepts of Biology*. 1st ed.

Gareau, M. G., Sherman, P. M. & Walker, W. A. (2010). Probiotics and the gut microbiota in intestinal health and disease. *Nat Rev Gastroenterol Hepatol*, 7 (9): 503-14. doi: 10.1038/nrgastro.2010.117.

Grangette, C., Müller-Alouf, H., Goudercourt, D., Geoffroy, M. C., Turneer, M. & Mercenier, A. (2001). Mucosal immune responses and protection against tetanus toxin after intranasal immunization with recombinant Lactobacillus plantarum. *Infect Immun*, 69 (3): 1547-53. doi: 10.1128/iai.69.3.1547-1553.2001.

Green, E. R. & Mecsas, J. (2016). Bacterial Secretion Systems: An Overview. *Microbiology spectrum*, 4 (1): 10.1128/microbiolspec.VMBF-0012-2015. doi: 10.1128/microbiolspec.VMBF-0012-2015.

Hamburger, Z. A., Brown, M. S., Isberg, R. R. & Bjorkman, P. J. (1999). Crystal Structure of Invasin: A Bacterial Integrin-Binding Protein. *Science*, 286 (5438): 291. doi: 10.1126/science.286.5438.291.

Henriques-Normark, B. & Tuomanen, E. I. (2013). The pneumococcus: epidemiology, microbiology, and pathogenesis. *Cold Spring Harbor perspectives in medicine*, 3 (7): a010215. doi: 10.1101/cshperspect.a010215.

Kajikawa, A., Nordone, S. K., Zhang, L., Stoeker, L. L., LaVoy, A. S., Klaenhammer, T. R. & Dean, G. A. (2011). Dissimilar properties of two recombinant Lactobacillus acidophilus strains displaying Salmonella FliC with different anchoring motifs.

Applied and environmental microbiology, 77 (18): 6587-6596. doi: 10.1128/AEM.05153-11.

Kandler, O. (1983). Carbohydrate metabolism in lactic acid bacteria. *Antonie van Leeuwenhoek*, 49 (3): 209-224. doi: 10.1007/BF00399499.

Kernéis, S., Bogdanova, A., Kraehenbuhl, J. P. & Pringault, E. (1997). Conversion by Peyer's patch lymphocytes of human enterocytes into M cells that transport bacteria. *Science*, 277 (5328): 949-52. doi: 10.1126/science.277.5328.949.

Kleerebezem, M., Boekhorst, J., van Kranenburg, R., Molenaar, D., Kuipers, O. P., Leer, R., Tarchini, R., Peters, S. A., Sandbrink, H. M., Fiers, M. W. E. J., et al. (2003). Complete genome sequence of Lactobacillus plantarum WCFS1. *Proceedings of the National Academy of Sciences*, 100 (4): 1990-1995. doi: 10.1073/pnas.0337704100.

Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes11Edited by F. Cohen. *Journal of Molecular Biology*, 305 (3): 567-580. doi: https://doi.org/10.1006/jmbi.2000.4315.

Kuczkowska, K., Mathiesen, G., Eijsink, V. G. H. & Øynebråten, I. (2015). Lactobacillus plantarum displaying CCL3 chemokine in fusion with HIV-1 Gag derived antigen causes increased recruitment of T cells. *Microbial Cell Factories*, 14 (1): 169. doi: 10.1186/s12934-015-0360-z.

Kuczkowska, K., Kleiveland, C. R., Minic, R., Moen, L. F., Øverland, L., Tjåland, R., Carlsen, H., Lea, T., Mathiesen, G. & Eijsink, V. G. H. (2016). Immunogenic Properties of Lactobacillus plantarum Producing Surface-Displayed Mycobacterium tuberculosis Antigens. *Applied and environmental microbiology*, 83 (2): e02782-16. doi: 10.1128/AEM.02782-16.

Kuczkowska, K., Myrbråten, I., Øverland, L., Eijsink, V. G. H., Follmann, F., Mathiesen, G. & Dietrich, J. (2017). Lactobacillus plantarum producing a Chlamydia trachomatis antigen induces a specific IgA response after mucosal booster immunization. *PloS one*, 12 (5): e0176401-e0176401. doi: 10.1371/journal.pone.0176401.

Kuczkowska, K., Copland, A., Øverland, L., Mathiesen, G., Tran, A. C., Paul, M. J., Eijsink, V. G. H. & Reljic, R. (2019). Inactivated Lactobacillus plantarum Carrying a Surface-Displayed Ag85B-ESAT-6 Fusion Antigen as a Booster Vaccine Against Mycobacterium tuberculosis Infection. *Frontiers in Immunology*, 10 (1588). doi: 10.3389/fimmu.2019.01588.

Kurien, B. T. & Scofield, R. H. (2015). Western blotting: an introduction. *Methods in molecular biology (Clifton, N.J.)*, 1312: 17-30. doi: 10.1007/978-1-4939-2694-7_5.

Lakshmana Rao, A. & Sankar, G. G. (2009). Caco-2 Cells: An Overview. *Asian Journal of Pharmaceutical Research and Health Care*: 16.

Lee, J. S., Poo, H., Han, D. P., Hong, S. P., Kim, K., Cho, M. W., Kim, E., Sung, M. H. & Kim, C. J. (2006). Mucosal immunization with surface-displayed severe acute respiratory syndrome coronavirus spike protein on Lactobacillus casei induces neutralizing antibodies in mice. *J Virol*, 80 (8): 4079-87. doi: 10.1128/jvi.80.8.4079-4087.2006.

Lycke, N. (2012). Recent progress in mucosal vaccine development: potential and limitations. *Nature Reviews Immunology*, 12 (8): 592-605. doi: 10.1038/nri3251.

Ma, S., Wang, L., Huang, X., Wang, X., Chen, S., Shi, W., Qiao, X., Jiang, Y., Tang, L., Xu, Y., et al. (2018). Oral recombinant Lactobacillus vaccine targeting the intestinal microfold cells and dendritic cells for delivering the core neutralizing epitope of porcine epidemic diarrhea virus. *Microb Cell Fact*, 17 (1): 20. doi: 10.1186/s12934-018-0861-7.

Mathiesen, G. (2004). *Studies of genes and promoters involved in peptide-pheromone regulated bacteriocin production in Lactobacillus and their use in gene expression systems*. Ås, Norway: Norwegian University of Life Sciences(NMBU).

Mathiesen, G., Sveen, A., Piard, J. C., Axelsson, L. & Eijsink, V. G. (2008). Heterologous protein secretion by Lactobacillus plantarum using homologous signal peptides. *J Appl Microbiol*, 105 (1): 215-26. doi: 10.1111/j.1365-2672.2008.03734.x.

Mathiesen, G., Sveen, A., Brurberg, M. B., Fredriksen, L., Axelsson, L. & Eijsink, V. G. (2009). Genome-wide analysis of signal peptide functionality in Lactobacillus plantarum WCFS1. *BMC genomics*, 10: 425-425. doi: 10.1186/1471-2164-10-425.

