How to Be a Bad Bug: Virulence Determinants of *Enterococcus faecium*

Theresa Wagner
A dissertation for the degree of Philosophiae Doctor – June 2018
A dissertation for the degree of Philosophiae Doctor

How to Be a Bad Bug:
Virulence Determinants of Enterococcus faecium

THERESA WAGNER

June 2018

Research Group of Host-Microbe Interactions
Department of Medical Biology
Faculty of Health Sciences
UiT – The Arctic University of Norway
Acknowledgments

This work has been carried out in the Research Group of Host-Microbe Interactions, Faculty of Medical Biology, UiT, the Arctic University of Norway. I express my gratitude to the Northern Norway Regional Health Authority Medical Research Program for the financial support of this PhD project.

I would like to express my sincere gratefulness to our research group and the research milieu of Tromsø and Utrecht for providing me an environment to discover, learn and develop. This thesis represents a group effort and I am deeply grateful to all who participated and helped in the process.

First and foremost, I would like to express my great appreciation to my supervisors Kristin Hegstad, Mona Johannessen and Torunn Pedersen. Thank you for welcoming me in Tromsø, for leading me along the way, always having an open door for me, giving your time so generously and supporting me throughout the final steps.

Most of all, my deep and heartfelt gratitude goes to my main supervisor, Kristin, for all the guidance and encouragement. Thanks for sharing your knowledge and enthusiasm, for giving me opportunities to explore ideas and helping me find my path. Kristin, you were a mentor one could only wish for.

I would like to give special thanks to my second supervisor, Mona. I appreciate all your valuable feedback and continuous help. Thank you for sharing your experience and expertise.

I wish to acknowledge my third supervisor, Torunn, for all your contributions, comments, and precise revisions.

I am very grateful to our collaborators in Utrecht, especially Rob Willems and Fernanda Paganelli, who made my stay abroad possible and enjoyable. Thank you for welcoming me to your lab, for teaching me all those methods, and thank you for many inspiring discussions.

A big thanks goes to my colleagues from our research group and associates for the good collaboration and providing such a nice work atmosphere. Thanks to the seniors and post-docs for constructive suggestions, to the technicians for continually keeping the lab running as well as excellent technical support, and to all my fellow PhD-students for making university so lively.

Thank you, Clement and Julia, for being dear friends through out my PhD journey, from the beginning to the end, thank you for being my best critics, lab sunshine, and companions to rely on.

Dear friends, I thank you for all your company, joy and fun. Thanks to those who made our house a home. Thanks to the faithful ones, who were with me from afar.

Thank you, Mayeul, for holding my hand, thank you for chasing summits with me.

A warm thanks goes to my family for all the support and care. Especially, my parents Antje and Andreas, you are the ones who encouraged my curiosity, who woke my scientific interest; you might well be my first and longest inspiration.

Last of all, my dear Enterococcus, thank you for twisting every experimental idea I had.

You will be remembered.
Contents

Introduction ......................................................................................................................... 1
  Characteristics of \textit{E. faecium} .................................................................................. 1
  Horizontal Gene Transfer and its Agents ..................................................................... 3
  \textit{E. faecium} as a Pathogen ....................................................................................... 8
  \textit{E. faecium} as a Pathogen....................................................................................... 8
  Bacterial TIR-domain Containing Proteins as Immune Evasion Factors .................. 16
  Extracellular Vesicles ................................................................................................. 20

Objective............................................................................................................................. 24

Choice of Methods ............................................................................................................. 25
  Virulence Factor Research .......................................................................................... 25
  Studying Membrane Vesicles ..................................................................................... 28
  Investigating the Genetic Content of Megaplasmids .................................................. 30

Summary of Main Results.................................................................................................. 33
  Paper I: \textit{Enterococcus faecium} TIR-domain Genes Are Part of a Gene Cluster Which Promotes Bacterial Survival in Blood ................................................................. 33
  Paper II: \textit{Enterococcus faecium} Produces Membrane Vesicles Containing Virulence Factors and Antimicrobial Resistance Related Proteins .............................................. 33
  Paper III: Megaplasmid-encoded Metabolic Factors Enhance Proliferation of \textit{Enterococcus faecium} in Human Blood ......................................................................................... 34

General Discussion .......................................................................................................... 35

Conclusion ......................................................................................................................... 41

References ......................................................................................................................... 42
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>Clonal Complex</td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular Matrix</td>
</tr>
<tr>
<td>EVs</td>
<td>Extracellular Vesicles</td>
</tr>
<tr>
<td>HGT</td>
<td>Horizontal Gene Transfer</td>
</tr>
<tr>
<td>IRF</td>
<td>Interferon Regulatory Factor</td>
</tr>
<tr>
<td>IS</td>
<td>Insertion Sequence</td>
</tr>
<tr>
<td>LTA</td>
<td>Lipoteichoic Acid</td>
</tr>
<tr>
<td>MAC</td>
<td>Membrane Attack Complex</td>
</tr>
<tr>
<td>MAL</td>
<td>MyD88 Adaptor Like Protein</td>
</tr>
<tr>
<td>MGE</td>
<td>Mobile Genetic Element</td>
</tr>
<tr>
<td>MLST</td>
<td>Multilocus Sequence Typing</td>
</tr>
<tr>
<td>MV</td>
<td>Membrane Vesicle</td>
</tr>
<tr>
<td>MyD88</td>
<td>Myeloid Differentiation Primary Response Gene 88</td>
</tr>
<tr>
<td>NFkB</td>
<td>Nuclear Factor κB</td>
</tr>
<tr>
<td>OMVs</td>
<td>Outer Membrane Vesicles</td>
</tr>
<tr>
<td>PAMPs</td>
<td>Pathogen Associated Molecular Patterns</td>
</tr>
<tr>
<td>PRRs</td>
<td>Pattern Recognition Receptors</td>
</tr>
<tr>
<td>Rep</td>
<td>Replicon</td>
</tr>
<tr>
<td>SARM</td>
<td>Sterile α and Armadillo Motif Containing Protein</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single Stranded DNA</td>
</tr>
<tr>
<td>ST</td>
<td>Sequence Type</td>
</tr>
<tr>
<td>T4CP</td>
<td>Type 4 Coupling Protein</td>
</tr>
<tr>
<td>T4SS</td>
<td>Type 4 Secretion System</td>
</tr>
<tr>
<td>TIR</td>
<td>Toll/Interleukin 1 Receptor</td>
</tr>
<tr>
<td>TirE</td>
<td>TIR-domain containing proteins of <em>E. faecium</em></td>
</tr>
<tr>
<td>TLRs</td>
<td>Toll-like Receptors</td>
</tr>
<tr>
<td>Tn</td>
<td>Transposon</td>
</tr>
<tr>
<td>TRAM</td>
<td>TRIF-related Adaptor Molecule</td>
</tr>
<tr>
<td>TRIF</td>
<td>TIR-domain Containing Adaptor Protein Inducing Interferon β</td>
</tr>
<tr>
<td>WGS</td>
<td>Whole Genome Sequencing</td>
</tr>
<tr>
<td>WTA</td>
<td>Wall Teichoic Acid</td>
</tr>
</tbody>
</table>
List of Papers
This thesis is based on the following manuscripts, cited by roman numerals

Paper I
Wagner T, Janice J, Paganelli FL, Willems RJ, Askarian F, Pedersen T, Top J, de Haas C, van Strijp JA, Johannessen M #, Hegstad K #.
*Enterococcus faecium TIR-domain genes are part of a gene cluster which promotes bacterial survival in blood*
(Submitted Manuscript)

Paper II
*Enterococcus faecium produces membrane vesicles containing virulence factors and antimicrobial resistance related proteins*
(Accepted; Journal of Proteomics)

Paper III
Wagner T, Janice J, Hegstad K, Pedersen T.
*Megaplasmid-encoded Metabolic Factors Enhance Proliferation of Enterococcus faecium in Human Blood*
(Manuscript)

*, # These authors contributed equally to this work.
Summary

Enterococcus faecium is a ubiquitous bacterium that only recently emerged as a nosocomial multi-resistant pathogen. Infections mostly affect immunocompromised patients and the high antimicrobial resistance of E. faecium often hampers treatment. To understand E. faecium’s pathogenicity research on virulence factors is needed. The focus of this thesis are features, which give pathogenic potential to E. faecium. First, TIR-domain-containing proteins (TirEs) are examined as immune evasion factors. Second, it is investigated if membrane vesicles (MVs) are produced by nosocomial E. faecium and whether they contain virulence-related cargo. Third, the content of a megaplasmid originating from a bacteremia isolate is characterized.

TirE proteins are described as novel virulence factors of E. faecium. The tirE encoding genes are exclusive to nosocomial E. faecium strains and localized on a putative mobile genetic element of phage origin. Comparison of the wild-type and its isogenic mutant which lacks the tirE locus showed that the tirE locus promotes bacterial proliferation in human blood. Both TirE proteins were detected in bacterial supernatant and one of them was associated with membrane vesicles derived from E. faecium (Paper I).

An isolation protocol using ultracentrifugation and purification of the vesicles over a density gradient was established. Using a proteomic approach the proteinaceous content of four different clinically relevant strains under two different growth conditions was described. In addition to virulence factors, the MVs were found to contain vaccine candidates and antimicrobial resistance related proteins. Thus, MVs are likely to be used by E. faecium to release virulence-promoting factors (Paper II).

A putative virulence factor-encoding megaplasmid was transferred to a commensal strain by filter-mating experiments. The megaplasmid enhanced the bacterial proliferation in a blood survival assay. DNA sequence analysis of the plasmid and its variants revealed candidate genes, which are potentially responsible for the observed effect. Most factors encoded by these candidate genes are predicted to be involved in primary metabolic processes and the genes are enriched among clinical strains. Megaplasmids are therefore likely to serve as vessels for various genes, including those needed for niche adaption in infection (Paper III).
Introduction

Characteristics of *E. faecium*

**General features of *E. faecium***

Enterococci are facultative anaerobic non-spore-forming Gram-positive bacteria of the phylum *Firmicutes*, and often occur in pairs as diplococci ¹, as illustrated in figure 1. Enterococci reside in sand, fresh and marine water sediment, soil and vegetation, but also colonizes the gastrointestinal tract of mammals, birds, reptiles and insects ². In humans enterococci make up less than 1% of the adult gut microbiota ³. The two most abundant enterococcal commensals of humans are *Enterococcus faecium* and *Enterococcus faecalis* ⁴. *E. faecalis* historically accounted for 80 to 90% of the clinical enterococcal isolates, whereas only 5 to 10% used to be *E. faecium* ⁵. Presently, however, *E. faecium* has risen to be the cause of more than 30% of enterococcal infections ⁶.

![Electron microscopy picture of *E. faecium* E155.](image)

*Figure 1: Electron microscopy picture of *E. faecium* E155. E. faecium typically occurs as diplococci.*

The clade structure of *E. faecium*

A deep phylogenic split of the hospital-associated clade (A) and the commensal clade (B) characterizes the clade structure of *E. faecium*. First, MLST (multilocus sequence typing), where isolates are compared based on their allelic profile of seven housekeeping genes, suggested that hospital linages belonged to a distinct genetic lineage called CC17 (Clonal Complex 17 with the
important lineages ST17, ST18, ST78) and clustered differently from community isolates. The relation of these STs (sequence types), based on the number of shared alleles, used to be assessed by eBURST. However, more recent analyses showed that the high recombination rate within \textit{E. faecium} makes this analysis unreliable and that eBURST networks wrongly link unrelated CCs. The current method of choice is therefore BAPS (Bayesian Analysis of Population Structure), which uses a statistical genetic model to subdivide molecular variation based on clonal ancestry and recombination patterns. Here, \textit{E. faecium} isolates cluster into numerous groups, but the majority of nosocomial isolates cluster into only two subgroups: 2-1, including the ST78 lineage, and 3-3, including the ST17 and ST18 lineages.

Now it is known, that the nosocomial clade A1 and the commensal clade B split 3,000 years ago, a time that coincides with increasing urbanization, domestication of livestock and the introduction of hygiene measures. Clade A and B differ in core genome level at 3.2 – 4\%\textsuperscript{12}. Approximately 75 years ago, a split occurred within clade A. The hypermutating clade A1, including the majority of clinical \textit{E. faecium} isolates, branched from clade A2, which mainly consists of animal isolates. Interestingly, this timeframe overlaps with the beginning of the modern antibiotic era. The clade structure of \textit{E. faecium} is depicted in figure 2. This figure also illustrates that clade A1 strains are predominantly isolated from infections.

\textbf{Figure 2: Clade structure of \textit{E. faecium}.} The origins of the strains and the dates for the split between the clades are indicated (ya, years ago). The infectivity score reflects the number of strains of a particular ST isolated from infection. Clades are color-coded: clade B in dark blue, clade A1 in red, and clade A2 in gray. The tree is based on alignments of DNA sequences of 1,344 single-copy core genes in 73 \textit{E. faecium} genomes. Reprinted with permission, modified from mBio\textsuperscript{11}.

Most recent phylogenetic reconstruction of whole genome sequenced \textit{E. faecium} strains, describing clonal relatedness of the strains based on pairwise single nucleotide polymorphisms comparisons,
Introduction

confirmed the relation of the nosocomial clade A1 and commensal clade B. But it found that the former A2 clade now appeared as a polyphyletic group in between clade A1 and B, and was therefore designated IG (intermediate group). IG is a multiclonal group comprising of clinical, animal, and community isolates and its topology varies significantly from the ladder-like topology of clade A1. However, as described earlier for A1 and A2, A1 and IG are closely related. 14 Especially nosocomial E. faecium strains have a highly malleable genome and the capacity to acquire and exchange large amounts of DNA. The hospital-associated clade A1 differs from the commensal clade by its larger core genome and also has a larger overall average genome size. Additionally, MGEs (mobile genetic elements), including plasmids, integrated phages and genomic islands, are enriched in clade A1 11. The high gene acquisition of clade A1 is likely to be promoted by its tendency to lack functional CRISPR (clustered regularly interspaced short palindromic repeats), a bacterial defense system giving immunity against incoming DNA, which would normally restrict the acquisition of novel genes 15–17. Further characteristics of clade A1 strains are typically the possession of insertion element IS16 18, pathogenicity islands and plasmids or genes associated with antimicrobial resistance as well as colonization and virulence 19.

Horizontal Gene Transfer and its Agents

Horizontal Gene Transfer Mechanisms

HGT (horizontal gene transfer) enables bacteria to disseminate genes among related and unrelated bacterial species, which is important for the adaption to new niches or challenges such as antimicrobial pressure 11. First studies on transferable antimicrobial resistance in enterococci date back to the 1970s 20. Enterococci have a highly plastic genome and especially nosocomial clade A1 strains harbor multiple MGEs (on average 5 plasmids/ genome in clade A1, compared to 2 plasmids/ genome in clade B strains; 2 phages/ genome in clade A1, compared to 1 phage/ genome in clade B strains; 36 kb of island-associated sequence/ genome, compared to 17 kb of island-associated sequence/ genome in clade B strains) 11. Importantly, enterococci have been shown to transfer resistance traits also to other more pathogenic species, such as S. aureus 21–23.

Genes, including those encoded by MGEs, are transferred horizontally in various ways and the three main mechanisms are transformation, transduction, and conjugation. Resistance genes or other advantageous genes can be transferred through all three mechanisms, as illustrated in Figure 3.
Figure 3: Main mechanisms of HGT between bacteria. A) Transformation occurs when naked DNA is released by an organism and taken up by another organism. A gene, encoding a resistance determinant or a virulence factor (red) can get integrated into a plasmid or the chromosome of the recipient. B) If genes are transferred by phages, the process is called transduction. C) Conjugation occurs by direct contact between two bacteria, which here is shown to exchange a plasmid. Transposons (yellow) can change location within the chromosome and plasmids. Reprinted with permission from Nat. Rev. Microbiol. 24.

Transformation is the uptake of naked DNA from the environment by a competent cell 24, but for enterococci natural transformation has never been reported 25. Transduction, which describes bacteriophage-mediated transfer of DNA from a donor to a recipient cell, is only beginning to be reported in enterococci. Bacteriophages may transfer bacterial genes, including resistance determinants, due to excision immediately adjacent to the integrated prophage, known as specialized transduction, or by packaging of any genomic DNA during their assembly, known as generalized transduction 26. Enterococcal phages have the capacity to mediate interspecies transfer of antimicrobial resistance genes 27 and even whole plasmids can be transferred through generalized transduction 28. Conjugation, the cell-cell contact dependent exchange of DNA between two live cells, is the primary way of gene mobility in enterococci. The metabolic burden of conjugation is decreased by...
Introduction

tight regulation of the process\textsuperscript{29}. Regulators register internal and external signals, such as environmental conditions, recipient availability and host cell physiology to initiate the process of conjugation. Subsequently, \textit{tra} genes, which by default are off, are expressed. The \textit{tra} genes are genes necessary for DNA transmission; they encode components of the T4SS (type 4 secretion system) as well as DNA transfer and replication\textsuperscript{29}.

The ssDNA transfer machinery of Gram-positive bacteria is seen as a minimized version of the T4SS found in Gram-negative bacteria. The differences arise from differences in cell wall architecture. In Gram-positive bacteria the T4SS spans only one membrane but needs a peptidoglycan hydrolase for local digestion of the cell wall. To substitute the lack of sex pili in the T4SS of Gram-positive bacteria, cell-cell contact formation involves adhesin-mediated condensation of extracellular DNA\textsuperscript{30}.

As illustrated in figure 4, mobilizable plasmids code only for a MOB (mobility) module consisting of an \textit{oriT} (origin of transfer), a relaxase, and need a T4SS of a co-resident conjugative system to become transmissible. Conjugative or self-transmissible plasmids encode this T4SS in addition to the MOB module and a T4CP (type 4 coupling protein). After the establishment of cell-cell contact through pili, a T4SS channel is formed and a rolling circle type replication or theta replication proceeds. The relaxase cleaves a specific site within \textit{oriT} to initiate conjugation and a DNA strand bound to the relaxase protein is displaced by ongoing DNA replication. The relaxase interacts with the T4CP and other components of the T4SS and gets transported to the recipient cell, with the ssDNA (single stranded DNA) bound to it. Subsequently, the ssDNA is pumped into the recipient by the ATPase activity of the T4CP and finally double stranded DNA is reconstituted in both cells\textsuperscript{31}.

![Figure 4: Scheme of transmissible plasmids and essential interactions during conjugation](image)

Conjugative plasmids code for the four components of a conjugative apparatus: an \textit{oriT} (violet), a relaxase (R, red), a type IV coupling protein (T4CP, green), and a type IV secretion system (T4SS, blue). Based on Smillie \textit{et al.} \textsuperscript{31}
Introduction

Besides these three main mechanisms of HGT, genetic material can also be transmitted through nanotubes and membrane vesicles (MV). Nanotubes have structures similar to cell membranes and can transfer intracellular molecules including metabolites, proteins, toxins, RNA, and plasmid DNA between neighboring and distant bacteria of the same and different genera\(^{32,33}\). Nanotubes have potential for a broad distribution of genes involved in antimicrobial resistance \(^{34}\). RNA and DNA containing MVs also mediate HGT in response to environmental stimuli \(^{35}\). MVs may spread antimicrobial resistance genes such as a carbapenem-resistance gene of *Acinetobacter baumannii* \(^{36}\) and a ciprofloxacin-resistance gene of mycoplasma \(^{37}\). Also MVs of Gram-positive bacteria contain DNA and transfer it between bacteria. *Clostridium perfringens* MVs contain toxin-encoding genes and MVs from *Ruminococcus* sp. transfer chromosomal DNA in heterogeneous bacterial communities \(^{38}\). Even though it has not been described yet, it stands to reason that nanotube or MV-mediated HGT mechanisms might exist in enterococci as well.

**Mobile Genetic Elements in E. faecium**

MGEs are agents of both intra- and intercellular DNA mobility. They can contain genes to replicate independently of the chromosome, recombine homologously or non-homologously, as well as to move DNA between cells. The MGEs accessory genes typically provide selective advantages to their host, for example genes encoding virulence factors or antimicrobial resistance determinants.

**Plasmids**

Plasmids are semi-autonomously replicating extra-chromosomal genetic elements of modular structure. Enterococcal plasmids range in size from 3 kb to above 300 kb and are present in low or high copy number, up to more than 20 copies per cell \(^{25,39}\). The backbone of a plasmid includes the genetic modules encoding self-replication and copy number control, and may also encode genetic modules involved in stable inheritance and the ability to transfer in between bacteria. Between these backbone modules, accessory plasmid modules are integrated. The accessory content gives rise to high genetic variability.

Differences in plasmid replication genes are used to type and classify plasmids \(^{40}\). Out of 20 replicon classes defined from Gram-positives, eleven classes have been identified in enterococci. Those eleven classes can be divided into four families: RepA\_N family plasmids encode replication initiators with the RepA\_N domain in their N-terminus, Inc18 family plasmids have initiator proteins related to the traditional Inc18 group, RCR family plasmids are a large group of plasmids replicating through a rolling circle mechanism, and Rep\_3 family plasmids contain the Rep\_3
Introduction

conserved domain in plasmid replication initiator proteins. In E. faecium rep17/pRUM and rep20/pLG1-type replicons (RepA_N), rep2_pRE25 (Inc18 family), rep14/EFNP1 (RCR family) and rep18/pEF418 (Rep_3 family) dominate.

An alternative classification scheme for plasmids is MOB typing, utilizing differences in relaxase encoding genes. Seven relaxase families have been described, out of which MOB_Q, MOB_C, MOB_T (found in conjugative elements), MOB_V (found in mobilizable plasmids) and MOB_P (found in conjugative elements and mobilizable plasmids) are present in enterococci.

Another older system of plasmid classification groups plasmids which fail to co-reside into Inc (incompatibility) groups, such as the enterococcal-streptococcal Inc18 group plasmids. Inc18 group plasmids, i.e. rep1/pIP501 and rep2/pER25 plasmids, often carry multiple antimicrobial resistance genes, and have a broad host range among streptococci, lactococci, staphylococci, and enterococci. The host range is also used to classify plasmids, examples of narrow host range plasmids are plasmids of the Rep_3 family, such as rep18/pEF418, rep11/pEF1071 and repCIZ2 plasmids, and plasmids of the RepA_N family. Pheromone-responsive plasmids are conjugative narrow host range plasmids. As the name suggests, their conjugative process is induced upon recognition of a pheromone-responsive plasmid-free recipient, such as E. faecalis pAD1, pCF10 and pAM373.

Plasmids conferring resistance to antimicrobial agents are called R-plasmids. In enterococci R-plasmids largely contribute to glycopeptide, quinupristin-dalfopristin, and high-level aminoglycoside resistance. To ensure stable inheritance in the absence of antimicrobial selection, R-plasmids often harbor plasmid addiction systems, namely toxin-antitoxin systems. R-plasmids encode labile anti-toxins and stabile toxins, which kill daughter cells that did not inherit the plasmid encoding these features. rep20/pLG1 plasmids are associated with the axe-txe locus, a plasmid stabilizing toxin-antitoxin system, and carry genes involved in antimicrobial resistance.

Especially in nosocomial E. faecium large conjugative plasmids (above 150 kb) are common, and besides antimicrobial resistance related genes may carry genes involved in colonization of the gastrointestinal tract among others. They are referred to as megaplasmids and have been reported to belong to the rep20/pLG1 type. Plasmids can furthermore help the transfer of other MGEs and with that serve as scribbling pads in formation and propagation of new operons.