Mathiesen, G., Øverland, L., Kuczkowska, K. & Eijsink, V. G. H. (2020). Anchoring of heterologous proteins in multiple Lactobacillus species using anchors derived from Lactobacillus plantarum. *Scientific Reports*, 10 (1): 9640. doi: 10.1038/s41598-020-66531-7.

Michon, C., Kuczkowska, K., Langella, P., Eijsink, V. G. H., Mathiesen, G. & Chatel, J.-M. (2015). Erratum to: Surface display of an anti-DEC-205 single chain Fv fragment in Lactobacillus plantarum increases internalization and plasmid transfer to dendritic cells in vitro and in vivo. *Microbial Cell Factories*, 14 (1): 184. doi: 10.1186/s12934-015-0366-6.

Michon, C., Langella, P., Eijsink, V. G. H., Mathiesen, G. & Chatel, J. M. (2016). Display of recombinant proteins at the surface of lactic acid bacteria: strategies and applications. *Microbial cell factories*, 15: 70-70. doi: 10.1186/s12934-016-0468-9.

Minic, R., Gavrovic-Jankulovic, M., Petrusic, V., Zivkovic, I., Eijsink, V. G., Dimitrijevic, L. & Mathiesen, G. (2015). Effects of orally applied Fes p1-displaying L. plantarum WCFS1 on Fes p1 induced allergy in mice. *J Biotechnol*, 199: 23-8. doi: 10.1016/j.jbiotec.2015.01.028.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, Gustavo A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., et al. (2020). Pfam: The protein families database in 2021. *Nucleic Acids Research*, 49 (D1): D412-D419. doi: 10.1093/nar/gkaa913.

Mobergslien, A., Vasovic, V., Mathiesen, G., Fredriksen, L., Westby, P., Eijsink, V. G. H., Peng, Q. & Sioud, M. (2015). Recombinant Lactobacillus plantarum induces immune responses to cancer testis antigen NY-ESO-1 and maturation of dendritic cells. *Human vaccines & immunotherapeutics*, 11 (11): 2664-2673. doi: 10.1080/21645515.2015.1056952.

Mohamadzadeh, M., E±¡ng, T., Sandwick, S., Hoover, T. & Klaenhammer, T. (2009). Dendritic cell targeting of Bacillus anthracis protective antigen expressed by Lactobacillus acidophilus protects mice from lethal challenge. *Proceedings of the National Academy of Sciences*, 106: 4331 - 4336.

Nguyen, H.-M., Pham, M.-L., Stelzer, E. M., Plattner, E., Grabherr, R., Mathiesen, G., Peterbauer, C. K., Haltrich, D. & Nguyen, T.-H. (2019). Constitutive expression and cell-surface display of a bacterial β-mannanase in Lactobacillus plantarum. *Microbial cell factories*, 18 (1): 76-76. doi: 10.1186/s12934-019-1124-y.

Norton, P. M., Wells, J. M., Brown, H. W. G., Macpherson, A. M. & Le Page, R. W. F. (1997). Protection against tetanus toxin in mice nasally immunized with recombinant Lactococcus lactis expressing tetanus toxin fragment C. *Vaccine*, 15 (6): 616-619. doi: https://doi.org/10.1016/S0264-410X(96)00241-1.

Nowakowski, A. B., Wobig, W. J. & Petering, D. H. (2014). Native SDS-PAGE: high resolution electrophoretic separation of proteins with retention of native properties including bound metal ions. *Metallomics : integrated biometal science*, 6 (5): 1068-1078. doi: 10.1039/c4mt00033a.

Nygaard, C. S. (2011). *Secretion and anchoring of invasin from Yersinia pseudotuberculosis in Lactobacillus plantarum.*: Norwegian University of Life Sciences, Department of Chemistry, Biotechnology and Food Science, Ås.

Oh, S.-H., Kim, S.-H., Jeon, J.-H., Kim, E. B., Lee, N.-K., Beck, S., Choi, Y.-J. & Kang, S.-K. (2021). Cytoplasmic expression of a model antigen with M Cell-Targeting moiety in lactic acid bacteria and implication of the mechanism as a mucosal vaccine via oral route. *Vaccine*, 39 (30): 4072-4081. doi: https://doi.org/10.1016/j.vaccine.2021.06.010.

Owen, J. L., Sahay, B. & Mohamadzadeh, M. (2013). New generation of oral mucosal vaccines targeting dendritic cells. *Curr Opin Chem Biol*, 17 (6): 918-24. doi: 10.1016/j.cbpa.2013.06.013.

Parham, P. (2015). *The Immune System*. Fourth edition ed.: Garland Science.

Pham, M.-L., Tran, A.-M., Kittibunchakul, S., Nguyen, T.-T., Mathiesen, G. & Nguyen, T.-H. (2019). Immobilization of β-Galactosidases on the Lactobacillus Cell Surface Using the Peptidoglycan-Binding Motif LysM. *Catalysts (Basel, Switzerland)*, 9 (5): 443-443. doi: 10.3390/catal9050443.

Pham, M. L., Tran, A. M., Mathiesen, G., Nguyen, H. M. & Nguyen, T. H. (2020). Cell Wall Anchoring of a Bacterial Chitosanase in Lactobacillus plantarum Using a Food-Grade Expression System and Two Versions of an LP TG Anchor. *Int J Mol Sci*, 21 (11). doi: 10.3390/ijms21113773.

Prabudiansyah, I., Kusters, I. & Driessen, A. J. M. (2015). In Vitro Interaction of the Housekeeping SecA1 with the Accessory SecA2 Protein of Mycobacterium tuberculosis. *PloS one*, 10 (6): e0128788-e0128788. doi: 10.1371/journal.pone.0128788.

Robinson, K., Chamberlain, L. M., Lopez, M. C., Rush, C. M., Marcotte, H., Le Page, R. W. F. & Wells, J. M. (2004). Mucosal and Cellular Immune Responses Elicited by Recombinant Lactococcus lactis Strains Expressing Tetanus Toxin Fragment C. *Infection and Immunity*, 72 (5): 2753-2761. doi: 10.1128/IAI.72.5.2753-2761.2004.

Rud, I., Jensen, P. R., Naterstad, K. & Axelsson, L. (2006). A synthetic promoter library for constitutive gene expression in Lactobacillus plantarum. *Microbiology (Reading)*, 152 (Pt 4): 1011-1019. doi: 10.1099/mic.0.28599-0.

Sainsbury, S., Bird, L., Rao, V., Shepherd, S. M., Stuart, D. I., Hunter, W. N., Owens, R. J. & Ren, J. (2011). Crystal structures of penicillin-binding protein 3 from Pseudomonas aeruginosa: comparison of native and antibiotic-bound forms. *Journal of molecular biology*, 405 (1): 173-184. doi: 10.1016/j.jmb.2010.10.024.

Shreiner, A. B., Kao, J. Y. & Young, V. B. (2015). The gut microbiome in health and in disease. *Current opinion in gastroenterology*, 31 (1): 69-75. doi: 10.1097/MOG.0000000000000139.

Siezen, R. J. & van Hylckama Vlieg, J. E. T. (2011). Genomic diversity and versatility of Lactobacillus plantarum, a natural metabolic engineer. *Microbial Cell Factories*, 10 (1): S3. doi: 10.1186/1475-2859-10-S1-S3.

Solberg, E. (2015). *Targeting of Lactobacillus to M-Cells to Improve its Potential as Vaccine Delivery Vector*: Norwegian University of Life Sciences, Department of Chemistry, Biotechnology and Food Science, Ås.

Sonnhammer, E., Heijne, G. & Krogh, A. (1998). A Hidden Markov Model for Predicting Transmembrane Helices in Protein Sequences. *Proceedings. International Conference on Intelligent Systems for Molecular Biology*, 6: 175-82.

Sørvig, E., Grönqvist, S., Naterstad, K., Mathiesen, G., Eijsink, V. G. H. & Axelsson, L. (2003). Construction of vectors for inducible gene expression in Lactobacillus sakei

and L. plantarum. *FEMS Microbiology Letters*, 229 (1): 119-126. doi: 10.1016/S0378-1097(03)00798-5.

Sørvig, E., Mathiesen, G., Naterstad, K., Eijsink, V. G. H. & Axelsson, L. (2005). High-level, inducible gene expression in Lactobacillus sakei and Lactobacillus plantarum using versatile expression vectors. *Microbiology*, 151 (Pt 7): 2439-2449. doi: 10.1099/mic.0.28084-0.

Tjalsma, H., Antelmann, H., Jongbloed, J. D., Braun, P. G., Darmon, E., Dorenbos, R., Dubois, J. Y., Westers, H., Zanen, G., Quax, W. J., et al. (2004). Proteomics of protein secretion by Bacillus subtilis: separating the "secrets" of the secretome. *Microbiol Mol Biol Rev*, 68 (2): 207-33. doi: 10.1128/mmbr.68.2.207-233.2004.

van Roosmalen, M. L., Geukens, N., Jongbloed, J. D. H., Tjalsma, H., Dubois, J.-Y. F., Bron, S., van Dijl, J. M. & Anné, J. (2004). Type I signal peptidases of Gram-positive bacteria. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1694 (1): 279-297. doi: https://doi.org/10.1016/j.bbamcr.2004.05.006.

Visweswaran, G. R. R., Leenhouts, K., van Roosmalen, M., Kok, J. & Buist, G. (2014). Exploiting the peptidoglycan-binding motif, LysM, for medical and industrial applications. *Applied microbiology and biotechnology*, 98 (10): 4331-4345. doi: 10.1007/s00253-014-5633-7.

Walker, M. J., Barnett, T. C., McArthur, J. D., Cole, J. N., Gillen, C. M., Henningham, A., Sriprakash, K. S., Sanderson-Smith, M. L. & Nizet, V. (2014). Disease manifestations and pathogenic mechanisms of Group A Streptococcus. *Clinical microbiology reviews*, 27 (2): 264-301. doi: 10.1128/CMR.00101-13.

Wells, J. (2011). Mucosal vaccination and therapy with genetically modified lactic acid bacteria. *Annu Rev Food Sci Technol*, 2: 423-45. doi: 10.1146/annurev-food-022510-133640.

Wells, J. M. & Mercenier, A. (2008). Mucosal delivery of therapeutic and prophylactic molecules using lactic acid bacteria. *Nature reviews. Microbiology*, 6 (5): 349-362. doi: 10.1038/nrmicro1840.

Woolfson, D. N. & Williams, D. H. (1990). The influence of proline residues on α-helical structure. *FEBS Letters*, 277 (1-2): 185-188. doi: https://doi.org/10.1016/0014-5793(90)80839-B.

Wyszyńska, A., Kobierecka, P., Bardowski, J. & Jagusztyn-Krynicka, E. K. (2015). Lactic acid bacteria--20 years exploring their potential as live vectors for mucosal vaccination. *Applied microbiology and biotechnology*, 99 (7): 2967-2977. doi: 10.1007/s00253-015-6498-0.

Zheng, J., Wittouck, S., Salvetti, E., Franz, C., Harris, H., Mattarelli, P., O'Toole, P., Pot, B., Vandamme, P., Walter, J., et al. (2020). A taxonomic note on the genus Lactobacillus: Description of 23 novel genera, emended description of the genus Lactobacillus Beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. *International journal of systematic and evolutionary microbiology*.

Zhou, M., Boekhorst, J., Francke, C. & Siezen, R. J. (2008). LocateP: Genome-scale subcellular-location predictor for bacterial proteins. *BMC Bioinformatics*, 9 (1): 173. doi: 10.1186/1471-2105-9-173.

Øverland, L. (2013). *Secretion and anchoring of proteins in Lactobacillus plantarum: Studies of a dendritic cell-targeted Mycobacterium tuberculosis antigen.* : Norwegian University of Life Sciences, Department of Chemistry, Biotechnology and Food Science, Ås.

7 Appendix

Explanation of color coding in table A-1 and A-2:

Proteins are marked in green if: They are predicted to contain a transmembrane helix and their probability of containing a cleavage site are below 30%.

Proteins are marked in yellow if:

- They are predicted to contain a transmembrane helix, and their probability of containing a cleavage site is between 30 and 50 %.
- They are predicted to contain a domain that would overlap with the part of the protein that would be used as the anchor sequence.

Proteins are marked in red if:

- They are not predicted to have a helix by TMHMM
- The probability for a signal peptide is above 0.50 (50%)
- They are predicted to have a $N_{out} – C_{in}$ orientation (in some proteins, the N-terminal region is extracellular, while the C-terminal region following the helix is intracellular).

Table A-1. List of predicted N-terminal transmembrane proteins in *L. plantarum* WCFS1. For the proteins that were predicted to have a helix in TMHMM, the length and localization of the helix is shown. The length of the intracellular tail (iT) preceding the helix is also included in the table. NB: hypothetical proteins and prophage proteins that were predicted to be N-terminally anchored by LocateP2, are not included in the table since they did not fulfill the preset criteria.