Transposable Elements

Transposable elements are movable units of DNA. The simplest transposable elements are IS (insertion sequence) elements, which encode only enzymes for their own transposition. Typical transposable elements with additional genes found in enterococci are composite transposons, Tn3 family transposons, and conjugative transposons. Composite transposons are flanked by two copies
of IS elements from the same family, which move the DNA in between them. Enterococcal high-level gentamicin resistance is associated with composite transposons (Tn4001 (Tn5281) and variants 52–54, Tn5384 55, Tn5385 56, Tn924 57). High-level glycopeptide resistance encoded by vanA in enterococci is associated with a Tn3 family transposon, namely Tn1546 58. Tn3 family transposons contain a transposase (tnpA) and a resolvase (tnpR) and move in a replicative fashion. Conjugative transposons encode all information for their own excision, conjugation, and integration, and were therefore later classified as integrative conjugative elements (ICE). Conjugative transposons have for example been associated with resistance to tetracycline (Tn916-like transposons encoding tet(M) 59) or vancomycin (Tn1549-like transposons encoding vanB2 60). Importantly, conjugative transposons may co-transfer other transposons, plasmids, and even large chromosomal fragments between strains and they may have a broad host range, like Tn916 61–63. Transposable elements, including IS elements, facilitate niche adaptation, contribute to genome plasticity and phenotypic variation, as they mediate rearrangements in chromosomal and plasmid DNA but also alter expression of genes by disrupting promoters or providing alternative promoters or regulators 18,47. Even though newly acquired enterococcal MGEs can pose a severe fitness cost at first, this costs can rapidly be reduced and beneficial plasmid-host association may emerge 64. Thus, advantages gained through MGEs may outweigh their cost. Even though genetic exchange occurs with or without selective pressures, such as antimicrobial treatment, selects for rare events and makes HGT more probable 24.

**E. faecium as a Pathogen**

**Enterococcal Infection and Disease**

In 1899, the term enterococci was introduced to describe a saprophytic intestinal coccus able to cause infection 65,66 and the first enterococcal infection, infective endocarditis, was described 67. *E. faecium* infections are typically established upon translocation to the bloodstream, through a perturbed intestine or uptake from contaminated hospital environment 68. In the past, the source of infection was mainly assumed to be the patient’s own endogenous flora. With the increase of nosocomial infections, however, transmission also occurs among patients in a hospital setting 4. Today, *E. faecium* mainly causes urinary tract infections, but also life-threatening infections such as bacteremia. Importantly, among the five main bacteremia causing pathogens from 2002 to 2008, the increase in *E. faecium* bacteremia was the highest 69. The source of bacteremia with endocarditis as
its major complication is usually the urogenital system, but may also arise from *E. faecium* intra-abdominal, pelvic or soft-tissue infections.\(^4\)

Enterococcal infections are most serious in immunocompromised patients, where risk factors include diabetes mellitus, transplantation, malignancy, indwelling catheters and chemotherapy\(^70\). Another important factor in enterococcal disease is exposure of hospitalized patients to antimicrobials against Gram-negative bacteria, which results in substantial changes in the regulation of the gut immune system, leading to overgrowth with enterococci\(^1\). Normally, the lipopolysaccharide and flagellin of Gram-negative bacteria stimulates the production of REGIII\(\gamma\), a C-type lectin. REGIII\(\gamma\) has antimicrobial activity against Gram-positives, including *E. faecium*. Consequently, a decrease in Gram-negatives through antimicrobial exposure will facilitate a down-regulation of REGIII\(\gamma\) and thus lead to overgrowth of antimicrobial resistant enterococci\(^71,72\).

**E. faecium**’s Antimicrobial Resistance and Options for Infection Treatment

*E. faecium* has been recognized as one of the ESKAPE pathogens, along with *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* species. ESKAPE pathogens cause most nosocomial multi-resistant infections and require special attention\(^73\). The first option of treatment for enterococci is mainly ampicillin and penicillin\(^4\), but high intrinsic and acquired antimicrobial resistance especially in *E. faecium* often hamper treatment and decrease therapy options.

Compounds which inhibit cell-wall synthesis are often only bacteriostatic against enterococci\(^74\) and *E. faecium* is intrinsically non-susceptible towards cephalosporins, clindamycin, penicillins, and aminoglycosides\(^1\). Additionally, the rate of acquired resistance towards modern antimicrobials in *E. faecium* is high. The main mechanisms of enterococcal antimicrobial resistance are summarized in figure 5.

Approximately 90% of modern nosocomial *E. faecium* isolates are resistant towards β-lactams. PBP5, a penicillin binding protein, is involved in this resistance\(^1\). Besides the low-level intrinsic resistance to aminoglycosides, acquired high-level resistance occurs. High-level streptomycin resistance is conferred by ribosomal mutations or the acquisition of an aminoglycoside nucleotidytransferase, Ant(3’’)-la or Ant(6’’)-la. The bifunctional aminoglycoside-modifying enzyme Aac(6’)-le-Aph(2’’)-la confers high-level resistance to all other aminoglycosides (gentamicin MICs above 128 mg/L)\(^1\).
Infections caused by β-lactam or aminoglycoside resistant enterococci were usually treated with the glycopeptide vancomycin. In 1986, however, the first vancomycin resistant enterococci (VRE) were described. Glycopeptides are cell-wall-active agents, which bind to the D-Ala-D-Ala termini of peptidoglycan precursors to inhibit the synthesis of the bacterial cell wall. Glycopeptide resistance is achieved, first, through replacing the terminal D-Ala of peptidoglycan precursors with D-Ser, resulting in low-level resistance (7-fold vancomycin affinity decrease), or with D-Lac, resulting in high-level resistance (1000-fold vancomycin affinity decrease); and second, by removing precursors ending in D-Ala-D-Ala. Gene clusters, van-clusters, comprising both regulatory and resistance genes, encode the determinants conferring vancomycin resistance. To date eight genotypic variants of acquired vancomycin resistance (vanA, vanB, vanD, vanE, vanG, vanL, vanM, vanN) and one type of intrinsic resistance (vanC) have been described in enterococci. The vanA, vanB, vanD and vanM clusters encode change of the D-Ala-precursor to D-Lac and may cause high-level vancomycin resistance, whereas the other van-gene clusters encode change of the D-Ala-precursor to D-Ser and cause low-level vancomycin resistance. Glycopeptide resistance dissemination may...
involve the transposition of transposons between different replicons, including plasmids which promote horizontal transfer, i.e. in case of vanA \(^{79}\). The clinically important van-clusters are vanA and vanB. Globally and within Europe, vanA is the most prevalent \(^{80}\), while vanB was the most abundant cluster found in Australia previously \(^{81}\) but also recently increased within Europe \(^{82,83}\). Both clusters are associated with successful MGEs, which may be the reason for their relatively high abundance compared to the other van-clusters. vanA is normally part of the Tn3-family transposon Tn1546 often found in on conjugative plasmids \(^{84}\) and vanB is mostly part of the integrative conjugative element Tn1549 \(^{60}\), which localizes both on the chromosome and on plasmids \(^{84,85}\). Importantly, VRE bacteremia has a 2.5-fold increase in mortality compared to vancomycin-susceptible Enterococcus bacteremia \(^{86}\).

In order to treat VRE, quinupristin-dalfopristin, linezolid, and daptomycin are used. However, several mechanisms reducing the activity of quinupristin-dalfopristin exist in E. faecium \(^{87,88}\). Even though linezolid resistance still has an overall low prevalence of <1% \(^{89}\), it is alarming that multiple mechanisms of linezolid resistance exist which do not alter the fitness of E. faecium \(^{90}\). Linezolid resistance can be acquired through mutations or gene acquisition, resulting in changes of the ribosomal target or efflux, but also cell wall thickness and biofilm formation are under discussion as resistance mechanisms (under submission \(^{91}\)). Daptomycin resistance is associated with structural changes of the cell envelope and involves mutations in multiple intrinsic genes (under submission \(^{91}\)). With this, resistance even to the last resort antimicrobial agents is found in enterococci.

As discussed above, enterococcal infections are therapeutically challenging, due to the intrinsic resistance towards many antimicrobial agents, and the ability to acquire resistance towards essentially every antimicrobial agent used against enterococci. Therefore, combination therapy of synergistic agents is often the only treatment option \(^{92}\). This approach has its limitations as well, since multi-resistances may rule out combination possibilities and toxic side effects are likely to occur \(^{1}\). Thus, novel approaches to treat enterococcal infections are needed.

**Persistence and Biofilm Formation**

Another factor, which makes E. faecium a resilient bacterium, is its ability to withstand harsh conditions. Optimal growth of enterococci is typically at 35°C, but their growth range is as wide as from 10 to 45°C \(^{93}\). Enterococci can also endure both hypotonic and hypertonic conditions \(^{94}\). Normally, enterococci grow in 6.5% NaCl but tolerate higher salt concentrations \(^{95}\). They further withstand chlorine and alcohol preparations frequently used as disinfectants \(^{96,97}\). Therefore, enterococci can survive over long time spans in virtually any kind of environment, including
hospital surfaces. For example, three CC17 isolates survived for more than 5 years in an in vitro setting without nutrition \(^98\); and the same multi-resistant high-risk clone caused infections in a Norwegian hospital one year apart from each other, and must have survived within the hospital during that time \(^90\).

Biofilms enhance the resilience of enterococci further, both ex vivo, such as on medical devices and in the hospital environment, as well as in vivo, where biofilms are protecting their inhabitants \(^99\). Biofilm formation is by definition the adherence of planktonic bacteria to each other and a surface, and the subsequent growth resulting in a structured community, often enclosed in an extracellular polymeric substance matrix. Biofilm production is regulated by environmental signals such as nutrients, pH, temperature, and presence of serum \(^99\). A number of enterococcal genetic determinants playing a role in biofilm formation and maturation have been identified, mostly in \(E.\ faecalis\). Among those are surface adhesins, autolysins and glycolipides, which mainly act early in the adhesion phase, as well as polysaccharides, extracellular DNA, and proteases, which predominate in biofilm maturation \(^100\). Once a biofilm is established, bacteria within biofilms can resist phagocytosis, which makes it extremely difficult for the host to clear the bacteria \(^99\). Additionally, biofilms protect the bacteria against antimicrobials: 10-1000 times higher antibiotic concentrations can be tolerated by a biofilm compared to the amount required to kill planktonic bacteria \(^101\). Biofilms are also hot spots for HGT, including exchange of antimicrobial resistance genes \(^102\).

Physical removal of infected medical devices or the infected tissue is usually the only option to attack biofilms. So far, the development of anti-biofilm compounds mainly addressed the prevention of biofilm formation on medical devices, but to date, there are no drugs available to disassemble a biofilm \(^100\).

**Virulence Factors of \(E.\ faecium\)**

Generally, \(E.\ faecium\) displays low levels of virulence and pathogenicity. However, due to its emergence as a nosocomial pathogen, further knowledge on virulence factors is warranted. Figure 6 summarizes the current knowledge on \(E.\ faecium\)'s virulence factors.

Except for virulence factors per se, cell wall components, such as capsular polysaccharides, LTA (lipoteichoic acid) and WTA (wall teichoic acid) play a role in virulence of Gram-positive bacteria. Together they contribute to phagocytosis-, complement- and antimicrobial-resistance, biofilm formation and host or surface attachment, as well as being important immunogenic components \(^103\). Colonization and binding to ECM (extracellular matrix) proteins is crucial for the establishment of enterococcal infection, thus virulence factors of \(E.\ faecium\) are mainly liked to biofilm formation.
and adhesion. *E. faecium* virulence factors encompass both secreted components and cell surface associated factors (reviewed in Gao et al. 103, figure 6).