Protein nr.	Nr. of protein *L.plantarum* (N-terminal trans-membrane protein)	Identifier	Total protein length (aa)	Nr. of predicted helices	Length of intra-cellular tail (iT) preceding the helix	Helix length	Helix localization (from start aa to end aa)
1	32	gi|380031133|ref|YP_004888124.1| regulator of two-component system, YycH family [Lactobacillus plantarum WCFS1]	441	1	8	23	9-31

2	33	gi|380031134|ref|YP_004888125.1| regulator of two-component system, YycI family [Lactobacillus plantarum WCFS1]	284	1	6	20	7-26
3	35	gi|380031136|ref|YP_004888127.1| serine protease HtrA [Lactobacillus plantarum WCFS1]	420	1	4	23	5-27
4	113	gi|380031214|ref|YP_004888205.1| linoleic acid isomerase [Lactobacillus plantarum WCFS1]	564	0		No predicted helix	
5	237	gi|380031338|ref|YP_004888329.1| transcriptional attenuator, cell envelope-related, LytR family [Lactobacillus plantarum WCFS1]	343	1	11	20	12-31
6	261	gi|380031362|ref|YP_004888353.1| spermidine/putrescine ABC transporter,substrate binding protein [Lactobacillus plantarum WCFS1]	356	1	4	18	5-22
7*	276	gi|380031377|ref|YP_004888368.1| membrane protein [Lactobacillus plantarum WCFS1]	300	1	1	23	2-24
8	313	gi|380031414|ref|YP_004888405.1| extracellular protein, lysine-rich [Lactobacillus plantarum WCFS1]	595	1	11	23	12-34
9	351	gi|380031453|ref|YP_004888444.1| bacteriocin ABC transporter, accessory factor PlnH [Lactobacillus plantarum WCFS1]	458	1	18	23	19-41
10	425	gi|380031529|ref|YP_004888520.1| sortase A [Lactobacillus plantarum WCFS1]	234	1	11	18	12-29
11	444	gi|380031548|ref|YP_004888539.1| cell surface protein, ErfK/YbiS/YcfS/YnhG family [Lactobacillus plantarum WCFS1]	464	1	6	23	7-29
12	450	gi|380031554|ref|YP_004888545.1| septum formation initiator [Lactobacillus plantarum WCFS1]	132	1	38	20	39-58
13	497	gi|380031601|ref|YP_004888592.1| transcriptional attenuator, cell envelope-related, LytR family [Lactobacillus plantarum WCFS1]	397	1	32	23	33-55
14	501	gi|380031605|ref|YP_004888596.1| extracellular zinc metalloproteinase, M10 family [Lactobacillus plantarum WCFS1]	230	1	6	23	7-29
15	517	gi|380031621|ref|YP_004888612.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	293	1	4	18	5-22
16	583	gi|380031687|ref|YP_004888678.1| DNA entry nuclease [Lactobacillus plantarum WCFS1]	325	0		No predicted helix	
17	696	gi|380031800|ref|YP_004888791.1| cell surface protein, YbbR-like family [Lactobacillus plantarum WCFS1]	325	1	11	18	12-29

18	945	gi|380032053|ref|YP_004889044.1| cell wall hydrolase/muramidase [Lactobacillus plantarum WCFS1]	213	1	25	20	26-45
19	960	gi|380032068|ref|YP_004889059.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	294	1	12	23	13-35
20	962	gi|380032070|ref|YP_004889061.1| lysozyme [Lactobacillus plantarum WCFS1]	258	1	20	23	21-43
21	963	gi|380032071|ref|YP_004889062.1| extracellular protein, membrane-bound [Lactobacillus plantarum WCFS1]	121	0		No predicted helix	
22	969	gi|380032077|ref|YP_004889068.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	288	1	4	23	5-27
23	974	gi|380032082|ref|YP_004889073.1| serine-type D-Ala-D-Ala carboxypeptidase [Lactobacillus plantarum WCFS1]	391	1	6	20	7-26
24	1003	gi|380032111|ref|YP_004889102.1| priming glycosyltransferase, polyprenyl glycosylphosphotransferase [Lactobacillus plantarum WCFS1]	225	1	43	22	44-65
25	1016	gi|380032126|ref|YP_004889117.1| UDP-galactopyranose mutase [Lactobacillus plantarum WCFS1]	377	1	6	19	7-25
26	1026	gi|380032138|ref|YP_004889129.1| priming glycosyltransferase,undecaprenyl-phosphate beta-glucosephosphotransferase [Lactobacillus plantarum WCFS1]	225	1	32	23	33-55
27	1035	gi|380032147|ref|YP_004889138.1| extracellular protein, NlpC/P60 family,gamma-D-glutamate-meso-diaminopimelate muropeptidase [Lactobacillus plantarum WCFS1]	243	1	4	18	5-22
28	1087	gi|380032199|ref|YP_004889190.1| methyltransferase, SAM-dependent [Lactobacillus plantarum WCFS1]	197	0		No predicted helix	
29	1098	gi|380032210|ref|YP_004889201.1| glycerol-3-phosphate ABC transporter, substrate binding protein [Lactobacillus plantarum WCFS1]	450	1	6	23	7-29
30**	1122	gi|380032234|ref|YP_004889225.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	112	1	3	23	4-26
31	1131	gi|380032243|ref|YP_004889234.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	356	1	22	23	23-45
32	1140	gi|380032252|ref|YP_004889243.1| arylsulfate sulfotransferase [Lactobacillus plantarum WCFS1]	579	1	6	19	7-25

33	1155	gi|380032269|ref|YP_004889260.1| cell surface protein, ErfK family [Lactobacillus plantarum WCFS1]	203	1	6	23	7-29
34	1162	gi|380032276|ref|YP_004889267.1| transpeptidase-transglycosylase (penicillin binding protein 2A) [Lactobacillus plantarum WCFS1]	709	1	43	19	44-62
35	1168	gi|380032282|ref|YP_004889273.1| carboxyterminal processing proteinase, S41 family, membrane-anchored [Lactobacillus plantarum WCFS1]	323	1	6	18	7-24
36**	1198	gi|380032312|ref|YP_004889303.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	107	1	4	23	5-27
37	1233	gi|380032347|ref|YP_004889338.1| molybdopterin biosynthesis protein MoeB [Lactobacillus plantarum WCFS1]	344	1	20	20	21-40
38	1288	gi|380032402|ref|YP_004889393.1| aminodeoxychorismate lyase [Lactobacillus plantarum WCFS1]	401	1	43	23	44-66
39	1295	gi|380032409|ref|YP_004889400.1| transpeptidase (penicillin binding protein 2B) [Lactobacillus plantarum WCFS1]	678	1	19	23	20-42
40**	1302	gi|380032416|ref|YP_004889407.1| rhodanese family protein [Lactobacillus plantarum WCFS1]	137	1	9	23	10-32
41	1305	gi|380032419|ref|YP_004889410.1| tRNA isopentenylpyrophosphate transferase [Lactobacillus plantarum WCFS1]	311	0		No predicted helix	
42	1338	gi|380032452|ref|YP_004889443.1| methionyl-tRNA formyltransferase [Lactobacillus plantarum WCFS1]	317	0		No predicted helix	
43	1421	gi|380032534|ref|YP_004889525.1| dipeptidase [Lactobacillus plantarum WCFS1]	479	0		No predicted helix	
44	1449	gi|380032562|ref|YP_004889553.1| D-methionine ABC transporter, substrate binding protein [Lactobacillus plantarum WCFS1]	285	1	4	23	5-27
45	1454	gi|380032567|ref|YP_004889558.1| transpeptidase-transglycosylase (penicillin binding protein 1A) [Lactobacillus plantarum WCFS1]	767	1	28	23	29-51
46	1540	gi|380032656|ref|YP_004889647.1| carboxy-terminal proteinase, S41 family,peptidoglycan-bound [Lactobacillus plantarum WCFS1]	492	1	22	23	23-45
47	1542	gi|380032658|ref|YP_004889649.1| extracellular protein [Lactobacillus plantarum WCFS1]	214	1	20	23	21-43

48	1543	gi|380032659|ref|YP_004889650.1| extracellular lipase/acylhydrolase with GDSL-like motif [Lactobacillus plantarum WCFS1]	314	1	6	23	7-29
49**	1558	gi|380032674|ref|YP_004889665.1| extracellular protein, with LysM peptidoglycan binding domain [Lactobacillus plantarum WCFS1]	210	1	43	23	44-66
50	1594	gi|380032710|ref|YP_004889701.1| phosphohydrolase [Lactobacillus plantarum WCFS1]	201	0		No predicted helix	
51	1637	gi|380032751|ref|YP_004889742.1| short-chain dehydrogenase [Lactobacillus plantarum WCFS1]	241	0		No predicted helix	
52	1644	gi|380032758|ref|YP_004889749.1| N-acetylmuramoyl-L-alanine amidase [Lactobacillus plantarum WCFS1]	282	1	6	23	7-29
53	1670	gi|380032787|ref|YP_004889778.1| poly(glycerophosphate chain) D-alanine transfer protein DltD [Lactobacillus plantarum WCFS1]	425	1	6	23	7-29
54	1718	gi|380032838|ref|YP_004889829.1| transcriptional attenuator, cell envelope-related, LytR family [Lactobacillus plantarum WCFS1]	394	1	20	23	21-43
55	1743	gi|380032863|ref|YP_004889854.1| polysaccharide biosynthesis polyprenyl glycosylphosphotransferase, priming glycosyltransferase [Lactobacillus plantarum WCFS1]	221	1	39	22	40-61
56	1767	gi|380032887|ref|YP_004889878.1| ComE operon protein 1, DNA receptor [Lactobacillus plantarum WCFS1]	241	1	11	20	12-31
57	1768	gi|380032888|ref|YP_004889879.1| extracellular protease Lon, peptidase S16 family, membrane-bound [Lactobacillus plantarum WCFS1]	348	1	10	23	11-33
58**	1787	gi|380032907|ref|YP_004889898.1| membrane protein [Lactobacillus plantarum WCFS1]	168	1	3	19	4-21
59	1821	gi|380032941|ref|YP_004889932.1| transpeptidase, penicillin binding protein 2B [Lactobacillus plantarum WCFS1]	713	1	20	23	21-43
60	1822	gi|380032942|ref|YP_004889933.1| cell division protein FtsL [Lactobacillus plantarum WCFS1]	131	1	49	23	50-72
61	1859	gi|380032979|ref|YP_004889970.1| bacterial type II secretion/trafficking system protein [Lactobacillus plantarum WCFS1]	194	1	37	23	38-60
62	1860	gi|380032980|ref|YP_004889971.1| bacterial type II secretion/trafficking system extracellular protein [Lactobacillus plantarum WCFS1]	157	1	6	23	7-29

63	1902	gi|380033023|ref|YP_004890014.1| metal dependent phosphohydrolase, HD family [Lactobacillus plantarum WCFS1]	519	1	1	23	2-24
64	1916	gi|380033037|ref|YP_004890028.1| cell shape determining protein MreC [Lactobacillus plantarum WCFS1]	281	1	11	19	12-30
65	1925	gi|380033046|ref|YP_004890037.1| septation ring formation regulator [Lactobacillus plantarum WCFS1]	573	1	2	23	3-25
66	1936	gi|380033057|ref|YP_004890048.1| beta-lactamase, class A [Lactobacillus plantarum WCFS1]	376	1	57	23	58-80
67**	1958	gi|380033079|ref|YP_004890070.1| H(+)-transporting two-sector ATPase, B subunit [Lactobacillus plantarum WCFS1]	171	1	14	22	15-36
68	2078	gi|380033199|ref|YP_004890190.1| amino acid regulated enzyme with ACT domain [Lactobacillus plantarum WCFS1]	133	0		No predicted helix	
69	2135	gi|380033258|ref|YP_004890249.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	254	1	6	20	7-26
70	2230	gi|380033353|ref|YP_004890344.1| diguanylate cyclase/phosphodiesterase, EAL domain [Lactobacillus plantarum WCFS1]	274	1	9	23	10-32
71	2250	gi|380033373|ref|YP_004890364.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	269	1	4	23	5-27
72	2313	gi|380033435|ref|YP_004890426.1| glycosyl hydrolase family protein [Lactobacillus plantarum WCFS1]	432	0		No predicted helix	
73	2314	gi|380033436|ref|YP_004890427.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	256	1	12	23	13-35
74	2343	gi|380033465|ref|YP_004890456.1| short-chain dehydrogenase [Lactobacillus plantarum WCFS1]	247	0		No predicted helix	
75	2353	gi|380033475|ref|YP_004890466.1| signal peptidase I [Lactobacillus plantarum WCFS1]	205	1	6	23	7-29
76	2354	gi|380033476|ref|YP_004890467.1| signal peptidase I [Lactobacillus plantarum WCFS1]	207	1	6	23	7-29
77	2376	gi|380033498|ref|YP_004890489.1| aldo/keto reductase family protein [Lactobacillus plantarum WCFS1]	342	0		No predicted helix	

78	2424	gi|380033548|ref|YP_004890539.1| lipase/esterase, subfamily of SGNH-hydrolases [Lactobacillus plantarum WCFS1]	233	1	6	23	7-29
79	2426	gi|380033550|ref|YP_004890541.1| TetR family transcriptional regulator [Lactobacillus plantarum WCFS1]	207	1	12	23	13-35
80	2466	gi|380033594|ref|YP_004890585.1| extracellular transglycosylase, with LysM peptidoglycan binding domain [Lactobacillus plantarum WCFS1]	204	1	3	23	4-26
81	2471	gi|380033599|ref|YP_004890590.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	174	1	3	20	4-23
82	2477	gi|380033605|ref|YP_004890596.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	139	1	4	23	5-27
83	2478	gi|380033606|ref|YP_004890597.1| short-chain dehydrogenase [Lactobacillus plantarum WCFS1]	247	0		No predicted helix	
84	2516	gi|380033646|ref|YP_004890637.1| extracellular protein [Lactobacillus plantarum WCFS1]	274	1	3	23	4-26
85	2527	gi|380033657|ref|YP_004890648.1| TetR family transcriptional regulator [Lactobacillus plantarum WCFS1]	193	0		No predicted helix	
86	2530	gi|380033660|ref|YP_004890651.1| lysozyme/muramidase, glycoside hydrolase family 25 [Lactobacillus plantarum WCFS1]	860	0		No predicted helix	
87	2533	gi|380033663|ref|YP_004890654.1| short-chain dehydrogenase [Lactobacillus plantarum WCFS1]	242	0		No predicted helix	
88	2553	gi|380033683|ref|YP_004890674.1| cell surface protein, CscC family [Lactobacillus plantarum WCFS1]	750	1	6	23	7-29
89	2567	gi|380033697|ref|YP_004890688.1| extracellular protein [Lactobacillus plantarum WCFS1]	123	1	3	23	4-26
90	2570	gi|380033700|ref|YP_004890691.1| PTS system transporter subunit IIB [Lactobacillus plantarum WCFS1]	100	0		No predicted helix	
91	2589	gi|380033727|ref|YP_004890718.1| cell surface protein, membrane-anchored [Lactobacillus plantarum WCFS1]	345	1	4	18	5-22
92	2615	gi|380033753|ref|YP_004890744.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	277	1	6	23	7-29