Secreted virulence factors comprise cytolysin Cyl, secreted antigen SagA, and gelatinase GelE. Cyl, as a haemolysin, damages various host cells, such as red blood cells, neutrophils and macrophages 104. SagA, notably the most abundant protein in enterococcal biofilms, contributes to colonization by binding to ECM proteins 14,105,106. GelE affects adherence and biofilm formation 107.

Cell surface-associated virulence factors contribute to phagocytosis–resistance, biofilm formation and adhesion to both the host and surfaces. Especially important for adhesion are ECM binding virulence factors, such as the collagen adhesins Acm 108,109, Scm 110 and CcpA 111,112, as well as the fibronectin/ fibrinogen-binding proteins Fnm 113 and PrpA 114. Cell surface exposure and collagen binding of Acm is influenced by the major autolysin AtLA 115. Further cell wall-associated adhesins are SgrA 116, EcbA 116 and CapD 117, where the latter is a capsular polysaccharide protein which is involved in the immune evasion. The pilus proteins PilA2 118, PilB/ Ebpfm (Pilus) 119 are associated with clinical strains and promote biofilm formation. The most widespread virulence factor in clinical *E. faecium* strains is the enterococcal surface protein Esp 103, which besides cell adhesion and biofilm formation 120 also contributes to the attachment to silicon-based surfaces 121.

Additionally, mannose family phosphotransferase systems (PTSclin) are enriched in the clinical lineage and have been associated with *E. faecium* virulence 122,123. The transmembrane proteins of the PTS phosphorylate carbohydrates, that are translocated across the bacterial cell membrane 122. PTS is therefore proposed to be involved in the adaption of the carbohydrate metabolism to intestinal colonization 122. As part of another putative PTS the permease BepA was described to be involved in endocarditis and biofilm formation in human serum 123.

Besides interacting with the host through virulence factors, enterococci have to compete with other bacteria within the gastrointestinal tract. Therefore, enterococci may utilize bacteriocins, called more specifically enterocins, which are peptidic toxins inhibiting the growth of other bacteria, including pathogenic microorganisms. For example enterocin P (EntP) inhibits the growth of the pathogens *Listeria monocytogenes*, *S. aureus*, *C. perfringens*, and *C. botulinum* 124. Enterocins are mainly plasmid encoded and further include EntA, EntB, EntL50 125, EntQ 126, Bac32 84 and Bac43 127. Because of the increasing antimicrobial resistance, bacteriocins gained attention since they might be used as novel drugs (reviewed by Nes et al. 128).
Introduction

Figure 6: Overview of *E. faecium* virulence factors. Virulence factors utilized by *E. faecium* are illustrated and described in short; including cell wall components, secreted virulence factors and membrane-bound virulence factors. Based on Gao *et al.* [103].
**Immune Response Towards *E. faecium***

A mechanical barrier that is virtually impermeable to pathogens is built by epithelial surfaces, consisting of epithelial cells joined by tight junctions and the removal of pathogens from surfaces is supported by desquamation of skin epithelium or movement due to peristalsis or cilia. In addition, chemical factors prevent bacterial growth, such as the low pH of gastric secretions or antimicrobial peptides \(^{129}\). Pathogens, which overcome physical and chemical barriers, encounter cells of the innate immune system, the host’s primary defense against microbial attacks, which mediates bacterial clearance and bridges to adaptive immunity.

Innate immunity is induced upon recognition of highly conserved microbial structures, called PAMPs (pathogen associated molecular patterns), by PRRs (pattern recognition receptors). Among PRRs TLRs (Toll-like receptors) are dominant \(^{130}\).

Out of the 10 human TLRs, TLR2 has a central role in the recognition of Gram-positive bacteria, as it recognizes peptidoglycan, lipoteichoic acid and lipoproteins. *E. faecium* is recognized by TLR2 \(^{131}\), a receptor expressed by various cells, including neutrophils, macrophages, and dendritic cells. CD14 functions as a co-receptor of TLR2 to recognize lipoteichoic acid and has also been shown to be involved in the TLR2-recognition of *E. faecium*. The dimerization partner determines the ligand specificity of TLR2 \(^{131}\). The formation of TLR1-TLR2 is induced by Gram-negative triacyl-lipopeptides; while TLR2-TLR6 formation is induced by Gram-positive diacyl-lipopeptides \(^{132}\). In a mice peritonitis model, early host defense against *E. faecium* was driven by TLR2-dependent signaling via MyD88 (myeloid differentiation primary-response gene 88) resulting in NFκB (nuclear factor-κB) activation and expression of proinflammatory mediators \(^{131,133}\).

Subsequent, peritoneal macrophages \(^{134}\), neutrophil influx \(^{135}\) and opsonization of *E. faecium* by the complement \(^{136}\) are important in the early immune response of the host.

The complement system consists of over 30 serum and cell-surface proteins. Its activation occurs through three different pathways, all of which lead to the cleavage of the complement component C3. In the classical and lectin pathway, induced by antigen-antibody interaction or recognition of carbohydrate structures, C3 is cleaved by the C4b2a. In the alternative pathway, which amplifies the two other pathways, C3bBb cleaves C3. This cleavage results in the release of the soluble anaphylatoxin C3a, attracting phagocytes, and the deposition of C3b and iC3b on the surface of the bacterium. C3b and iC3b, in turn, are ligands for cellular receptors on leukocytes, phagocytosing the labeled bacterium. Cleavage of C3 also results in the formation of MAC (membrane attack complex, C5–C9), which directly lyases Gram-negative bacteria. \(^{137–139}\)
The thick cell wall of Gram-positive bacteria, including enterococci, is generally thought to prevent complement-mediated MAC lysis. However, commensal *E. faecium* strains have recently been shown to be susceptible to a serum phospholipase whereas clinical strains were resistant\(^\text{140}\). In clinical strains, genes involved in carbohydrate metabolism and nucleotide biosynthesis are essential for growth in human serum\(^\text{141}\).

Both complement and antibodies play an especially important role in phagocyte recruitment, and thus clearance of enterococci\(^\text{142–144}\). Opsonization by the complement also enhances phagocytosis by neutrophils and macrophages, the cells first encountered by the bacteria during infection\(^\text{136}\). Upon endocytosis of a pathogen and subsequent phagosome-formation, the phagosome fuses with a lysosome to form a phagolysosome, in which the pathogen is killed\(^\text{145}\). However, certain enterococcal strains are able to survive even within phagocytes\(^\text{146–148}\) and might thereby spread systemically\(^\text{149}\).

### Bacterial TIR-domain Containing Proteins as Immune Evasion Factors

#### Recognition of Bacteria by the Human Innate Immunity and TIR-domain Signaling

As mentioned above, the host innate immune system provides the primary defense against microbial attacks and is induced, when PRRs, among them TLRs, recognize PAMPs. PAMP-TRLR interaction consequently leads to the recruitment of adaptor molecules, whose stimulation results in pro-inflammatory mediator expression\(^\text{150,151}\).

The cytosolic side of TLRs and their adaptors all contain TIR (Toll/interleukin-1 receptor)-domains, through which homotypic interactions are mediated\(^\text{152}\), illustrated in figure 7.

The five TLR adaptor molecules are MyD88, MAL (MyD88-adaptor-like protein), TRIF (TIR-domain containing adaptor protein inducing interferon-β), TRAM (TRIF-related adaptor molecule) and SARM (sterile α- and armadillo-motif-containing protein). The universal adaptor MyD88 interacts with all TLRs, except for TLR3, while MAL and TRAM act as bridging adaptors. MAL recruits MyD88, and TRAM recruits TRIF. SARM, by contrast, is not recruited to TIR-domains to initiate signaling like the other four adaptors but inhibits TRIF-dependent signaling\(^\text{152}\). The TLR-adaptors couple to protein kinase pathways, leading to activation of transcription factors such as NFκB, as well as members of the IRF (interferon-regulatory factor) family, which in turn induce inflammatory cytokines and co-stimulatory molecules. This results in an host-mediated immune response against the invading pathogen\(^\text{153}\).
**Introduction**

![Diagram showing TLR-complexes and their adaptors](image)

**Figure 7: TIR-domains are mediating the interaction between TLRs and their adaptors.** TLR complexes recognize various ligands (green colors) and utilize each adaptor (blue colors) differently, with exception of SARM, which is negatively interfering with TRIF and TRAM. All TIR-domains are colored in grey. Based on O’Neill et al. 152.

**TIR-domains as Conserved Modules in Different Organisms**

Originally defined as the cytoplasmic domain of TLRs 154, the TIR-domain is a protein-protein interaction domain of approximately 125-200 amino acids. It is widely distributed among different species, including animal, plants, and bacteria. TIR-domains function through self-association or interact homotypic with other TIR-domains to build up protein scaffolds for the formation of protein complexes, especially immune signaling complexes.

The TIR-domain typically adopts a flavodoxin-like fold built up by a central five-stranded parallel β-sheet (βA- βE) surrounded by five α-helices (αA- αE) and connected by flexible loops, which are named based on the secondary structures they connect 155, as illustrated in figure 8. Despite the low sequence identity among TIR-domains, the core fold of the TIR-domain is conserved 156. The loops and helical regions, create differences in the structure of the TIR-domain among the TIR-containing proteins. This variability is assumed to be responsible for the specificity required for protein-protein interaction and signaling. Especially the BB loop (connecting strand βB and helix αB) and DD loop (connecting strand βD and helix αD) of the TIR-domains are important for TIR-TIR interaction 155,157.

Within the TIR-domains the sequence motifs, named box 1, box 2 and box 3, are found 156. Box 1 corresponds to the βA strand, and maintains structural stability of the TIR-domain, for example in TLR2 158. Box 2 corresponds to the βB strand and the BB loop, and box 3 corresponds to the αE helix 156. However, structural and biochemical studies of multiple TIR-domain proteins of different
species found that boxes 2 and 3 are rather poorly conserved and that other regions are also of functional importance\textsuperscript{156,159,160}.

**Figure 8: Illustration of TIR domain structures.** On the left, a cartoon representation of the MyD88 TIR-domain crystal structure is shown, labeled with key secondary structure elements. The five parallel β-strands are colored orange, the rest cyan. On the right, the MyD88 structure (cyan) is superposed to the structures of the TIR-domains from TLR2 (lime), TLR1 (pale green), and *Paracoccus* sp. PdTLP (magenta). Reprinted with permission from Proc. Natl. Acad. Sci. U.S.A. \textsuperscript{161}.

**Interference with TLR Signaling by Bacterial TIR-domain Containing Proteins**

TLR signaling provides an interesting point of vantage to pathogens, because of the TLR’s frontline role in innate immunity initiation. Counteraction of this initial recognition may be conducted by bacteria through molecular mimicry of key structures, such as TIR-domains\textsuperscript{162}. TIR-domains are widespread among bacteria, suggesting that they function as general protein-protein interaction domains, which may adopt diverse uses\textsuperscript{163}. The TIR-domain is located in the N- or C-terminal region of the protein and the remaining protein domains can vary\textsuperscript{156}. The first study identifying bacterial TIR-homologues found more than 200 TIR-domain proteins in a wide range of both Gram-negative and Gram-positive species, including the pathogens *Salmonella enterica* and *S. aureus*\textsuperscript{164}.

Most studies on bacterial TIR-proteins explored their role as putative virulence factors interfering with host signaling, especially NFκB activation. The putative mode of action of bacterial TIR-domain containing proteins, illustrated in figure 9, is competition with adaptor molecules\textsuperscript{164–167} or degradation of adaptor molecules\textsuperscript{168}. However, the specific targets of most bacterial TIR-domain containing proteins have not been identified.

The first functionally characterized bacterial TIR-domain containing proteins, *S. enterica* serovar Enteritidis TlpA, was shown to impair TLR- and MyD88-dependent activation of NFκB when TlpA was expressed in the eukaryotic cytoplasm. Additionally, a *tlpA* deletion mutant was attenuated in both a cell culture and a mouse infection model\textsuperscript{164}.
Figure 9: Bacterial TIR-domain containing proteins negatively interfere with innate immune recognition. Left) Bacteria are recognized by TLRs (light red), which subsequently interact with adaptor molecules (blue) via TIR-TIR interactions. Signaling cascades are induced leading to the activation of transcription factors (TF, such as NFκB), which results in phagocytosis of the bacteria and consequently the induction of an inflammatory response and microbial killing. Right) Bacteria with the gene encoding TIR-domain containing proteins (dark red), express and release those proteins. The bacterial TIR-domain containing proteins interfere with TLR-adaptor signaling and thus the bacterium evades the innate immune response. All TIR-domains are illustrated in grey.

In line with these observations, *Brucella* TcpB and *Escherichia coli* TcpC were shown to interact with MyD88 and studies with their respective deletion mutants found them to play a role in pathogenesis\(^\text{165}\). Similar observations were also made for the Gram-negative *Yersinia pestis* YpTdp\(^\text{169}\) and *Paracoccus denitrificans* PdTlp\(^\text{166}\) as well as for the Gram-positive *S. aureus* TirS\(^\text{167}\) and *E. faecalis* TcpF\(^\text{170,171}\).

How bacterial TIR-domain containing proteins are released has not been finally resolved, but stands to reason, that these proteins are secreted, even though most of them lack a signal-peptide. Studies in *E. coli* TcpC\(^\text{161}\) and *S. aureus* TirS\(^\text{167}\) showed, that no bacterium-cell contact is required for the virulence factor to exert its function.

Genes encoding TIR-domain containing proteins often localize in regions of phage origin, examples of this are tlpA\(^\text{164}\), tcpC and tcpB\(^\text{165}\). It is therefore likely that tir-genes are disseminated through lateral gene transfer. Additionally, the staphylococcal tirS-gene localizes within SCC (staphylococcal chromosomal cassette), a mobile genetic element containing a fusC gene conferring fusidic acid resistance\(^\text{172}\). The fact that sub-inhibitory concentration of fusidic acid increase TirS expression links regulation of virulence and antimicrobial resistance.
Besides their role as virulence factors in pathogens, TIR-domain proteins were also identified in non-pathogenic bacteria as well as fungi and archaea, and showed functions deviating from the common theme. For example, the two staphylococcal TIR-proteins, SaTlp1 and SaTlp2, found in zoonotic ST398, up-regulated NFκB signaling instead of down-regulating it. Therefore it has been suggested that TIR-domains are general protein-protein interaction domains, which are not primarily virulence factors but were secondarily adapted to host mimicry and immune evasion.

**Extracellular Vesicles**

**Extracellular Vesicles Exist in All Domains of Life**

EVs (extracellular vesicles) are spheres of lipid-bilayer enclosing liquids, formed from membranes either in a physiological process or upon mechanical disruption of membranes. Whether EVs are artifacts of lipid-self-assembly and released during cell death or whether vesiculation is an actual physiological process was subject of continuing debates. Nevertheless, multiple studies report that an active metabolism is required for the EV-production and that killed cells do not produce EVs. This underlines that the observed EVs are due to a physiological process. EVs are found in all domains of life investigated till this day, including eukaryotes, archaea, and bacteria, suggesting a universal phenomenon. As a common theme, vesicles can bud directly from the membrane of organisms without cell wall but have to escape the cell wall barrier in cell-walled organisms. EV-release provides an additional mode of secretion, which gives the flexibility to respond to environmental cues, especially in inter-microbial and host-microbe interaction.

**Archaeal Vesicles**

The evolutionary most basic process of vesicle production is probably represented by archael membrane vesicles, which combine features found in both eukaryotic and bacterial vesicle production. Archeal membrane vesicles are surface derived and released through budding of the cell surface. They play a role in inter-microbial communication and genetic exchange. For example, the 90-230 nm sized *Sulfolobus* derived vesicles contain toxins, sulfolobicins, which inhibit the growth of competing species, and *Thermococcales* derived membrane vesicles contain DNA and hence might mediate gene transfer.

**Eukaryotic Vesicles**

Eukaryotic cells, including animal, fungal and plant cells, release vesicles of at least three types with different modes of biogenesis. The first type, apoptotic bodies are products of membrane
blebbing during programmed cell death. They are the most heterogeneous and biggest extracellular vesicles with a size of 1000-5000 nm. The second type, shedding microvesicles are 100-1000 nm in size and directly bud from the cell surface. The third type, exosomes with a size of 40-100 nm are derived from multivesicular bodies within the cell and then released\textsuperscript{178}.

Originally it was thought, that animal cell-derived exosomes are part of a cellular waste disposal system. However, recently it was found that they mediate intercellular communication, modulate immune responses and shuttle nucleic acids\textsuperscript{179}. Since exosomes are released by all types of human cells and are found in all body fluids, a variety of therapeutic and diagnostic purposes has been suggested for exosomes, such as cancer diagnostics, targeted drug delivery and immunotherapy\textsuperscript{179}.

The eukaryotic parasite \textit{Leishmania donovani} releases extracellular vesicles containing up to 98\% proteins without a secretion signal, among them putative virulence factors\textsuperscript{180}. This underlines the importance of vesicles as an alternative secretion tool for eukaryotic cells.

Since fungi possess a cell wall, an open question is how fungal vesicles transit this barrier. It has been hypothesized, that vesicles escape through pores, in areas where the cell wall is thinned, which happens for instance during daughter cell budding, or by remodeling the cell wall through enzymes\textsuperscript{174}. Similar to other eukaryotic microbes, fungal vesicles may be utilized by pathogens to release virulence-associated components. For example, virulence factors are vesicle-associated in the human pathogens \textit{Cryptococcus neoformans} and \textit{Candida albicans}\textsuperscript{181,182}. However, virulence is unlikely to be the only function of fungal vesicles, since also the non-pathogenic \textit{Saccharomyces cerevisiae} releases vesicles\textsuperscript{183}.

Even plant cells were found to release extracellular vesicles, particular upon pathogen infection\textsuperscript{184}. Plant extracellular vesicles may facilitate the export of proteins without secretion signals, required for transport through the standard secretory pathway, and contain important defense compounds, such as hydrogen peroxide\textsuperscript{184}.

**Bacterial Membrane Vesicles**

In the 1960, the first bacterial vesicles were described in \textit{E. coli}\textsuperscript{185,186}, and later in several other Gram-negative species including \textit{Shigella} sp.\textsuperscript{187}, \textit{Salmonella} sp.\textsuperscript{188}, and \textit{Vibrio} sp.\textsuperscript{189}. In Gram-negative bacteria, vesicles derive from the outer membrane and are therefore called OMVs (outer-membrane vesicles). They contain outer membrane components but also inner membrane constituents and cytoplasmic elements. Even though vesiculation has been studied intensively in Gram-negative bacteria, attempts to generate a mutant deficient in OMV production failed, indicating that vesiculation is only partially under genetic regulation and additionally dependent on physical and biochemical processes. Since LPS is an integral component of the Gram-negative cell
surface, its disruption impacts OMV release. A model to explain vesiculation, especially in only partially hydrated environments is that nanopods are formed to release vesicles\textsuperscript{190}. The most recent model to describe OMV biogenesis in hydrated environments is based on phospholipid accumulation in the outer leaflet of the outer membrane, whereupon vesicles pinch off from the outer membrane\textsuperscript{191}. In addition to conventional one-bilayer OMVs, a second less abundant two-bilayer vesicle type, so-called outer-inner membrane vesicles (O-IMVs), was described, first in the Antarctic bacterium\textit{Shewanella vesiculosa}\textsuperscript{192} and then also in Gram-negative pathogens\textsuperscript{193}–\textsuperscript{195}. O-IMVs are formed when both plasma and outer membrane stretch out into the extracellular milieu and pull cytoplasmic content in the vesicle. O-IMVs can therefore contain cytoplasmic proteins and DNA\textsuperscript{193}.

OMVs derived from pathogens have been shown to be involved in bacterial virulence, by exerting functions such as delivery of biomolecules, including toxins, biofilm formation, adherence to the host and stress response\textsuperscript{196}. In non-pathogenic bacteria, OMVs play roles similar to other extracellular vesicles, such as cellular communication, surface modification and clearance of unwanted molecules\textsuperscript{175}.

In the past, the existence of MVs was out of question in Gram-positive bacteria, based on the fact that vesicular escape was precluded by the thick cell wall of these bacteria. Nevertheless, nowadays studies are exploring the field of MVs derived from Gram-positive bacteria\textsuperscript{174}. The composition and cargo of both Gram-negative and Gram-positive vesicles are illustrated in figure 10.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure10.png}
\caption{Model of the predicted composition of vesicles from Gram-negative and Gram-positive bacteria. The differences in cell wall architecture of Gram-negative and Gram-positive bacteria cause differences in vesicular composition as well. Gram-negative OMVs are typically surrounded by the outer membrane and contain periplasmic proteins. Gram-positive MVs, on the contrary, carry cytoplasmic components and are encased by the cytoplasmic membrane. Additional cargo is indicated. Reprinted with permission from Int. J. Mol. Sci.\textsuperscript{197}.}
\end{figure}
The first reports of Gram-positive derived MVs in *Bacillus cereus* and *Bacillus subtilis* \(^{198}\) dates back as far as the 1990s. However, these studies were not regarded for the next 20 years until MVs were described in 2009 in *S. aureus* \(^{199}\) and since then gained increased attention \(^{174}\). Data on the genetic regulation of MV biogenesis is still warranted. Also, how MVs escape the Gram-positive cell wall has not been conclusively clarified. However, similar to mechanisms suggested for fungal cell wall transit \(^{174}\), cell wall modification might be involved in MV release. For example, *S. aureus*-derived MVs contain peptidoglycan-degrading enzymes, which can alter the cell wall structure \(^{199}\). Another recently proposed mechanism is phage-triggered vesiculogenesis, where phage endolysins locally digest the cell wall to form holes through which the cytoplasmic membrane can stretch into the extracellular space to form vesicles \(^{200}\).

A wide range of cargo reflecting the cellular content has been associated with MVs from Gram-positive, including DNA, proteins and lipids. For example in *C. perfringens*, it was shown that DNA within MVs was protected from exonucleases \(^{201}\). This implies that DNA is protected from eventually harsh environmental conditions and may be transported over long distances and thus, MVs might play a role in gene transfer. The cargo of MVs is also associated with drug resistance, an example of this are *S. aureus*-derived MVs containing a penicillin-binding protein \(^{202}\). MVs can also act as vehicles for antimicrobial resistance-conferring proteins, for instance, *S. aureus*-derived MVs deliver β-lactamase proteins to ampicillin-susceptible bacteria which make them resistant to β-lactams \(^{203}\). Additionally, MVs may contain proteins which are important in immune system evasion and bacterial pathogenesis, such as virulence factors and toxins \(^{174}\). MVs may furthermore be key players in host-pathogen interactions, as they cause disease without the living bacterial cell \(^{204}\) and induce strong host responses \(^{205}\).
Objective

The overall aim of this PhD project was to investigate determinants contributing to virulence of nosocomial *E. faecium*. These include virulence factors, membrane vesicles as possible means of conveyance of these factors, and megaplasmids harboring virulence traits.

The specific objectives of **Paper I** are

1) To explore the epidemiology and genetic context of ORFs encoding TIR-domain containing proteins in *E. faecium*;
2) To evaluate the role of TIR-domain containing proteins in bacterial virulence through functional studies; and
3) To study whether TIR-domain containing proteins in *E. faecium* are secreted and whether they may be associated with membrane vesicles.

The specific objectives of **Paper II** are

1) To detect MVs derived from clinically relevant *E. faecium* and describe their morphology; as well as
2) To characterize the proteinaceous content of *E. faecium* derived MVs.

The specific objectives of **Paper III** are

1) To describe the genetic content of a megaplasmid originating from an *E. faecium* blood culture isolate and characterize traits contributing to bacterial virulence; and
2) To investigate the prevalence of these traits in the *E. faecium* population and their association to *rep* _pLG1_-type megaplasmids.
Choice of Methods

In general, the aim of research is to create reproducible and sustainable knowledge. Thus, methods, which meet the requirements of the modern scientific principles and deliver trustworthy data, are to be chosen. Different approaches illustrating the same phenomenon should be combined and the experimental design kept as simple as possible. Once a hypothesis is formulated, other explanations should still be taken into consideration. The methods should, therefore, be chosen to shine light on the research question from different angles. Still, every method will have its inevitable limitations, which have to be kept in mind when interpreting the data.

Virulence Factor Research

Prediction of Virulence Factors Using Computational Methods

The identification and subsequent characterization of virulence factors are important to understand the mechanisms underlying bacterial virulence as well as to identify novel points of intervention such as drug and vaccine targets. With the growing mass of available data on genome and protein sequences generated in the postgenomic age, it is an important first step to identify and describe virulence factors and other proteins of unknown function on their sequence information. Thereafter this information may be used to guide the way for experimental design of functional assays in the laboratory. The choice of in silico methods to predict virulence factors may be challenging, as platforms or data may be incomplete. Thus multiple approaches should be combined to validate or discard the initial finding.

The first bacterial TIR-domain containing proteins were found by detecting domains in silico and subsequent characterization. In this study, we choose a similar approach, where two putative E. faecium proteins annotated with a TIR-domain were identified in the protein database at NCBI and characterized using computational methods regarding properties such as protein family, secretion signals, and structure. Similar methods were also used to investigate the open reading frames adjacent of the identified TIR-domain containing proteins. After virulence factor identification it should be kept in mind that predicted domains and functions are still only predicted. Therefore, epidemiological and functional approaches, which were guided by the results of the in silico predictions, were included in the presented study.
**Epidemiology of Virulence Factors**

An epidemiological approach was used to assess the distribution of the identified virulence factors among *E. faecium* isolates of different origin. This strategy will show whether the virulence factor is rare or common and in what kind of isolates it is present. If a virulence factor-encoding gene is unique to a certain kind of isolate it can accordingly serve as an isolate-specific marker.

In this study, a combination of PCR screenings and WGS (whole genome sequencing) was used. On one hand, PCR might give a false negative result, for example, if the DNA was degraded, or the reaction does not work and depends on the utilized primers. Still, PCR is very simple, cheap and fast. WGS, on the other hand, is more time and resource intensive. However, it gives not only information about the presence of a gene, but the data can be further used to explore the genetic context of the gene and build phylogenetic trees. This is particularly interesting when looking at clonal expansion and transmission events. In the past, MLST-typing and PCRs were used, however, it recently becomes more and more apparent, that MLST typing is not reliable in *E. faecium*, because of the high recombination frequency in this bacterium. Phylogenetic trees can be especially interesting tools to study the emergence and spread of a virulence factor. From this data, one might speculate what kind of selective pressure lead to events in the evolution of a virulence factor.

**Functional Characterization of Putative Virulence Factors**

For functional studies, a combination of methods based on the purified virulence factor and methods comparing the characteristics of the wild-type and its isogenic mutant are used as complementary approaches.

The advantage of using purified recombinant proteins is, that the system allows for tight control. Additionally, purified recombinant proteins are used to generate antibodies, which are useful tools to detect the protein. However, the purification of a virulence factor might be challenging and many previous studies failed to purify full-length bacterial TIR-domain containing proteins. The disadvantage of systems using purified proteins is that the conditions of choice might be less physiological than other systems. Still, in combination with other – more physiological – systems, studies of the purified protein can provide important information.

In order to evaluate the immune evasive properties of a putative virulence factor, two main approaches are typically used. In the first approach, eukaryotic cells are transfected with an eukaryotic expression plasmid encoding the bacterial immune evasion factor to study its interference with the host intracellular signaling. When a receptor of the eukaryotic cell is now
stimulated it can be evaluated whether ectopic expression of the immune evasion factor interferes with the signaling of the receptor. In the second approach, the purified immune evasion factor is added to the host cells. Thereafter the receptor is stimulated and again it is evaluated whether the virulence factor interferes with the signaling of the receptor. For both approaches, the evaluation whether the virulence factor interferes with the signaling of the receptor can be done through a luciferase assay or by measuring the cytokines in the cell culture supernatant. In order to perform a luciferase assay the cells have to be transfected with another plasmid encoding a luciferase under NFκB control. When the receptors signaling now results in NFκB production, the luciferase is transcribed and can be quantified.

In this study, the second approach, where the purified protein is added to the cell culture, was used since proteins expressed in bacteria are more likely to have the correct conformation and we furthermore speculated that the virulence factors would be released by the bacterium and taken up by the cell. Whether the protein interferes with the receptor signaling was assessed by measuring cytokines in the supernatant, since cytokine release is a result of stimuli.

An isogenic mutant, lacking the virulence factor-encoding genes was created to compare it to the wild type and describe the phenotypic manifestation of the absence of the virulence factor. E. faecium is known to be difficult to manipulate genetically. In highly genetic unstable organisms such as E. faecium the generation of knockouts can be challenging, i.e. the deletion of only one gene or also the generation of complementary strains might be biologically impossible. Especially in those organisms, it should be verified that no mutations occur in the process; therefore the generated isogenic mutant was sequenced. The assays in which the isogenic mutant and the wild type were compared were chosen to represent challenges E. faecium would be likely to face upon infection, such as serum deposition, phagocytosis, and intracellular survival.

To prove the importance of a virulence factor in bacterial infection, mouse infection models are mostly warranted. However, nosocomial E. faecium are well adapted to the human host and even in humans act as rather opportunistic pathogens, which only cause disease under certain preconditions, such as immunodeficiency. Thus, it is questionable how well a mouse infection model could represent the situation under which E. faecium would cause human infection and how much can be extrapolated from a mouse model.

Therefore fresh human blood was used as the model of choice in this study. Still, also this model does not take into account the pre-adaption E. faecium might undergo during the transition from the intestine or the environment to the bloodstream. However, the blood survival model represents the environment E. faecium would encounter upon infection rather closely.
Studying Membrane Vesicles

Isolation of Vesicles and Their Morphological Characterization

Vesicular research was hampered in its very beginning because it was thought that observed vesicle-like structures were only cellular debris. Even after the description of OMVs in Gram-negative bacteria, the scientific community precluded the existence of MVs in Gram-positive bacteria based on the assumption that vesicles could not escape a barrier such as the Gram-positive cell wall. Now it is known, that living cells are required for vesicle production, and that dead cells did not produce vesicles, showing that vesiculation is a physiological process \(^{174}\).

However, there is still no standardized protocol for the preparation of a pure vesicle sample. Based on observation in different bacterial species it is likely that the method needs optimization for every bacterial strain and the chosen method should be verified using different methods for every preparation to ensure the desired quality \(^{209}\). A typical workflow of MV preparation is shown in figure 11. The isolation of MVs from different \(E.\ faecium\) strains was performed based on an isolation method described previously for the Gram-positive \(S.\ aureus\) \(^{210,211}\). First, a bacterial culture is obtained under the desired growth conditions, from which then the supernatant is sterilized through filtration. Second, the bacteria-free supernatant is concentrated in centrifugal filters with a size cut off at 100 kDa and third, the concentrate is ultracentrifuged to pellet MVs. At this stage, the vesicular pellet may be contaminated with proteins, pili or phages. Therefore, a purification step using density gradient centrifugation is included. The density gradient medium of choice is OptiPrep, an iodixanol based solution. OptiPrep was shown to be superior to other density gradient solutions since it prevents damage of frail membranous structures \(^{212}\). The vesicle sample is mixed with a high-density solution and transferred to the bottom of an ultracentrifuge tube and overlaid with step gradients of low-density solutions. The vesicles have a lower density then cell debris due to the lipid content. Thus, during centrifugation, the vesicles migrate to a layer according to their density and form a ring. The fractions are then collected and the density gradient medium may be exchanged to PBS. Another approach is top loading, where the vesicle sample is placed on top of the last low-density solution. During the cause of this study also top loading was tested, but bottom loading gave the best results for the purification of \(E.\ faecium\) MVs. We experienced that with bottom loading contaminating proteins and cell debris stay in the bottom of the tube with high-density solution, while the lighter MV containing fractions travel upwards to low-density solution.
Typically, the quality and the properties of the MV sample and the properties of MVs are assessed through visualization and content analysis. This is done using transmission electron microscopy, which allows the characterization of the MVs sizes, shapes, and overall appearance. Moreover, this method will assess the purity of the sample, regarding non-MV material such as protein aggregates, phages or pilis. As a complementary approach, atomic force microscopy was conducted to visualize the release of MVs by the bacterium to the environment. This technique provides a three-dimensional surface profile, and MVs may be seen next to the bacterium. Furthermore, the size distribution, as well as the membrane potential, was measured. Additionally, lipid staining and immunoassay for the presence of peptidoglycan, the key cell wall component, were performed.

![Diagram of MV preparation workflow](image)

**Figure 11: Typical MV preparation workflow.** First, bacteria are cultured in liquid media (Medium composition, harvesting time point, and stress condition may be varied). Second, bacteria are removed by sterile filtration. Third, the culture supernatant is concentrated. Fourth, MVs are purified according to the required purity. Reprinted with permission from Microbiol. Res. 213.
Identification of Vesicle-associated Proteins in Mass Spectrometry

MV-associated proteins were identified by means of mass spectrometry. This is done through sample digestion and fractionation prior to introduction in a mass spectrometer. The experimentally obtained fragmentation spectra are matched with a database for identification. This database contains simulated spectra generated through \textit{in silico} digestion of protein sequences extrapolated from DNA sequences. There are a few alternative approaches, such as protein purification followed by residue-by-residue sequencing. However, this approach is not feasible at proteome level. Moreover, one should be aware that mass spectrometry holds its limitation regarding protein digestion and peptide-matching algorithm \(^{214}\). In addition, it should be taken into account that the databases are being generated from a specific DNA sequence \(^{214}\). The whole genome sequence of each \textit{E. faecium} strain itself was used as database for the respective sample to include all strain-unique proteins. The mass spectrometry obtained data was used in a descriptive way to characterize MV-associated proteins.

Different computational tools were used to describe the localization origin and the amount of secreted MV-associated proteins. The combination of different tools allowed comparison and validation of the data obtained by each individual approach. Further analysis of MV-associated proteins included searching for known virulence factors, vaccine candidates, and antimicrobial resistance determinants. It should be noted that merely the presence of a protein with a certain ascribed function does not necessarily translate to the protein executing that function in the context it was found. Due to the above-mentioned challenges and limitations, the MV study in \textit{E. faecium} was kept descriptive. However, functional studies are warranted to describe the physiological role of enterococcal MVs.

Investigating the Genetic Content of Megaplasmids

Identification of Megaplasmid-Encoded Factors Contributing to Virulence

Since the goal of this research was to identify factors, which contribute to the virulence of \textit{E. faecium}, the contents of a megaplasmid from a blood culture isolate was investigated.

Factors other than antimicrobial resistance determinants were suggested to be involved in enterococcal gastrointestinal colonization, and genes encoding these factors were located on a megaplasmid \(^50\). Transferable megaplasmids larger than 145 kb harboring the \textit{hyl\textsubscript{\textit{E}fm}}-gene, named after a putative glycosyl hydrolase, were shown to promote virulence in peritonitis \(^{215}\), but the
hyI_Efa-gene itself was found not to be involved in this effect \(^{216}\). Thus, it was suggested that the plasmid content should be investigated for other virulence promoting factors.

Initially, megaplasmids of nosocomial *E. faecium* strains were sequenced and putative secreted proteins, including hypothetical secreted proteins, were selected for further analysis. The hypothesis was that secreted megaplasmid-encoded factors could interfere with the host’s immune defense and enable *E. faecium* to invade the host. The software SignalP was used to predict putative secretion signals encoded by the ORFs. RT-PCR was then used to assess whether these putative secretory factors are transcribed. Furthermore, their prevalence in nosocomial versus commensal strains was investigated by screening through PCR. A set of more than ten putative secretory factors, which were expressed and found in blood culture isolates, was heterologously expressed in *E. coli* without the secretion signal and successfully purified. The purified proteins were tested, one at a time, in a wide range of immune evasion assays, including binding to leucocytes, complement evasion, serum killing, red blood cell lysis and phagocytosis assays. However, none of the recombinant proteins showed an activity in these assays (unpublished results), and we hence choose to examine the whole megaplasmid in a more biologically relevant setting.

Therefore, it was investigated whether the megaplasmid could be transferred into a plasmid-free strain and enhance its survival in human blood. Exposing the bacteria to whole human blood was used to mimic the environment *E. faecium* would be facing during bacteremia, one of the most serious enterococcal infections. The megaplasmid from a bacteremia isolate (K60-39) was therefore transferred into other strains. The first transfer was done into a plasmid-free laboratory strain, which has been used in previous mating experiments (BM4105-RF \(^{217}\)). The plasmid was additionally transferred into a plasmid-free community isolate (E1007 \(^{140}\)), which was previously shown to be serum susceptible \(^{140}\). Different transconjugants resulted from the mating and comparing them in a blood survival assay gave us the possibility to identify factors, which contribute to the blood survival advantages.

**Characterization of Virulence Determinants in Megaplasmids**

To find a function of putative virulence factors, an often-used strategy is to purify the candidate protein and test it in a multitude of immune evasion assays, as mentioned above. For example, putative virulence factors could bind to blood cells and thus hinder bacterial recognition, play a role in complement evasion, interfere with serum killing, hinder phagocytosis and/or help bacterial intracellular survival.

However, the heterologous expression and purification of putative virulence factors may be difficult or in some cases even impossible and thus rule out the characterization of certain proteins.
Furthermore, these assays are mainly used to test one protein at a time and it is therefore difficult to examine proteins, which function as multimers or act together in a biological pathway. Another limitation is, that evasion assays are designed based on previously described phenomena but might not be able to detect novel functions.

Since the putative virulence factor candidates did not show a function in any of the immune evasion assays, we therefore choose to use bioinformatic tools to describe factors putatively contributing to *E. faecium* survival in blood.

The transconjugants were sequenced and comparison of the ones replicating in human blood and the plasmid, which did not promote survival, revealed a list of putative blood survival promoting factors. These factors were then bioinformatically characterized and screened for in the NCBI database. Based on these data, hypotheses could be constructed on how these factors might aid bacterial survival. Still, functional studies would be warranted to conclude their mechanism of action.

To assemble the chromosome and plasmids of the original donor, it was sequenced using PacBio, a long-read high-quality technique. A hybrid sequencing approach was used to describe the genomic contents of the transconjugants. At first, the transconjugants were sequenced using Illumina, a very accurate sequencing technique that produces short reads. At a second stage, the transconjugants were sequenced using Oxford Nanopore Technology. Since enterococci contain numerous long repetitive sequences, such as IS elements, the long reads of Nanopore are necessary to pass these regions and close the plasmids (ongoing analysis and unpublished results).
Summary of Main Results

Paper I: *Enterococcus faecium* TIR-domain Genes Are Part of a Gene Cluster Which Promotes Bacterial Survival in Blood

*E. faecium* TIR-domain containing proteins (TirEs) were identified, and their structure showed a typical TIR-fold. The genes encoding them, *tirEs*, were part of a putative mobile genetic element of phage origin exclusive to nosocomial *E. faecium* isolates.

TirEs were detected within the bacterial supernatant even though they do not have a signal peptide. A novel release mechanism was identified since TirE2 was found within enterococcal membrane vesicles.

The *tirE* gene cluster did not influence cytokine release, phagocytosis or bacterial survival within macrophages. However, the *tirE* genes were predominant in blood culture isolates, expressed upon blood exposure of *E. faecium* and the *tirE* gene cluster promoted bacterial proliferation in human blood.

Paper II: *Enterococcus faecium* Produces Membrane Vesicles Containing Virulence Factors and Antimicrobial Resistance Related Proteins

Membrane vesicles (MV) were produced by *E. faecium*. They were spherical in shape, range in size from 37±23 nm to 83±29 nm, and were negatively charged. *E. faecium* derived MVs were associated with lipids and peptidoglycan.

The proteinaceous content and the number of MV-associated proteins depended on strain and growth condition, but the overall proteomic pattern regarding the percentage of proteins that derived from the core genome, their origin of localization, and whether they had a signal peptide or not, was similar among *E. faecium* derived MVs.

11 previously known virulence factors were found to be MV-associated. Among them were adhesins (CapD, PrpA), collagen binding proteins (Acm, Scm), and fibronectin binding protein (Fnm), as well as biofilm promoting factors (AflA, Esp, SagA). Further, vaccine candidates, as well as antimicrobial resistance related proteins, were MV-associated, including all proteins of the *vanA*-cluster of a vancomycin-resistant strain.
Paper III: Megaplasmid-encoded Metabolic Factors Enhance Proliferation of Enterococcus faecium in Human Blood

A repLG1-type megaplasmid encoding high-level gentamicin resistance and a T4SS was transferrable from a bloodstream isolate via a laboratory strain to a commensal strain. In the commensal strain the plasmid conferred a significant increase in bacterial proliferation in human blood.

The genes encoded by this megaplasmid were mainly involved in primary metabolic processes, such as carbohydrate metabolism, membrane transport and information processing. The replication initiation protein-encoding gene repLG1 was prevalent in E. faecium of diverse origin, including bloodstream as well as commensal isolates. However, the blood proliferation promoting candidate genes were significantly enriched among clinical E. faecium isolates. We found indications for exchange of gene clusters between plasmid and chromosome, suggesting that megaplasmids serve as a gene reservoir for niche adaption.
General Discussion

From a harmless gut commensal, *E. faecium* has undergone a transition towards a multi-resistant nosocomial pathogen, posing a major threat to human health, especially in immunocompromised individuals\(^1,68\). Not only does the hardy bacterium survive over prolong periods of time on various surfaces\(^98\), resist extreme conditions and even alcohol preparations\(^96,97\) but *E. faecium* also acquires genetic material tremendously fast\(^11\), which enables it to adapt to new selective pressures.

The aim of this thesis was to shed light on virulence-related aspects of nosocomial *E. faecium* from different angles, including virulence factors utilized for immune evasion (Paper I), membrane vesicles containing virulence factors amongst other cargo (Paper II) and megaplasmids encoding factors aiding *E. faecium* survival in human blood (Paper III).

In *E. faecium* the search for new virulence factors might be hampered since the bacterium is foremost a commensal and does not natively display high virulence\(^208\). In enterococci, many factors contribute to fitness within the human host as well as fitness in other ecological niches. Thus, pathogenicity- or fitness-enhancing factors are probably not only virulence factors in a strict sense. “Virulence factors” are defined as substances necessary for causing disease in the host, but not necessary for survival in other contexts\(^208\). However, determinants, which also play a role in other contexts than causing disease, such as metabolic genes, can be important in colonization and/or in out-competing other bacteria. It seems that the virulence of enterococci depends on more than a few main players, and is likely to depend on strain variable combinations of factors that jointly enhance infection\(^208\). Also from the highly plastic and heterogeneous genomes, especially of the nosocomial clade\(^11,14\) it is verisimilar that *E. faecium* does not employ just one winning strategy, but that multiple factors aid its adaption to the hospital environment. Our findings reinforce this hypothesis since various proteins – resembling metabolic as well as virulence factors – were found to be associated with MVs of nosocomial *E. faecium* (Paper II). Also, the here investigated megaplasmids harbor multiple – mainly metabolic – factors (Paper III).

The prevalence of megaplasmids and their genes promoting proliferation in human blood among nosocomial isolates indicates that these metabolism-related genes contribute to the infectious phenotype of *E. faecium* (Paper III). The function of the factors was predicted based on bioinformatics analysis, which provides a useful basis to forecast how they work, but further
functional studies would be needed to unravel their specific roles and interplay. The main functions of the megaplasmid encoded genes are presumably primary metabolic pathways, especially carbohydrate metabolism (Paper III) and it consequently stands to reason that E. faecium has an advantage if it can utilize different nutrient sources in the course of infection where it would face varying environments. Carbohydrate uptake and metabolism genes are also the genes with the largest difference between human commensal and the nonsocial strains\(^{11}\) and carbohydrate utilization has been described as a driver in adaption to niches like a hospital environment and host colonization\(^{218,219}\). A limitation of this study is that we focused on analyzing genes, which are present, but it is difficult to conclude on the importance of the absence of genes. It is apparent, that smaller plasmids might pose a lower cost and therefore be advantageous. It could also be advantageous for the bacterium if genes, which have an adverse effect on blood survival, are lost. A description of the exact topology and closing of the plasmids might help to reveal other differences between the two transconjugants, which both enhanced proliferation in blood but to varying degrees, such as gene synteny and regulation by introduction or removal of mobile elements. Furthermore, to generalize these finding, a larger number of megaplasmids would need to be investigated, since this study presents foremost a case study of one megaplasmid and the transfer of it to the commensal might have been only possible because of the interim laboratory strain (Paper III). Still, the findings of this study can guide future work.

Even if it might be difficult to narrow the many facets of E. faecium down to single players, virulence factor research is important in order to understand pathogenicity. Here, we describe multiple virulence factors to be associated with MVs (Paper II), among them TirEs, which we describe as novel virulence factors of E. faecium (Paper I). A strategy to search for virulence factors is to predict those with secretion signals, using tools like SignalP or PrediSi, as it is anticipated that secreted proteins are more likely to interact with the host. Generally, exported proteins are synthesized with an N-terminal signal peptide directing them to the respective transport pathway\(^ {220}\). However, if virulence factors, as described here for TirE (Paper I) and other virulence factors (Paper II), are found in the bacterial supernatant even though they lack a secretion signal peptide and/or packed in MV, one would have to think differently when predicting putative virulence factors. Since we found that the TirE locus enhances blood survival and purified TirE down-regulated the cytokine (IL-8) response (Paper I), we hypothesize that TirE will also interact with the human host in a bloodstream infection. Unfortunately, we were not able to make knockouts of the single genes
of the TirE locus and we thus could not assign individual functions to the single genes. Besides their function as immune evasion proteins, bacterial TIR-domain containing proteins probably also have bacterial proteins as interaction partners. They might not necessarily have their primary function as agents of subversion since TIR-domains are foremost protein-interaction domains and also found in non-pathogenic bacteria. Bacterial TIR-domains were for example proposed to be involved in nucleic acid binding or as nucleases. Recently, an intrinsic NADase activity was described in bacterial, mammalian and achaeal TIR-domains and TIR-domain containing proteins proposed to influence metabolic and energetic pathways. \textit{S. aureus} TirS degrades NAD$^+$ also in mammalian cells, suggesting that its function as virulence factor could be due to its NADase activity. The precise interaction partner of TirE, of bacterial and/or host origin, remains to be identified by future studies.

An important factor in pathogenicity is the ability of virulence factors to spread, both in terms of the virulence factor protein itself to reach its target but also virulence encoding genes to other bacteria. In enterococci risks arise from HGT because these rather less-pathogenic bacteria are able to exchange genes encoding antimicrobial resistance and virulence factors with other microbes, including more pathogenic bacteria. For example, MRSA (methicillin-resistant \textit{S. aureus}) became high-level resistant to vancomycin through an \textit{E. faecium} multiresistance conjugative plasmid with Tn1546 harboring \textit{vanA} resulting in a virtually untreatable VRSA (vancomycin-resistant \textit{S. aureus}). Here, we furthermore found that a megaplasmid carrying determinants for high-level gentamicin resistance and determinants involved in blood survival is transferable from nosocomial \textit{E. faecium} via a laboratory strain to a commensal \textit{E. faecium} strain (Paper III). This points out the potential of \textit{E. faecium} to gather and spread resistance and/or virulence genes to both pathogens and commensals, increasing the virulence of its cohabiters.

We found that the genes encoding two TirEs along with the intermediate gene Hp1 and adjacent genes are localized on a putative MGE flanked by integrases and is exclusive to clade A, suggesting lateral transfer (Paper I). A mobile genetic context is found for other \textit{E. faecium} virulence factors. Genes encoding phage integrase family proteins flank the gene cluster of the carbohydrate phosphotransferase system (Pts) which is enriched in clinical strains and absent in commensal isolates. Furthermore, a putative transposase was found upstream of \textit{fmr} \textit{113}. Except for that, none of the other virulence factors of \textit{E. faecium} were yet described to be associated with putative MGEs. Similar to genes encoding TirEs, the virulence factors SgrA, EcbA, PilA and PilB, BepA were enriched in nosocomial isolates; and genes encoding Esp, Acm, and SagA were found in
different variants in community versus nosocomial isolates. Thus it would be interesting to explore the genetic context of these virulence factors to link them to MGEs, since their predominance in the nosocomial clade suggests lateral transfer. Subsequently, HGT of *E. faecium* virulence factors localized on putative MGEs should be verified experimentally. We tested the transferability of the putative MGE-*tirE* in a preliminary experiment. We used a strain where *tirE* was replaced by a gentamicin resistance gene as a donor in a filter-mating experiment and plated the transconjugants on gentamicin plates. Unfortunately, we could not confirm the transferability of the putative MGE (unpublished result), but we speculate that it might need the help of a conjugative plasmid, as was previously described for other *E. faecium* MGEs\(^\text{224,225}\). It would hence be worthwhile to test mobilization of MGE-TirE and other *E. faecium* genes encoding virulence factors located on putative MGEs by introduction of a conjugative plasmid known to mobilize genes in *E. faecium*.

Gene organization of bacterial TIR-containing proteins were previously studied in the closely related *E. faecalis*\(^\text{170}\) and in the zoonotic *S. aureus* ST398\(^\text{173}\). In *E. faecalis* two variants of gene arrangement were found: a TIR-domain containing tcpF gene appeared by itself or the tcpF gene was intersected by a GTPase\(^\text{170}\). However, no functional studies were conducted on the intersected variant. In *S. aureus* ST398 the two genes, *saTlp1* and *saTlp2*, encoding TIR-domain containing proteins lie adjacent to each other\(^\text{173}\). Hp1 is inserted between the two genes encoding *E. faecium* TIR-containing proteins presented here (Paper I). Since we were able to link these genes by RT-PCR, we suggest that they are expressed as one operon. BLAST-searches identified Hp1 as a SIR2-like domain-containing protein. These proteins have been described as deacetylases\(^\text{226}\), but in initial experiments, no deacetylase activity was detected (unpublished result). The function of Hp1 thus remains elusive and should be investigated in future studies.

Interestingly a correlation between the lack of *hp1* in the *tirE*-gene cluster and the presence of *vanB* was found (Paper I). In previous studies in German and Polish isolates, *hp1* was found to be an entry point for the *vanB*ICE\(^\text{82,83}\). This could present a link between virulence factors and antimicrobial resistance genes, similar to an MGE of *S. aureus*. TirS localizes within the MGE *SCCmec* together with genes encoding fusidic acid resistance and exposure to fusidic acid further increased the expression of *tirS*\(^\text{172}\). Thus, it would be interesting to investigate in strains lacking Hp1 but encoding *vanB* whether *hp1* is an entry site for the *vanB*ICE and if vancomycin exposure has an impact on *tirE* expression. Moreover, *vanB* could potentially be mobilized together with *tirE*, as it would localize on a putative MGE if *hp1* serves as an entry point, linking two clinicopathologic features, immune evasion and antimicrobial resistance and the spread of both.
A novel mode of *E. faecium* cargo dissemination discovered here are MVs (*Paper II*). Whether *E. faecium* MVs contain nucleic acids and act as means for HGT, remains to be examined by future studies. We investigated the proteinaceous content of MVs from four different clinically relevant strains and found both antimicrobial resistance-conferring proteins as well as virulence factors associated with the MVs (*Paper II*). We hence hypothesize that other bacteria in the community potentially could utilize these factors. The virulence factors we discovered, however, represent a snapshot of what the bacteria release in medium in monoculture, and their contents might be altered upon exposure to stressors such as antimicrobials, the presence of a host or competing microorganisms.

The findings of the study on *E. faecium* MVs are strengthened using four different clinically relevant strains in two different growth conditions. This first report of *E. faecium* MVs was purely descriptive, and functional studies are needed to further understand the purpose of MVs for *E. faecium* itself. An active metabolism is required for the MV production, and MVs are created through a physiological process rather than being artifacts. Still, it is not clear whether *E. faecium* MV release is an active process including sorting mechanisms or mainly whether MVs act as a waste disposal system.

Since we found various proteins to be MV-associated (*Paper II*), it is likely that enterococcal MVs play important roles in bacterial communities similar to other bacterial MVs, for example by delivering enzymes to break down extracellular material into nutrients or by acting as decoys for bacteriophages or antimicrobial agents in the bacterial immune defense. Apart from functions of *E. faecium* MVs in bacterial physiology, their impact on host responses will be an interesting field to study, especially for the utilization of MVs as probiotics or vaccines. For example, *E. coli* OMVs have potential as probiotics and this application might also be interesting for *E. faecium* as they normally co-exist as commensals within the human intestine.

While investigating MV-associated proteins, we found previously described vaccine candidates (*Paper II*). In times of increasing antimicrobial resistance vaccines are one of the most promising alternatives, since they bear a low risk to develop resistance and are used prophylactically, rather than after an infection is established and the host’s health is impaired. One might ask whether it is a good idea to vaccinate against bacteria such as *E. faecium*, which mainly co-exist as commensals being a beneficial part of the healthy intestinal flora. However, a vaccination should mainly prevent bloodstream infections, since IgG antibodies would primarily reside in the blood and not the intestine where commensals reside. Another strategy could be, to find vaccine
candidates exclusive to nosocomial strains and hereby exploit the differences between commensal
and nosocomial strains. For the previously described vaccine candidates, antibodies against SagA
mediate opsonic killing in human blood $^{231}$, and the two metal binding lipoproteins elicit IgG
response $^{232}$. It is thus likely that MVs containing vaccine candidates would elicit a similar
response. Capsular polysaccharides, as the previously described vaccine candidates, are often poor
immunogens, but their immunogenicity can be enhanced by conjugation with a carrier protein $^{233}$. If
the whole vesicle preparation would be used though, the vesicle itself could be the immunogenic
$^{234}$. Still, an enterococcal vaccine should be carefully evaluated to rule out that it does not eradicate
asymptomatic carriage resulting in undesired changes in microflora with unknown future
consequences. Also, a vaccine would probably be most useful if given only to certain well-defined
risk groups or on passive immunotherapy since patients at bacteremia risk are likely to have an
impaired immune response $^{233}$.

Other suggested therapeutic uses for MVs include drug-delivery in cancer therapy $^{235}$, as antibiotic-
delivery to treat biofilms $^{236}$, and alleviation of allergy and mediate immune regulation $^{237,238}$. Immune regulation could also be an application of bacterial TIR-proteins, which interfere with TLR
signaling (Paper I). Potential drugs could be designed as a more potent version of bacterial TIR-
proteins and be used in diseases with excessive TLR signaling (e.g. rheumatoid arthritis, systemic
lupus erythematosus, colitis) $^{239}$.

Even plasmids, which normally are the main players in spread of antimicrobial resistance and
virulence (Paper III), could bear potential as use for antibacterial treatment and a number of
approaches have been suggested. Conjugation could be inhibited to prevent the transfer of plasmids
between bacteria and several relaxase inhibitors have been identified $^{240-242}$. To eliminate multi-
resistance plasmids from the bacterial population and resensitizing the bacteria to antimicrobial
agents, plasmid replication could be inhibited, for example by mimicking plasmid incompatibility
through small molecules $^{243}$. Furthermore, plasmid maintenance systems could be exploited to
directly and selectively induce death in drug-resistant bacteria $^{244,245}$. No compounds based on these
strategies have made it yet to clinical application but the high prevalence and threat of resistance
plasmids makes them attractive target points $^{246}$ and studying their biology and content will
hopefully help to advance the application of the knowledge gained on megaplasmids.
Conclusion

Over the last decades, *E. faecium* has arisen as a multi-resistant pathogen and research on both antimicrobial resistance and virulence of the bacterium has increased. Still, many questions remain unsolved and the aim of this thesis was to expand the knowledge on determinants contributing to the virulence of nosocomial *E. faecium*.

The findings of this thesis advanced the field of *E. faecium*’s virulence in different aspects. First, TIR-domain containing proteins, TirEs, of *E. faecium* are identified and characterized as novel virulence factors located on a putative MGE of phage origin and exclusive to nosocomial strains (Paper I). Second, membrane vesicles, MVs, are described for the first time in enterococci and found to contain varying cargo, including virulence factors, antimicrobial resistance related proteins as well as vaccine candidate proteins and thus might play an important role in virulence (Paper II). Third, a megaplasmids of nosocomial *E. faecium* is shown to promote blood survival, presumably through genes encoding factors involved in metabolic processes (Paper III).

Our results are in line with previous studies, stating that the nosocomial clade harbors different traits from the commensal clade. The described virulence factors, TirEs, are exclusive to nosocomial strains (Paper I) and also the megaplasmid-encoded factors promoting blood survival are significantly enriched in nosocomial strains (Paper III). With the MVs we identified a novel mode of cargo dissemination of *E. faecium* (Paper II) and this research bears much potential for future functional studies, in regards of interaction with the host as well as the bacterial community, nutrient acquisition, or antimicrobial resistance.

In the future, this research on the virulence of nosocomial *E. faecium* will hopefully provide valuable groundwork to understand *E. faecium*’s relationship to the human host and inspire novel therapy development.
References


4. Higuita, N. I. A. & Huycke, M. M. Enterococcal Disease, Epidemiology, and Implications for Treatment. (Massachusetts Eye and Ear Infirmary, 2014).


References


References


170.  Kraemer, T. D., Quintanar Haro, O. D., Domann, E., Chakraborty, T. & Tchatalbachev, S. The TIR Domain Containing Locus of *Enterococcus faecalis* Is Predominant among Urinary


195. Devos, S. et al. Membrane vesicle secretion and prophage induction in multidrug-resistant 
3937 (2017).

196. Roier, S., Zingl, F. G., Cakar, F. & Schild, S. Bacterial outer membrane vesicle biogenesis: 


198. Dorward, D. W. & Garon, C. F. DNA Is Packaged within Membrane-Derived Vesicles of 

(2009).


201. Jiang, Y., Kong, Q., Roland, K. L. & Curtiss, R. Membrane vesicles of *Clostridium 
perfringens* Type A strains induce innate and adaptive immunity. *Int. J. Med. Microbiol. IJMM* 


203. Lee, J. et al. *Staphylococcus aureus* extracellular vesicles carry biologically active β-

204. Hong, S.-W. et al. Extracellular vesicles derived from *Staphylococcus aureus* induce atopic 


genomic relationships between isolates with independent emergence of vancomycin resistance. 

207. Bardoel, B. W. et al. Evasion of Toll-like receptor 2 activation by staphylococcal 