93	2670	gi|380033808|ref|YP_004890799.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	160	1	9	21	10-30
94	2757	gi|380033901|ref|YP_004890892.1| cell surface hydrolase, membrane-bound [Lactobacillus plantarum WCFS1]	312	1	6	23	7-29
95	2805	gi|380033950|ref|YP_004890941.1| cell surface protein, CscB family [Lactobacillus plantarum WCFS1]	230	1	6	23	7-29
96	2806	gi|380033951|ref|YP_004890942.1| cell surface protein, CscB family [Lactobacillus plantarum WCFS1]	190	0		No predicted helix	
97	2845	gi|380033990|ref|YP_004890981.1| shikimate 5-dehydrogenase [Lactobacillus plantarum WCFS1]	294	0		No predicted helix	
98	2956	gi|380034102|ref|YP_004891093.1| sorbitol operon activator [Lactobacillus plantarum WCFS1]	171	1	3	23	4-26
99	2986	gi|380034132|ref|YP_004891123.1| sorbitol operon activator [Lactobacillus plantarum WCFS1]	168	1	1	20	2-21
100	3002	gi|380034148|ref|YP_004891139.1| signal peptidase I [Lactobacillus plantarum WCFS1]	195	1	6	23	7-29

*: This protein is annotated as a hypothetical membrane protein in UniProt.

**: These proteins are predicted to have an $N_{out} – C_{in}$ orientation by TMHMM.

Table A-2. List of predicted N-terminal transmembrane proteins and their signal peptide probabilities (predicted in SignalP).

Protein nr.	Nr. of protein *L.plantarum* (N-terminal transmembrane anchor)	Identifier	Signal peptide (Sec/SPI)	TAT Signal peptide	Lipoprotein signal sequence (Sec/SPII)
1	32	gi|380031133|ref|YP_004888124.1| regulator of two-component system, YycH family [Lactobacillus plantarum WCFS1]	0.1034	0.0022	0.016
2	33	gi|380031134|ref|YP_004888125.1| regulator of two-component system, YycI family [Lactobacillus plantarum WCFS1]	0.5916	0.0245	0.0831
3	35	gi|380031136|ref|YP_004888127.1| serine protease HtrA [Lactobacillus plantarum WCFS1]	0.7446	0.0086	0.0598
4	113	gi|380031214|ref|YP_004888205.1| linoleic acid isomerase [Lactobacillus plantarum WCFS1]	0.4505	0.0031	0.0215
5	237	gi|380031338|ref|YP_004888329.1| transcriptional attenuator, cell envelope-related, LytR family [Lactobacillus plantarum WCFS1]	0.6679	0.0113	0.0037
6	261	gi|380031362|ref|YP_004888353.1| spermidine/putrescine ABC transporter,substrate binding protein [Lactobacillus plantarum WCFS1]	0.726	0.0014	0.258
7*	276	gi|380031377|ref|YP_004888368.1| membrane protein [Lactobacillus plantarum WCFS1]	0.1319	0.0006	0.0646
8	313	gi|380031414|ref|YP_004888405.1| extracellular protein, lysine-rich [Lactobacillus plantarum WCFS1]	0.6411	0.0266	0.2113
9	351	gi|380031453|ref|YP_004888444.1| bacteriocin ABC transporter, accessory factor PlnH [Lactobacillus plantarum WCFS1]	0.0478	0.0014	0.0084
10	425	gi|380031529|ref|YP_004888520.1| sortase A [Lactobacillus plantarum WCFS1]	0.476	0.0055	0.0126
11	444	gi|380031548|ref|YP_004888539.1| cell surface protein, ErfK/YbiS/YcfS/YnhG family [Lactobacillus plantarum WCFS1]	0.3196	0.0091	0.0291
12	450	gi|380031554|ref|YP_004888545.1| septum formation initiator [Lactobacillus plantarum WCFS1]	0.0158	0.0058	0.0068
13	497	gi|380031601|ref|YP_004888592.1| transcriptional attenuator, cell envelope-related, LytR family [Lactobacillus plantarum WCFS1]	0.0458	0.001	0.0223
14	501	gi|380031605|ref|YP_004888596.1| extracellular zinc metalloproteinase, M10 family [Lactobacillus plantarum WCFS1]	0.857	0.0155	0.0799
15	517	gi|380031621|ref|YP_004888612.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	0.0375	0.0029	0.0115

16	583	gi|380031687|ref|YP_004888678.1| DNA entry nuclease [Lactobacillus plantarum WCFS1]	0.1711	0.0044	0.04
17	696	gi|380031800|ref|YP_004888791.1| cell surface protein, YbbR-like family [Lactobacillus plantarum WCFS1]	0.756	0.0055	0.0541
18	945	gi|380032053|ref|YP_004889044.1| cell wall hydrolase/muramidase [Lactobacillus plantarum WCFS1]	0.0379	0.0012	0.0109
19	960	gi|380032068|ref|YP_004889059.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	0.376	0.1693	0.0155
20	962	gi|380032070|ref|YP_004889061.1| lysozyme [Lactobacillus plantarum WCFS1]	0.0499	0.0013	0.0078
21	963	gi|380032071|ref|YP_004889062.1| extracellular protein, membrane-bound [Lactobacillus plantarum WCFS1]	0.5337	0.0691	0.3131
22	969	gi|380032077|ref|YP_004889068.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	0.7953	0.0121	0.1428
23	974	gi|380032082|ref|YP_004889073.1| serine-type D-Ala-D-Ala carboxypeptidase [Lactobacillus plantarum WCFS1]	0.7621	0.0661	0.088
24	1003	gi|380032111|ref|YP_004889102.1| priming glycosyltransferase, polyprenyl glycosylphosphotransferase [Lactobacillus plantarum WCFS1]	0.0123	0.0002	0.0018
25	1016	gi|380032126|ref|YP_004889117.1| UDP-galactopyranose mutase [Lactobacillus plantarum WCFS1]	0.0422	0.0094	0.0059
26	1026	gi|380032138|ref|YP_004889129.1| priming glycosyltransferase,undecaprenyl-phosphate beta-glucosephosphotransferase [Lactobacillus plantarum WCFS1]	0.0077	0.0004	0.0027
27	1035	gi|380032147|ref|YP_004889138.1| extracellular protein, NlpC/P60 family,gamma-D-glutamate-meso-diaminopimelate muropeptidase [Lactobacillus plantarum WCFS1]	0.9862	0.0036	0.0057
28	1087	gi|380032199|ref|YP_004889190.1| methyltransferase, SAM-dependent [Lactobacillus plantarum WCFS1]	0.5465	0.0339	0.0269
29	1098	gi|380032210|ref|YP_004889201.1| glycerol-3-phosphate ABC transporter, substrate binding protein [Lactobacillus plantarum WCFS1]	0.9231	0.0159	0.0238
30**	1122	gi|380032234|ref|YP_004889225.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	0.1385	0.0005	0.8183
31	1131	gi|380032243|ref|YP_004889234.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	0.001	0.0003	0.0236
32	1140	gi|380032252|ref|YP_004889243.1| arylsulfate sulfotransferase [Lactobacillus plantarum WCFS1]	0.4199	0.0043	0.1139
33	1155	gi|380032269|ref|YP_004889260.1| cell surface protein, ErfK family [Lactobacillus plantarum WCFS1]	0.6265	0.0342	0.0233
34	1162	gi|380032276|ref|YP_004889267.1| transpeptidase-transglycosylase (penicillin binding protein 2A) [Lactobacillus plantarum WCFS1]	0.0064	0.0002	0.0022

35	1168	gi|380032282|ref|YP_004889273.1| carboxyterminal processing proteinase, S41 family, membrane-anchored [Lactobacillus plantarum WCFS1]	0.7104	0.0387	0.1143
36**	1198	gi|380032312|ref|YP_004889303.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	0.1077	0.0957	0.0777
37	1233	gi|380032347|ref|YP_004889338.1| molybdopterin biosynthesis protein MoeB [Lactobacillus plantarum WCFS1]	0.3694	0.1438	0.0068
38	1288	gi|380032402|ref|YP_004889393.1| aminodeoxychorismate lyase [Lactobacillus plantarum WCFS1]	0.0072	0.0005	0.0121
39	1295	gi|380032409|ref|YP_004889400.1| transpeptidase (penicillin binding protein 2B) [Lactobacillus plantarum WCFS1]	0.1965	0.0013	0.0246
40**	1302	gi|380032416|ref|YP_004889407.1| rhodanese family protein [Lactobacillus plantarum WCFS1]	0.0057	0.0009	0.0086
41	1305	gi|380032419|ref|YP_004889410.1| tRNA isopentenylpyrophosphate transferase [Lactobacillus plantarum WCFS1]	0.4654	0.0263	0.03
42	1338	gi|380032452|ref|YP_004889443.1| methionyl-tRNA formyltransferase [Lactobacillus plantarum WCFS1]	0.0791	0.0037	0.0034
43	1421	gi|380032534|ref|YP_004889525.1| dipeptidase [Lactobacillus plantarum WCFS1]	0.0584	0.004	0.0177
44	1449	gi|380032562|ref|YP_004889553.1| D-methionine ABC transporter, substrate binding protein [Lactobacillus plantarum WCFS1]	0.7961	0.0026	0.1663
45	1454	gi|380032567|ref|YP_004889558.1| transpeptidase-transglycosylase (penicillin binding protein 1A) [Lactobacillus plantarum WCFS1]	0.0303	0.0016	0.0087
46	1540	gi|380032656|ref|YP_004889647.1| carboxy-terminal proteinase, S41 family,peptidoglycan-bound [Lactobacillus plantarum WCFS1]	0.0699	0.0099	0.0037
47	1542	gi|380032658|ref|YP_004889649.1| extracellular protein [Lactobacillus plantarum WCFS1]	0.2322	0.0052	0.0239
48	1543	gi|380032659|ref|YP_004889650.1| extracellular lipase/acylhydrolase with GDSL-like motif [Lactobacillus plantarum WCFS1]	0.3746	0.0026	0.0289
49**	1558	gi|380032674|ref|YP_004889665.1| extracellular protein, with LysM peptidoglycan binding domain [Lactobacillus plantarum WCFS1]	0.0058	0.0057	0.0074
50	1594	gi|380032710|ref|YP_004889701.1| phosphohydrolase [Lactobacillus plantarum WCFS1]	0.1043	0.1492	0.0211
51	1637	gi|380032751|ref|YP_004889742.1| short-chain dehydrogenase [Lactobacillus plantarum WCFS1]	0.192	0.0609	0.0386
52	1644	gi|380032758|ref|YP_004889749.1| N-acetylmuramoyl-L-alanine amidase [Lactobacillus plantarum WCFS1]	0.406	0.006	0.0745

53	1670	gi|380032787|ref|YP_004889778.1| poly(glycerophosphate chain) D-alanine transfer protein DltD [Lactobacillus plantarum WCFS1]	0.5229	0.0144	0.0139
54	1718	gi|380032838|ref|YP_004889829.1| transcriptional attenuator, cell envelope-related, LytR family [Lactobacillus plantarum WCFS1]	0.0894	0.0011	0.0404
55	1743	gi|380032863|ref|YP_004889854.1| polysaccharide biosynthesis polyprenyl glycosylphosphotransferase, priming glycosyltransferase [Lactobacillus plantarum WCFS1]	0.0038	0.0002	0.0014
56	1767	gi|380032887|ref|YP_004889878.1| ComE operon protein 1, DNA receptor [Lactobacillus plantarum WCFS1]	0.3195	0.007	0.0864
57	1768	gi|380032888|ref|YP_004889879.1| extracellular protease Lon, peptidase S16 family, membrane-bound [Lactobacillus plantarum WCFS1]	0.2101	0.0043	0.0061
58**	1787	gi|380032907|ref|YP_004889898.1| membrane protein [Lactobacillus plantarum WCFS1]	0.0481	0.0289	0.0032
59	1821	gi|380032941|ref|YP_004889932.1| transpeptidase, penicillin binding protein 2B [Lactobacillus plantarum WCFS1]	0.0413	0.0009	0.0029
60	1822	gi|380032942|ref|YP_004889933.1| cell division protein FtsL [Lactobacillus plantarum WCFS1]	0.0125	0.0006	0.0127
61	1859	gi|380032979|ref|YP_004889970.1| bacterial type II secretion/trafficking system protein [Lactobacillus plantarum WCFS1]	0.0244	0.0032	0.0204
62	1860	gi|380032980|ref|YP_004889971.1| bacterial type II secretion/trafficking system extracellular protein [Lactobacillus plantarum WCFS1]	0.1882	0.0257	0.0464
63	1902	gi|380033023|ref|YP_004890014.1| metal dependent phosphohydrolase, HD family [Lactobacillus plantarum WCFS1]	0.0805	0.0011	0.033
64	1916	gi|380033037|ref|YP_004890028.1| cell shape determining protein MreC [Lactobacillus plantarum WCFS1]	0.6175	0.0093	0.0756
65	1925	gi|380033046|ref|YP_004890037.1| septation ring formation regulator [Lactobacillus plantarum WCFS1]	0.0578	0.0008	0.0165
66	1936	gi|380033057|ref|YP_004890048.1| beta-lactamase, class A [Lactobacillus plantarum WCFS1]	0.0376	0.0068	0.0252
67**	1958	gi|380033079|ref|YP_004890070.1| H(+)-transporting two-sector ATPase, B subunit [Lactobacillus plantarum WCFS1]	0.0108	0.0007	0.0032
68	2078	gi|380033199|ref|YP_004890190.1| amino acid regulated enzyme with ACT domain [Lactobacillus plantarum WCFS1]	0.0791	0.0028	0.0872
69	2135	gi|380033258|ref|YP_004890249.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	0.1675	0.0144	0.0618
70	2230	gi|380033353|ref|YP_004890344.1| diguanylate cyclase/phosphodiesterase, EAL domain [Lactobacillus plantarum WCFS1]	0.0036	0.0004	0.0036
71	2250	gi|380033373|ref|YP_004890364.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	0.3203	0.0064	0.1548

72	2313	gi|380033435|ref|YP_004890426.1| glycosyl hydrolase family protein [Lactobacillus plantarum WCFS1]	0.9881	0.0047	0.0031
73	2314	gi|380033436|ref|YP_004890427.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	0.6335	0.0586	0.0831
74	2343	gi|380033465|ref|YP_004890456.1| short-chain dehydrogenase [Lactobacillus plantarum WCFS1]	0.1895	0.0282	0.0329
75	2353	gi|380033475|ref|YP_004890466.1| signal peptidase I [Lactobacillus plantarum WCFS1]	0.0888	0.0007	0.0075
76	2354	gi|380033476|ref|YP_004890467.1| signal peptidase I [Lactobacillus plantarum WCFS1]	0.0534	0.0005	0.0059
77	2376	gi|380033498|ref|YP_004890489.1| aldo/keto reductase family protein [Lactobacillus plantarum WCFS1]	0.3205	0.244	0.281
78	2424	gi|380033548|ref|YP_004890539.1| lipase/esterase, subfamily of SGNH-hydrolases [Lactobacillus plantarum WCFS1]	0.6555	0.0118	0.1473
79	2426	gi|380033550|ref|YP_004890541.1| TetR family transcriptional regulator [Lactobacillus plantarum WCFS1]	0.3795	0.0147	0.0054
80	2466	gi|380033594|ref|YP_004890585.1| extracellular transglycosylase, with LysM peptidoglycan binding domain [Lactobacillus plantarum WCFS1]	0.9953	0.0029	0.0013
81	2471	gi|380033599|ref|YP_004890590.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	0.852	0.0036	0.0989
82	2477	gi|380033605|ref|YP_004890596.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	0.8094	0.008	0.0672
83	2478	gi|380033606|ref|YP_004890597.1| short-chain dehydrogenase [Lactobacillus plantarum WCFS1]	0.0526	0.0216	0.0072
84	2516	gi|380033646|ref|YP_004890637.1| extracellular protein [Lactobacillus plantarum WCFS1]	0.992	0.0034	0.004
85	2527	gi|380033657|ref|YP_004890648.1| TetR family transcriptional regulator [Lactobacillus plantarum WCFS1]	0.0513	0.0095	0.0061
86	2530	gi|380033660|ref|YP_004890651.1| lysozyme/muramidase, glycoside hydrolase family 25 [Lactobacillus plantarum WCFS1]	0.636	0.1369	0.0405
87	2533	gi|380033663|ref|YP_004890654.1| short-chain dehydrogenase [Lactobacillus plantarum WCFS1]	0.1337	0.0197	0.0074
88	2553	gi|380033683|ref|YP_004890674.1| cell surface protein, CscC family [Lactobacillus plantarum WCFS1]	0.9449	0.0283	0.0197
89	2567	gi|380033697|ref|YP_004890688.1| extracellular protein [Lactobacillus plantarum WCFS1]	0.7367	0.0006	0.0202
90	2570	gi|380033700|ref|YP_004890691.1| PTS system transporter subunit IIB [Lactobacillus plantarum WCFS1]	0.278	0.0036	0.0084

91	2589	gi|380033727|ref|YP_004890718.1| cell surface protein, membrane-anchored [Lactobacillus plantarum WCFS1]	0.5132	0.0025	0.2169
92	2615	gi|380033753|ref|YP_004890744.1| cell surface hydrolase [Lactobacillus plantarum WCFS1]	0.3315	0.0048	0.0709
93	2670	gi|380033808|ref|YP_004890799.1| extracellular protein, membrane-anchored [Lactobacillus plantarum WCFS1]	0.1107	0.0025	0.2345
94	2757	gi|380033901|ref|YP_004890892.1| cell surface hydrolase, membrane-bound [Lactobacillus plantarum WCFS1]	0.1689	0.0033	0.1542
95	2805	gi|380033950|ref|YP_004890941.1| cell surface protein, CscB family [Lactobacillus plantarum WCFS1]	0.789	0.0131	0.152
96	2806	gi|380033951|ref|YP_004890942.1| cell surface protein, CscB family [Lactobacillus plantarum WCFS1]	0.9859	0.001	0.0123
97	2845	gi|380033990|ref|YP_004890981.1| shikimate 5-dehydrogenase [Lactobacillus plantarum WCFS1]	0.5336	0.0254	0.0372
98	2956	gi|380034102|ref|YP_004891093.1| sorbitol operon activator [Lactobacillus plantarum WCFS1]	0.0344	0.0012	0.0523
99	2986	gi|380034132|ref|YP_004891123.1| sorbitol operon activator [Lactobacillus plantarum WCFS1]	0.0269	0.002	0.101
100	3002	gi|380034148|ref|YP_004891139.1| signal peptidase I [Lactobacillus plantarum WCFS1]	0.109	0.0006	0.0287

*: This protein is annotated as a hypothetical membrane protein in UniProt.

**: These proteins are predicted to have an N_{out} – C_{in} orientation by TMHMM.

```
seq <- read.table("locatep2_prediction N-terminale ankre.txt", fill
= TRUE, header = TRUE, sep = "\t")


seq$Identifier <- paste0(">",seq$Identifier)


fasta <- paste(seq$Identifier, seq$Sequence, sep = "\n")

fasta <- gsub("\"", "", fasta)


writeLines(fasta, "locatep2_prediction N-terminale ankre.fasta")
```

Figure A-1. Script from Rstudio that was used to convert the list of amino acid sequences in the excel document: “Locatep2_prediction N-terminale ankre” into FASTA format.

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway