The Influence of *Streptococcus pneumoniae* Serine Proteases on Pneumococcal Pathogenesis

I n a u g u r a l d i s s e r t a t i o n

zur

Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Universität Greifswald

vorgelegt von

Murtadha Qays Ali
geboren am 13.08.1990
in Babylon, Irak

Greifswald, 22.04.2022

Dekan: **Herr Prof. Dr. Gerald Kerth**

1. Gutachter: **Herr Prof. Dr. Sven Hammerschmidt**

2. Gutachter: **Frau Prof. Dr. Simone Bergmann**

Tag der Promotion: **22.04.2022**

إِنَّمَا يُوَفَّى الصَّابِرُونَ أَجْرَهُم بِغَيْرِ حِسَابٍ

"Those who endure with patience will be rewarded without measure."

(Al Qur'an, Al- Zummar 39:10)

Biofilm formation by *Streptococcus pneumoniae* on live epithelial cells

300nm
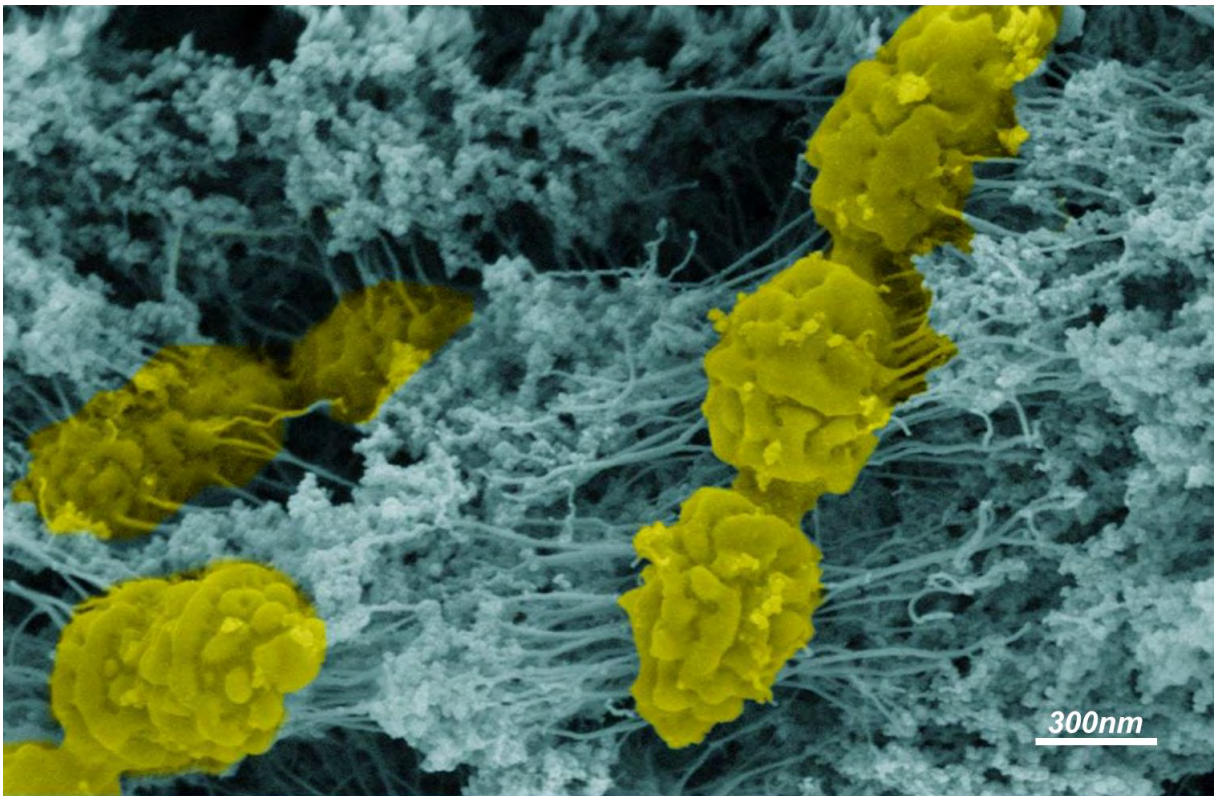
"Pneumococcus is altogether an amazing cell. Tiny in size, simple in structure, frail in make-up, it possesses physiological functions of great variety, performs biochemical feats of extraordinary intricacy and, attacking man, sets up a stormy disease so often fatal that it must be reckoned as one of the foremost causes of human death.

Dr. Benjamin White, 1938 in "The biology of pneumococcus".

<u>Dedication</u>

I dedicate my thesis to……

Grandmother, Father & Mother

Your inspiration, love, encouragement, and prayers of day and night made my success possible and honor…

To…

Sister and brothers

Along with full honor and respect to ….
Those who left and their spirit still around me…

The martyrs my uncles

Funding

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1
ZUSAMMENFASSUNG

1.   Zusammenfassung

Streptococcus pneumoniae (Pneumococcus) ist ein opportunistischer, humanpathogener Mikroorganismus, der lebensgefährliche Krankheiten wie z.B. Lungenentzündungen, Sepsis, Meningitis oder nichtinvasive Infektionen, wie Otitis media, verursacht. Pneumokokken haben mehrere Strategien entwickelt, um das Epithelium der oberen Atemwege von Menschen zu besiedeln. Die Besiedlung des Epithels wird direkt durch Interaktionen mit den Rezeptoren der Wirtzelle oder indirekt durch die Bindung an die extrazelluläre Matrix (ECM) vermittelt. Allerdings muss *S. pneumoniae* während einer Infektion Gewebebarrieren überwinden, die u.a. vom Immunsystem beschützt werden. Deswegen haben Pneumokokken eine Vielzahl von Mechanismen entwickelt, um die antibakterielle Funktion des Immunsystems zu umgeh n wie, z.B. die Serinprotease-Aktivitäten einiger Proteine. Serinproteasen. Sie gehören zu der am meisten verbreiteten und funktionell diversen Gruppe von Proteinen in eukaryotischen und prokaryotischen Organismen. Jedoch sind die Epithelbarrieren, Integrine und andere Zelloberflächenrezeptoren zunächst unzugänglich für Pneumokokken im Nasenrachenraum. Daher benutzen Pneumokokken auch extrazelluläre Serinproteasen des Wirtes, wie z.B. Plasmin(ogen), um ECM und Mucus zu degradieren, was zu einer verstärkten Bindung der Pneumokokken an Epithel- und Endothelzellen führt. Serotyp-abhängig exprimiert *S. pneumoniae* bis zu vier Serinproteasen: HtrA, SFP, PrtA und CbpG, die sich alle auf der bakteriellen Oberfläche befinden. Die Enzyme sind der Kategorie Trypsin-ähnlicher oder Subtilisin-ähnlicher Proteasen zugeordnet. Sie sind durch die Anwesenheit von drei konservierten Aminosäuren im aktiven Zentrum, nämlich Ser-His-Asp, charakterisiert. Diese katalytische Triade ist entscheidend für die Spaltung von Peptidbindungen. Studien, die sich mit der Inaktivierung von einzelnen Serinproteasen der Pneumokokken beschäftigen, sind wegen der potenziellen, kompensatorischen Effekte durch die anderen Serinproteasen schwierig zu interpretieren.

Zunächst wurde eine umfassende *in silico* Analyse der Verteilung und Konservierung von Serinproteasen der Pneumokokken in dieser Arbeit durchgeführt. Interessanterweise waren die Gene, die für PrtA, HtrA und CbpG kodieren, in den 11 analysierten Stämmen hoch konserviert. Überraschenderweise war das Subtilisin-ähnliche Protein SFP in einigen den untersuchten Stämmen nicht im Genom vorhanden. Daher besitzen Pneumokokken mindestens drei der vier möglichen Serinproteasen, so wie es für den Stamm 19F_EF3030 in dieser Arbeit gezeigt wurde. Computer gestützte Analysen zur Struktur dieser Proteasen zeigten bei den 3D Strukturmodellen eine hohe Ähnlichkeit in den katalytischen Domänen zwischen HtrA und CbpG bzw. zwischen PrtA und SFP.

Der Fokus dieser Arbeit war die Funktion von einzelnen extrazellulären Serinproteasen von Pneumokokken in Bezug auf die Pathogenese zu untersuchen; insbesondere den Einfluss auf die Adhärenz, Besiedelung, Virulenz, und Entstehung von Biofilmen. Deswegen wurden Doppel- und Dreifach-Deletionsmutanten im *S. pneumoniae* Serotyp 19F Stamm EF3030 und im invasiven Serotyp 4 Stamm TIGR4 generiert. In Adhärenzstudien, in denen humane Detroit-562 Epithelzellen eingesetzt wurden, konnte gezeigt werden, dass sowohl TIGR4Δ*cps* als auch 19F Mutanten, die keine oder nur eine Serinproteasen (CbpG, HtrA oder PrtA) exprimierten, eine verminderte Fähigkeit haben, an Detroit-562 Zellen zu binden. In einem experimentellen Kolonisierungs-Mausmodell wurde gezeigt, dass die Inaktivierung von Serinproteasen im Stamm 19F_EF3030 die nasopharyngeale Kolonisierung bei CD-1-Mäusen beeinflusst. Die Bakterienlast im Nasenrachenraum wurde über einen Zeitraum von 14 Tagen beobachtet. Die untersuchten Mutanten zeigten hierbei eine signifikant geringere Bakterienzahl im Nasenrachenraum an den Tagen 2, 3, 7 und 14 nach der intranasalen Inokulation.

Anschließend wurde der Einfluss der Serinproteasen bei invasiven Infektionen mittels eines *in vivo* Pneumonie-Maus-Infektions-Modells und durch *in vitro* Phagozytose überprüft. Dazu wurden Mäuse intranasal mit dem invasiven TIGR4 (Biolumineszenz-markiert) Wildtyp oder mit einem der 3-fach-Deletionsmutanten, die nur CbpG, HtrA, PtrA oder SFP exprimierten, infiziert. Die akute Lungeninfektion wurde in Echtzeit mithilfe eines IVIS®-Spektrum visualisiert. Die TIGR4*lux*-Mutante, die nur PrtA exprimiert, zeigte eine signifikante Abmilderung der Infektion im akuten Pneumonie-Modell. Die Versuche zur Phagozytose wurden mit murinen J77A.1 Makrophagen durchgeführt. Die 3-fach Serinprotease-Mutanten lagen in signifikant verminderter Anzahl im Vergleich zum Wildtyp-Stamm intrazellulär vor. Weiterhin zeigten Biofilmversuche mit dem Serotyp 19F die Bedeutung der CbpG-Protease in diesem Prozess. Zwei verschiedene Ansätze der Biolfilmtests wurden dafür verwendet. Zum einen die Biofilmbildung auf einer abiotischen Oberfläche (Glas) und zum anderen auf einer biotischen Oberfläche (lebende Epithelzellen). Der Biofilm war nach 48 und 72 Stunden auf lebenden Epithelzellen höher und dichter als auf der Glasoberfläche. Vor allem entwickelte die Serinprotease-Mutante, die nur CbpG+ exprimierte, einen erhöhten und wesentlich dichteren Biofilm als der parentale Wildtypstamm und die anderen Serinprotease-Mutanten. Bei der Mutante, bei der alle Serinproteasen deaktiviert waren, wurden in erheblichem Umfang Bakterien freigesetzt.

Die Nasenrachenraum-Besiedelung ist eine Voraussetzung für invasive Krankheiten und Übertragung von Menschen zu Mensch. Diese Enzyme könnten vielversprechende Kandidaten für die Entwicklung von antimikrobiellen Komponenten sein, die in der Lage sind, Pneumokokken-Besiedelung und -Übertragung zu reduzieren. Serinproteasen von

Pneumokokken sind unabdingbar für eine Nasenrachenraumbesiedelung und beschleunigen dadurch den Zugriff zu eukaryotischen Zelloberflächen-Rezeptoren durch die Spaltung von denselben ECM-Proteinen.

Chapter 2

SUMMARY

2. Summary

Streptococcus pneumoniae (the pneumococcus) is an opportunistic human pathogen that causes life-threatening diseases including pneumonia, sepsis, meningitis but also non-invasive local infections such as otitis media. Pneumococci have evolved versatile strategies to colonize the upper respiratory tract (URT) of humans. Binding to epithelial surfaces is thereby mediated through direct interactions with host cell receptors or indirectly via binding to components of the extracellular matrix (ECM). However, successful colonization and subsequent infection require *S. pneumoniae* to cross tissue barriers protected by the immune system of the host. Pneumococci have therefore evolved a wide range of mechanisms to circumvent the antibacterial activity of the immune system such as the acquisition or expression of serine protease activity. Serine protease enzymes have emerged during evolution as one of the most abundant and functionally diverse groups of proteins in eukaryotic and prokaryotic organisms. However, the epithelial barriers, integrins, and other cell surface receptors are often initially inaccessible for pneumococci colonizing the nasopharyngeal cavity. Therefore, pneumococci recruit host-derived extracellular serine proteases such as plasmin(ogen) for extracellular matrix and mucus degradation, which results in enhanced binding to epithelial and endothelial cells. *S. pneumoniae* expresses four surface-anchored or surface-associated serine proteases depending on the serotype: HtrA, SFP, PrtA, and CbpG. These enzymes belong to the category of trypsin-like or subtilisin-like family proteins, which are characterized by the presence of three-conserved amino acid residues, Ser-His-Asp. The catalytic triads are critical for the cleavage of peptide bonds. Studies focusing on the deletion of single pneumococcal serine proteases are difficult to interpret due to the compensatory effects of the other serine proteases. Initially, a comprehensive *in silico* analysis of the distribution and genes organization of pneumococcal serine proteases was carried out in this study. Interestingly, the genes encoding PrtA, HtrA, and CbpG were highly conserved among the 11 analyzed strains. Surprisingly, the gene encoding the subtilisin-like protein SFP was not present in some of the strains and seems to be strain-dependent. Therefore, pneumococci have at least three serine proteases as shown e.g., for serotype 19F_EF3030 strain. Computer-assisted analyses of the structure of pneumococcal serine proteases showed high similarities in the catalytic domains between HtrA and CbpG or between PrtA and SFP in 3D structural models.

The focus of this study lies on the impact of single extracellular pneumococcal serine proteases on pneumococcal pathogenesis during adherence, colonization, virulence and biofilm formation. Therefore, double and triple deletion mutants were generated in the colonizing

S. pneumoniae serotype 19F strain EF3030 and the more invasive serotype 4 strain TIGR4, respectively. In adherence studies with human Detroit-562 epithelial cells, we demonstrated that both TIGR4Δ*cps* and 19F_EF3030 mutants without serine proteases or expressing only CbpG, HtrA, or PrtA have a reduced ability to adhere to Detroit-562 cells. In a mouse colonization model, the inactivation of serine proteases in strain 19F_EF3030 strongly reduced nasopharyngeal colonization in CD-1 mice. The bacterial load in the nasopharynx was thereby monitored for a period of 14 days. Mutant strains showed significantly lower bacterial numbers in the nasopharynx on days 2, 3, 7, and 14 post-inoculations.

Following up on pneumococcal pathogenesis, an *in vivo* acute pneumonia mouse infection model and *in vitro* phagocytosis was used to analyze the impact of single serine proteases during infection and phagocytosis. Mice were intranasally infected with the bioluminescent TIGR4*lux* wild-type or isogenic triple mutants expressing only CbpG, HtrA, PrtA, or SFP. The acute lung infection was monitored in real-time by using an IVIS®-Spectrum *in vivo* imaging system. The TIGR4*lux* mutant expressing only PrtA showed a significant attenuation and was less virulent in the acute pneumonia model. Phagocytosis assays were conducted using murine J77A.1 macrophages. The number of triple *serine protease* mutants internalized by macrophages were significantly reduced in comparison to the isogenic wild-type.

Finally, two different experimental biofilm models were used to study the influence of serine proteases on biofilm formation grown on an abiotic surface (glass) and a biological surface. Biofilm development on living epithelial cells was stronger after 48 and 72h than on the glass surface. On epithelial substratum, the serine protease mutant with only CbpG+ showed higher and denser biofilm development after 48h and 72h of incubation compared to the parental strains and other *serine protease* mutants. Moreover, the bacterial dispersal from biofilms was significantly more in the mutant strains lacking serine proteases than in the wild type.

In conclusion, nasopharyngeal colonization is a prerequisite for invasive diseases and transmission. Pneumococcal serine proteases are indispensable for nasopharyngeal colonization and facilitate access to eukaryotic cell-surface receptors by the cleavage of ECM proteins. Thus, serine proteases could be promising candidates for developing antimicrobials to reduce pneumococcal colonization and transmission.

INTRODUCTION

3.   Introduction

3.1 *Streptococcus pneumoniae*

Streptococcus pneumoniae (also called pneumococcus) was isolated for the first time independently in 1881 by two scientists, US Army surgeon George Sternberg (Sternberg, 1881) and the french scientist Louis Pasteur (Pasteur, 1881). The pneumococcus was classified as follows:

Kingdom: Bacteria (Cavalier-Smith, 2002).
  Phylum: Firmicutes (Gibbons and Murray, 1978).
    Class: Bacilli (Ludwig et al., 2010).
      Order: Lactobacillales (Ludwig et al., 2010).
        Family: Streptococcaceae (Deibel and Seeley, 1974).
          Genus: *Streptococcus* (Rosenbach, 1884).
            Species: *S. pneumoniae* (Klein, 1884).

3.1.1 General characteristics

S. pneumoniae is an opportunistic, well-characterized human commensal and pathogenic bacterium. Pneumococcus is a Gram-positive bacterium that can grow successfully in microaerobic and anaerobic environments. Pneumococcus grows in pairs (diplococci) or short chains with a lanceolate morphology, and a diameter of 1.0-1.5μm. The colonies appear in small, dome-shaped, and shining colonies after 24h of incubation on blood agar plates. Interestingly, pneumococcus produces prodigious amounts of hydrogen peroxide (H_2O_2) as a competitive mechanism against other bacteria in their biological niche[1, 2]. Pneumococci can be divided into more than 98 serotypes based on the polysaccharide composition of the capsular polysaccharide (capsule; CPS)[3-6]. The capsule is one of the most important virulence factors, protecting pneumococci from host immune phagocytic cells[7-10]. The autolytic enzyme LytA is responsible for the lysis of pneumococci by degradation of the cell wall.

S. pneumoniae regularly colonizes asymptomatically epithelia of the upper respiratory tract (URT), and humans colonized with pneumococci are known as carriers[11]. Carriers serves as the main reservoir for bacterial transmission within the population via person-to-person contact in the form of airborne droplets[11]. Adherence to a mucosal surface, predominantly to components of the extracellular matrix (ECM), is a prerequisite for establishing stable colonization of the host[12]. However, pneumococci can disseminate from the nasopharynx to deeper tissues and the blood under certain circumstances, leading to invasive diseases such as pneumonia, meningitis, otitis media or sepsis[4, 13, 14].

3.1.2 The burden of pneumococcal diseases

S. pneumoniae can cause a wide range of life-threatening diseases also known as invasive pneumococcal diseases (IPD) such as pneumonia, bacteremia, and meningitis. In addition, pneumococci can also cause community-acquired pneumonia (CAP), which is mostly acquired outside the hospital and frequently circulating in the community. CAP is a major respiratory health disease with a high incidence and variable severity in the general population[15, 16]. Epidemiological studies estimate that 27-65% of children and around 10% of adults are pneumococcal carriers[17, 18]. The more vulnerable populations, which are more affected by IPD are the children, the elderly, and immunocompromised persons due to their undeveloped or weakened immune system, respectively[19-21]. In particular, children under the age of <2-5 and adults over 65 years [22]. In the World Health Organization (WHO) report from 2017, about antibiotic-resistant "priority pathogens", *S. pneumoniae* was considered as one of the 12 bacteria causing the greatest threat to human health[23, 24]. According to the WHO pneumonia and meningitis are the most important infectious diseases[25, 26]. Moreover, every year, around 300 million cases of acute otitis media (AOM) caused by *S. pneumoniae* are recorded worldwide[27]. The highest mortality was estimated in children under the age of five years with about 700,000 to 1 million dead children every year worldwide[28]. Thus, pneumococci were also called "The Forgotten Killer of Children" as mentioned by UNICEF and WHO[29, 30]. One of the reasons therefore could be that pneumococcus is more likely to be acquired from the environment by children, especially in day-care facilities where children closely interact with each other. In contrast, adults are more frequently infected with *S. pneumoniae* from their households, with stronger acquisition tendencies from younger children, who are more often colonized with pneumococci[31].

Pneumococcal colonization of the nasopharynx commonly occurs in healthy individuals without causing diseases. In childhood, 20 to 60% are colonized with *S. pneumoniae*, whereas only 2 to 20% of the adults are colonized. The colonization rates can reach up to 90% in developing countries[32-37]. However, acute pneumonia is the most common disease caused by pneumococci in the adults group. In fact, pneumonia correlates with high morbidity and mortality by bacteremia within 10-30% of the cases in the group of adults between 19 to 64 years old[38, 39].

The disease burden remains remarkable worldwide in both, developed and developing countries despite the advances or discovery of antimicrobial treatments, vaccinations, and antibiotic treatments [40]. On one hand, antibiotic treatment was discovered and used to treat pneumococcal

infections of otherwise deadly pneumococcal diseases[41]. On the other hand, antibiotic treatment had only a minor effect on the pneumococcal carriage, which is still prevalent in both children and adults worldwide[42-44].

The treatment of pneumococcal diseases includes antibiotics such as penicillin, erythromycin, tetracycline, and chloramphenicol. In 1998, analysis of IPD isolates in the USA revealed that 24% of the isolates carried resistance to penicillin and 14% were resistant to multiple antibiotics[45]. In 2014, the WHO global report on antimicrobial resistance revealed that the number of pneumococcal isolates resistant to penicillin has increased by 50% in the western pacific, America, and Europe[46]. In recent years, the frequency of penicillin and erythromycin-resistance to pneumococci has risen[47]. The incidence of pneumococcal infections is affected by various factors such as immunological state, age, and geographical region[38]. In 2000, the pneumococcal disease burden was estimated at 14.5 million in children worldwide, with 735,000 deaths per year[19]. In 2015, pneumococcal mortality decreased to 51% among children under five years and was estimated to be 45/100,000 population globally[28].

Further breakdown of this data depicted 15.8/100,000 cases of IPD in adults aged 65 years and 11.9/100,000 in children under the age of five[47]. Up until 2015, pneumonia was responsible for around 16% of all deaths in children under the age of five worldwide. Surprisingly, death rates among the elderly generally do not seem to change[48]. The ECDC reported that the pneumococcal mortality and morbidity remained high between 2012 to 2016 in the EU/EEA in adults aged 65 years and children under one year[47]. Vaccine introduction has significantly reduced IPD caused by serotypes covered in the vaccines[49, 50]. However, the emergence of non-vaccine serotype (NVT) in carriage and IPD, a phenomenon known as serotype replacement has also been raised with the use of the vaccination[51, 52]. Therefore, there is a need to improve vaccines and develop new antimicrobial therapies to reduce the pneumococcal burden of disease. The WHO in 2007 highly recommended and encouraged countries to implement surveillance of IPD by including pneumococcal conjugate vaccines (PCV) into infant immunization programs[30]. Therefore, in the subsequent years, the burden of pneumococcal disease is expected to drop dramatically due to the access to 10- and 13-valent pneumococcal conjugate vaccines (PCV10/13)[53].

3.1.3 Pneumococcal core genome and genome flexibility

The first whole-genome sequencing of *S. pneumoniae* strains TIGR4 (serotype 4) and R6 (serotype 2) were reported in 2001. The two genomes contain about 2-2.2 million base pairs (Mbp) and include 2.100-2.200 coding regions[54, 55]. The majority of serotypes were responsible for pneumococcal invasive diseases such as TIGR4, 6B, 9V, 14, 18C, 19F, and 23F[56].

The *S. pneumoniae* genome highly varies between the strains due to the natural genome flexibility. This phenomenon enables the pneumococci to uptake the single-stranded DNA from their environment and incorporate it for example via homologous recombination into its genome. The natural transformation was observed for the first time by Frederick Griffith in 1928, related to the phenomenon of capsular switching[57]. Later on, Avery et al. in 1944 demonstrated that the transforming material was different from proteins and suggests DNA was the transforming agent[58]. The pneumococcal core genome is comparable in all strains and includes around two-thirds of the total gene content, whereas genes of the accessory regions encode the remaining third. The core genome encodes genes for metabolism, virulence and regulation[59]. The accessory genes include the capsule locus and some virulence factors such as PsrP or pili[60 61].

The nasopharynx of humans is believed to be the primary site of genetic material exchange by horizontal gene transfer (HGT) between pneumococcal strains and *Streptococcus mitis*, therefore, for the acquisition of virulence factors[62, 63]. The genome of *S. pneumoniae* contains a wide range of genes encoding for virulence factors, with more than 300 genes related to virulence[64]. Some genes have a direct role in virulence, whereas others are generally necessary for bacterial fitness[64]. Bacterial pathogenicity correlates in general with the number of virulence factors. Elimination of these factors from the bacteria leads to diminished virulence. Pneumococci exploit its armament of virulence factors to allow the bacterium to colonize the host, interact with microbiota and host defenses and invade into the upper and lower respiratory tract[65]. The most important known virulence factors which actively participate in pneumococcal pathogenesis include the capsule, the cell wall, choline-binding proteins, sortase-anchored proteins (with an LPXTG cleavage motif), lipoproteins, pneumococcal adhesins and virulence factor A/B (PavA/PavB), enolase, pneumolysin and the major autolysin A (LytA)[65].

3.2 Pneumococcal virulence factors

3.2.1 The pneumococcal capsule

The pneumococcal capsular polysaccharide layer of repeating oligosaccharides surrounds the pneumococcus[66]. It is considered as one of the primary virulent determinants with a thickness of 200-400 nm that protects the pneumococcus from the human immune system[67]. Most of the *S. pneumoniae* clinical isolates are encapsulated and differ clearly from the nonencapsulated pneumococci by their colony morphology. Encapsulated pneumococci form smooth colonies, while nonencapsulated have a rough colony morphology. Due to the capsule thickness nature, the CPS can easily prevent the antibiotics from penetration, thus reducing the killing of pneumococci[68]. One of the most important functions of CPS is the prevention of opsonisation by complement factor C3b/iC3b and specific IgGs. This prevents the pneumococci from being phagocytosed by immune cells[7, 69]. Otherwise, these processes would result in the pneumococcal killing from the host by phagocytic immune cells[7]. Importantly, the pneumococcus can also resist mucus trapping in the upper respiratory system during nasopharyngeal colonization because of the negative charge of the capsule[66, 68, 70, 71]. Additionally, pneumococci are classified into more than 98 serotypes summarized in 48 serogroups sharing serologic characteristics such as cross-reactivity to antibodies, based on their capsular architecture and serologic properties[5, 72]. Of concern, capsular switching occurs frequently in pneumococcal populations, which amongst others, is due to the introduction of vaccines and the widespread use of antibiotics treatment[73].

Capsule biosynthesis is complex and regulated by the *cps* locus (*cpsA-D*) genes, which encode specific enzymes involved in the uptake and synthesis of monosaccharide components. Except for serotypes 3 and 37, which have a synthase-dependent mechanism, most pneumococcal serotypes are synthesized via the flippase (Wzy) pathway[6, 74].

The CPS biosynthesis places a considerable metabolic burden on pneumococci, and therefore needs to be strictly regulated. Although CPS enables pneumococci to overcome the mucosal barrier, it simultaneously inhibits their attachment to the epithelium via repulsion and masking of surface-localized adhesins, which are required for stable colonization. Therefore, pneumococci undergo phase variation of CPS expression by the development of opaque and transparent phenotypes[75]. The opaque phenotype is more virulent and more associated with systemic infections due to a high expression of capsular polysaccharides. The transparent variants in contrast express less CPS than opaque variations and are therefore more efficient in

mucosal surface colonization[76]. Previous studies have shown that the CPS thickness of serotype 3 strain was decreased during host cell invasion and further reduced during colonization in both *in vivo* and *in vitro*[77]. Furthermore, the capsule gene expression has been shown to be downregulated during colonization and biofilm formation[78, 79]. Therefore, capsule reduction may improve adhesion and biofilm development during colonization[12].

3.2.2   Pneumococcal cell wall

Like other Gram-positive bacteria, pneumococci are surrounded by a thick multi-layered cell wall that underlies the pneumococcal polysaccharide capsule. The cell wall of *S. pneumoniae* is complex and plays a crucial role in cell shape maintenance, growth, cell division, and interactions with human host components[80]. Peptidoglycan (PGN) and teichoic acids (TA) are the main components of the cell wall[81, 82]. The wall teichoic acids (WTA) are covalently linked to the peptidoglycan, while lipoteichoic acids (LTA) are anchored in the cytoplasmic membrane[80]. Macromolecular peptidoglycan is composed of alternating glycan chains composed of *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc)[83]. A pentapeptide link exists between the propionic acid of MurNAc. Pneumococcal WTA and LTA are structurally identical and consist of a repeated motif of repeating units possessing phosphorylcholine (*P*Cho) residues. These phosphorylcholine residues promote attachment of choline-binding proteins (CBPs) to the pneumococcal cell surface[84] (**Figure 3.4**).

3.2.3 Pneumococcal surface proteins

Several surface-exposed proteins of pneumococci are considered as important virulence factors because of their direct involvement in host-pathogen interactions. These proteins are involved in different stages of pneumococcal infection, including adhesion, colonization, biofilm formation, invasion, and immune evasion, allowing pneumococci to systemically spread into sterile host tissues and organs[65, 85]. In general, pneumococcal surface proteins can be divided based on the anchoring mechanism. *S. pneumoniae* expresses four prominent families of surface-exposed proteins **i)** choline-binding proteins (CBPs), **ii)** pneumococcal lipoproteins-attached to the bacterial cell membrane, **iii)** sortase-anchored proteins (LPXTG cleavage motif) anchored covalently to the peptidoglycan, and **iv)** non-classical surface proteins (NCSP) without a leader peptide or membrane-anchoring motif[65, 86-89]. The following section provides a brief overview of each family, focusing on pneumococcal adherence and virulence factors, with a focus on surface-exposed pneumococcal serine proteases and their interaction with the host.

3.2.3.1 Choline-binding proteins (CBPs)

S. pneumoniae expresses a special class of surface proteins known as choline-binding proteins. Choline is one of the essential nutrients that pneumococcus needs to grow and survive. *S. pneumoniae* imports the choline from the growth medium or environment, which can be incorporated into the repeating units of WTA and LTA[90, 91]. Pneumococcal choline-binding proteins are a family of polypeptides found in all pneumococcal strains[92]. The common feature is the modular organization with at least two domains: a functional module (FM) and a choline-binding module (CBM). CBPs can only be found in pneumococci or closely related species[38, 93-95]. The repetitive sequences of CBMs attach CBPs in a non-covalent manner to the cell wall by their interaction with phosphoryl-choline residues of PGN-anchored WTA and membrane-anchored LTA[86, 92]. The CBMs, generally located at the C-terminal part, consist of three to eighteen characteristic repetitive sequences of about 20 amino acids, the so-called choline-binding repeats (CBRs)[85, 96, 97]. Exception are the CBPs LytB and LytC, where the CBMs are located at the N-terminal part of the protein, whereas the FM is located in the C-terminal region[86].

CBPs play an essential role in the physiology of the cell wall, colonization processes, interaction with host cells, and biofilm formations[92, 98-101]. Furthermore, CBP plays a vital role by modifying the pneumococcal cell surface, inhibiting the binding of specific immune receptors secreted in response to pneumococcal infection[92, 102].

A comprehensive analysis of pneumococcal choline-binding proteins from different studies was extensively reviewed by Maestro and Sanz in 2016[85, 92]. Among the well-studied CBPs, are the pneumococcal surface protein A (PspA), and pneumococcal surface protein C (PspC, also known as CbpA). Four pneumococcal CBPs were identified as hydrolytic enzymes, including the N-acetylmuramoyl-L-alanine amidase (LytA), LytB, LytC, and choline-binding protein E (CbpE, also known as Pce). Further members are CbpF, CbpL, CbpD, and the serine protease CbpG. It is noteworthy to highlight that the deletion of CbpD, CbpE, CbpG, LytB, and LytC leads to a dramatically decreased adherence to eukaryotic cells and nasopharyngeal colonization[98, 100, 103, 104]. More importantly, the serine protease of interest CbpG has a multi-functional role. The enzymatic activity and adherence function of CbpG could play an essential role in mucosal colonization and sepsis[100, 105, 106].

3.2.3.2 Pneumococcal lipoproteins

Lipoproteins are another class of surface proteins containing a lipobox motif with a conserved cysteine (LX$_1$X$_2$C). The presence of this motif enables covalent anchoring of the protein to the cytoplasmic membrane[107, 108]. The anchoring of lipoproteins to the cell membrane generally requires two different enzymes[109]. First, the pre-lipoprotein is lipidated in an invariant cysteine residue by the lipoprotein diacylglycerol transferase (Lgt). The prolipoprotein is then covalently attached to the cell membrane through diacylglycerol moieties. The signal peptide is cleaved by the lipoprotein signal peptidase II (Lsp)[109]. Mature lipoproteins are involved in many functions, including bacterial fitness, substrate uptake, ABC transporters and enzyme activity[107, 108]. As an example, the deletion of Lgt leads to a reduction in surface anchored lipoproteins which is accompanied by a declined nasopharyngeal colonization in a murine colonization model[110]. Moreover, the deletion of *lgt* severely impaired pneumococcal virulence in mice models of pneumonia and septicemia[110, 111].

3.2.3.3 Sortase A-anchored proteins (LPxTG proteins)

Sortase A-anchored proteins are one of the most common types of cell surface anchored proteins in pneumococci. These proteins are covalently anchored to the PGN of the cell wall through their LPXTG motif. Sortase A, a transpeptidase, a transpeptidase, recognizes the LPXTG motif and cleaves between the threonine and glycine residues and covalently anchors the protein to the peptidoglycan of the cell wall[88, 112, 113]. Interestingly, pneumococcal sortase-anchored proteins have a wide range of functions in pneumococcal colonization, virulence, and proteolytic activities. Hydrolases and other lytic enzymes such as zinc-metalloprotease, StrH, NanA, and PrtA are part of pneumococcal sortase-anchored proteins[54, 55, 114-116]. These proteins

are well studied, like the pneumococcal adhesion and virulence factor B (PavB, also known as PfbB), the pneumococcal collagen-like protein A (PclA), and plasmin- and fibronectin-binding protein A (PfbA), which are implicated in pneumococcal adhesion and colonization [117, 118]. Other examples are the zinc-metalloprotease A, B, C (ZmpA, ZmpB, ZmpC), and IgA1 protease[54, 55, 117]. Importantly, the substrates of these proteases include various glycoproteins that have essential impacts on pneumococcal invasion. The proteins collectively degrade specific ECMs components thereby facilitating the adhesion of pneumococci to nasal and lunge compartments[119, 120]. As a striking examples, the deletion of *srtA* showed that LPXTG anchored proteins play a pivotal role in bacterial adhesion to host epithelial cells and the interaction with human ECM[121, 122]. In this experiment, inactivation of the sortase gene *srtA* showed that pneumococci cannot adhere to nasopharyngeal cells and the release of neuraminidase was observed. Neuraminidases are involved in biofilms formation and AOM[123]. In summary, these proteins seem to facilitate the adhesion of pneumococci to host epithelial cells and the degradation of ECM glycoproteins to gain access to sterile host compartments during infection[119].

3.2.3.4 Non-classical surface proteins (NCSPs)

Non-classical surface proteins, also known as moonlighting proteins, are another class of pneumococcal proteins that lack a typical signal peptide or anchoring motif [88]. Despite this, NCSPs are found on the pneumococcal surface, whereas the process of secretion and anchoring is unknown[85, 88]. The majority of the proteins in this category have been identified as intracellular metabolic enzymes such as the enolase protein[88], the glycerate-3-phosphate dehydrogenase (GAPDH),[54, 55] the 6-phosphogluconate dehydrogenases (6-PGD)[124, 125] and the pneumococcal adherence and virulence factor A (PavA)[85, 126]. In general, NCSPs play an important role as adhesins by interacting with human ECM proteins, facilitating the adhesion and transmigration of the pneumococci through the epithelium. One well-studied example for NCSPs is PavA, which binds to human fibronectin and promotes attachment to host epithelial cells. Also enolase and GAPDH were shown to bind human plasmin(ogen), facilitating pneumococcal invasion[127].

3.3 Host-pathogen interactions

The ability of a microorganism to cause disease, known as "pathogenicity", is mainly determined by its virulence factors. Pneumococci are endowed with a plethora of virulence factors contributing to pneumococcal adhesion, colonization, invasion, immune evasion and host cell damage[13, 65, 89, 128]. The initial steps of pneumococcal pathogenesis require an intimate and specific adherence to host structures and modulation or evasion of innate immune clearance mechanisms[13].

Pneumococci can interact with the host either directly via surface components engaging host cell receptors on respiratory epithelial cells or indirectly by using proteins of the ECM as bridging molecules. These interactions are also required for pneumococcal biofilm formation, which is often accompanied by increased antibiotic resistance. The contribution of pneumococcal surface proteins to colonization and dissemination to the lower respiratory tract is still a crucial issue to understand. Pneumococcus has several surface-associated factors that allow both a commensal and a pathogenic lifestyle[129]. These factors enable the pneumococcus to survive during colonization and facilitate invasion of host tissues and the blood. The detailed survival advantages gained by pneumococci from these factors are going to be thoroughly addressed in the following section.

3.3.1 Pneumococcal adherence, colonization, and disease transmission

The URTs are an ecological niche for many different bacterial species. Among them are also pathogenic species, including *S. pneumoniae*, *Haemophilus influenzae*, and *Moraxella catarrhalis*[130, 131]. Every person is likely to be colonized by one of these microorganisms at least once in a lifetime. In most cases, the carriage is asymptomatic; nevertheless, colonization is a prerequisite for invasive diseases[132, 133]. In the case of pneumococci, the first step towards invasive disease is nasopharyngeal colonization, which is the main reservoir of the bacteria. Many factors are required to colonize and remain on the nasopharyngeal mucosal surface at a density and duration sufficient for pneumococcal transmission to new susceptible hosts [13]. As an example, pneumococci, produce different enzymes, such as the attenuator of drug resistance (Adr), and *N*-acetylglucosamine deacetylase (PgdA), which hydrolyzes the PGN and render it resistant to the lytic effects of lysozyme, which is prevalent on the mucosal surface of the URT[134].

The initial step of pneumococcal colonization requires tight adherence to the host epithelial tissues of the URT. Subsequently, pneumococcus needs to overcome mucosal innate and

adaptive immunity, evade mucociliary flow clearance, and has to tolerate the limited nutrient availability. Pneumococcal binding to cell-surface carbohydrates (*N*-acetyl-glucosamine) in non-inflamed resting epithelium results, amongst others in asymptomatic colonization. Importantly, carcinoembryonic antigen-related cell adhesion molecule and intercellular adhesion molecule 1 (ICAM1) are two types of cell adhesion receptors exploited by pathogenic bacteria[135]. With regards to this, *S. pneumoniae* increases the expression of epithelial surface receptors, thereby facilitating the adhesion[136]. Furthermore, phosphorylcholine of the TAs (PCho) binds to the platelet-activating factor receptor (PAFR), while PspC (known also as CbpA: choline-binding protein A) binds to the secretory component of polymeric immunoglobulin receptor R (pIgR)[13, 137]. Internalization of pneumococci is induced by binding to these receptors, which facilitates pneumococcal transcellular migration through the respiratory epithelium and vascular endothelium, resulting in an invasion of living bacteria[13, 138].

Mucus entrapment on URT epithelia is the first line of host defense against pneumococci in the nasopharyngeal cavity. Pneumococci have to avoid entrapment in the mucus layer and subsequent clearance by the host immune system[13, 88]. There are many antimicrobial peptides and immunoglobulins, which are present in the glycocalyx that covers the URT epithelium[7]. However, the mucus layer keeps bacteria away from the underlying cell surface. Pneumococcal adherence to the mucin glycans enables the bacteria to persist on the mucosal surface and provides favourable nutrients[13]. Apart from avoiding mucus entrapment, pneumococci overcome the harsh environmental conditions by expressing a polysaccharide capsule. CPSs resist sialic acid-rich mucopolysaccharides in the mucus due to the negative charge of the capsule[75]. There are almost six serotypes which have a neutral charge (7F, 7A, 15, 33F, 37, 1)[5]. This ability enables repulsion from the anionic mucous barrier, thereby preventing trapping and mucus-mediated host clearance. However, pneumococci need to downregulate capsule expression to maintain stable adherence to host cells. In contrast, the thick capsule facilitates pneumococci to pass through mucus. The capsule also prevents adherence to the epithelium and prevents colonization of the host. Consequently, pneumococci elevate phase variation, a process that allows the bacteria to switch the capsule expression on or off. The bacterial colony phenotype switches from the "opaque" (more capsule) to the "transparent" (less capsule) phenotype, and vice versa[75, 77].

The initial attachment of pneumococci by surface-associated glycosyl hydrolases such as neuraminidase (NanA), and β-galactosidase BgaA are rather weak. In this context, pneumococcal surface-exposed proteins mediating adherence to epithelial cells or the ECM are

unmasked by capsule reduction[77]. Pneumococcal adhesins bind directly to host cell surface proteins or indirectly via ECM glycoproteins like fibronectin, fibrinogen, vitronectin, thrombospondin-1, and collagen[65, 89, 139-141]. *In vitro* cell culture models have provided considerable knowledge to understand the complex host-pneumococcus interactions. In this sense, pneumococci use various strategies to interact with the epithelial cell surface and avoid entrapment by mucus. The pneumococcal surface-exposed adhesion proteins are highly abundant on the bacterial surface. For example, the pneumococcal surface protein C (PspC, also known as CbpA), PavB, or PsrP, which were shown to interact directly with host cell receptors[76, 87, 142-145]. PspC binds to the secretory component of the polymeric immunoglobulin receptor (hpigR), which increases migration through the mucosal barrier[146]. In addition, PspC also binds to the host complement protein factor H and vitronectin, this interaction promotes adherence to host cells[4, 147]. Moreover, the close interaction of pneumococci with the nasopharyngeal host cells is initially prevented by mucus and ciliary beating of the microvilli on the apical pole of mucosal epithelial cells[148]. Pneumolysin inhibits ciliary beating[149], and enzymes like the pneumococcal NanA and hyaluronidase (Hyl) contribute to receptor exposure on the surface of host cells[13].

More importantly, the interaction between bacterial proteins known as microbial surface components recognizing adhesive matrix molecules such as the enolase, or pneumococcal adherence and virulence factor A and B (PavA, PavB) and extracellular matrix (ECM) components like fibronectin, vitronectin, thrombospondin-1, and plasminogen promotes binding to host cells [89, 111, 140, 143].

Pneumococci exhibit the ability to hijack host-derived serine protease proteolytic activities by binding of plasmin(ogen), to break down the mucosal surface. Hence degrade ECM components facilitate the tight interaction with host cells during colonization and dissemination[13, 150, 151]. Weiser and co-workers have shown the function of the pneumococcal IgA1-protease. They revealed that in the presence of secretory IgA (sIgA), which can be found on mucosal surfaces, pneumococci showed increased adherence to lung epithelial cells. It has been hypothesized that the cleavage of opsonizing IgA leads to a change in surface charge and therefore to increased physical contact of pneumococcal cell-wall choline to the platelet-activating factor receptor[152, 153].

Indeed, pneumococcal surface-exposed/or surface-associated proteases have been shown to play an important role during adherence and colonization. To name a few, CbpG, HtrA, CbpD, Caseinolytic Protease P (ClpP), C3-degrading Protease CppA, and aminopeptidase N (PepN). In support of this notion, PepN has been shown to reduce *in vitro* adherence to host epithelial

cells (A549) and *in vivo* colonization and thus is involved in pneumococcal transmission[154]. HtrA is a surface-exposed serine protease and has a strong influence on pneumococcal adherence and colonization[155]. A CbpG deficient mutant revealed a significant reduction in rat and mice colonization models and also reduced adherence to human epithelial cells[106]. Therefore, serine protease HtrA and CbpG could be promising targets for the prevention of colonization[100, 155].

3.3.2 Biofilm formation

3.3.2.1 Background

Bacterial biofilms are highly complex surface-attached bacterial communities embedded in self-produced polymer matrixes. These cellular communities are primarily comprised of extracellular polymeric substances (EPS) such as polysaccharides, proteins, lipids, and nucleic acids (extracellular eDNA) with a high percentage of water[156, 157]. Bacteria have different mechanisms to form biofilms depending on their environment and specific strain. According to the National Institutes of Health (NIH), more than 80% of bacterial infections are associated with biofilm formation[158]. Bacterial biofilms are naturally resistant to antimicrobial agents, and the protective biofilm matrix permits bacteria to evade the host immune system[79, 159]. Quorum sensing as cell-cell communication is also important during biofilm formation, which includes exchanging molecules and genetic information between bacteria[160].

Bacteria in biofilms can employ several survival strategies to evade the host. They can survive on the tissues of patients causing persistent infections and dissemination of bacteria to sterile tissues and organs[161]. The vast majority of pathogenic bacteria in the human body persists within well-structured biofilms, associated with many infections. These include chronic rhinosinusitis (*Staphylococcus aureus*)[162], recurrent urinary tract infection (*E. coli*), cystic fibrosis (*Pseudomonas aeruginosa*), and AOM (*S. pneumoniae*)[163]. Dental plaque is a notable example of a natural dental biofilm composed of different species of bacteria, particularly streptococci (*Streptococcus mitis* group), which causes dental caries[164]. However, the role of pneumococcal biofilms in diseases is not yet clear.

3.3.2.2   The implication of biofilm in pneumococcal pathogenesis

In streptococci, biofilm formation is a typical phenotype in both, commensal and pathogenic species. Commensal streptococcal biofilms represent their natural life, while pathogenic biofilms are considered important determinants of infectious diseases[164].

The development of bacterial biofilms is a multi-step process which is summarized in **Figure 3.1**. Biofilm formation is common in all pathogenic streptococcal species. The planktonic bacterium (free-living bacteria) transforms into an immotile lifestyle (sessile bacteria) during biofilm formation, resulting in a complex well-structured biofilm[165, 166]. For this, the initial step is the primary adherence of planktonic bacteria to biotic or abiotic surfaces. The second step of biofilm formation is bacterial aggregation into multicellular structures. A three-dimensional structure develops through the maturation of microcolonies and successive production of the EPS matrix. Bacterial biofilm formation limits the diffusion of nutrients and waste products in the core of the biofilm[167]. For *S. pneumoniae*, bacteria survive host desiccation for many days, and bacteria in the biofilm remain more viable than planktonic bacteria[168, 169]. The disruption of the biofilm matrix, on the other hand, can cause bacterial release from the biofilm (known as biofilm dispersal). Single pneumococci can then disseminate from the biofilms into sterile deeper tissues, causing invasive infections, such as middle ear infection, pneumonia and sepsis[79, 170-172].
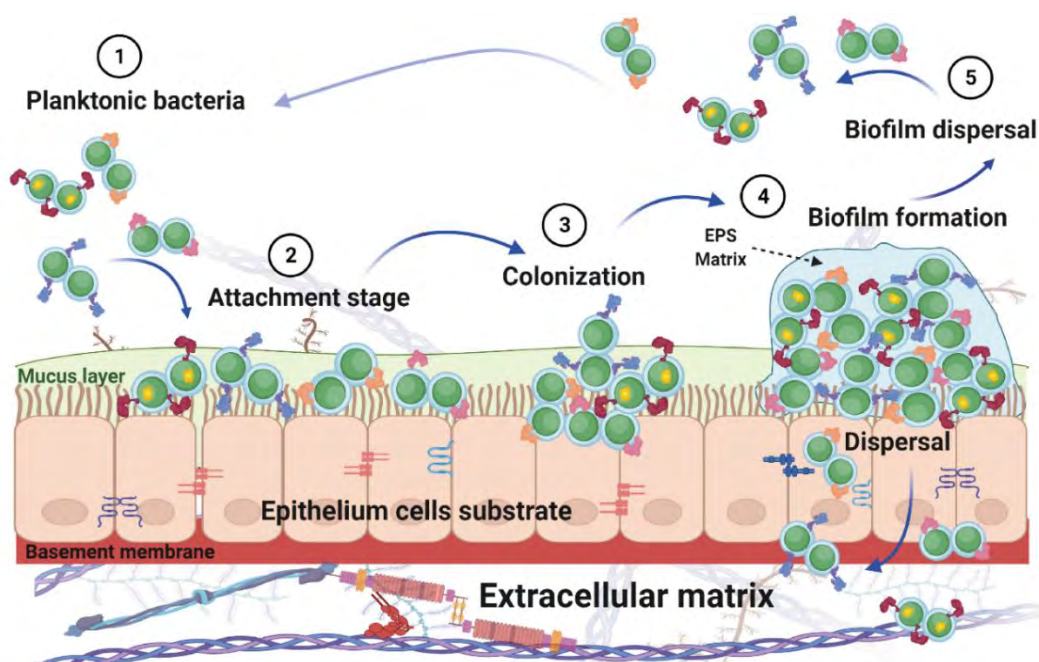
Figure 3.1: Schematic diagram presents the biofilm formation life cycle.

The diagram shows the development of a biofilm, which is a multistep process that can be divided into several steps. (**1**) Planktonic bacteria, which are free-living bacteria floating that consider the main reservoir of a bacterial biofilm. (**2**) The initial adherence of planktonic

bacteria to epithelial cells leads to an irreversible attachment along the surface and the formation of a single layer. (**3**) Microcolony formation pneumococci colonize the epithelial cells and grow upward. (**4**) Biofilm maturation with the production of EPS matrix by resident bacterial cells. (**5**) The biofilm matrix can be enzymatically degraded and pneumococci can be dispersed from the mature biofilm as dispersed bacteria can serve for invasion. Figure created with **BioRender**.com.

S. pneumoniae biofilm formation has been described particularly during otitis media[3, 173, 174] but is also suspected to occur during nasopharyngeal colonization. A well-structured biofilm was observed in a murine colonization model studied by Marks and co-workers[12]. The experiments showed that the 19F_EF3030 strain was able to form biofilms on epithelial cells *in vitro* and this directly correlates with the colonization of the murine nasopharynx. More importantly, several studies on *in vivo* pneumococcal biofilm formation have shown that bacterial clusters form during chronic rhinosinusitis and pneumonia[79, 175, 176]. Thus, biofilm formation seems to play an important role in the pathogenesis of pneumococcal diseases[177-181]. More recent data indicate that bacteria from biofilms can be dispersed to sterile sites. Therefore, the biofilm serves as a reservoir from which virulent bacteria may seed off under the right conditions[79, 182]. Pneumococci use biofilm formation during colonization for several purposes. For example, biofilm protects pneumococci from the host immune system by forming sessile colonies embedded in an extracellular matrix composed of different EPS[183]. The EPS most likely consist of different components such as carbohydrates. The integrity of the EPS depends on eDNA and probably on proteins that are incorporated in the matrix, because protease treatment was shown to alter the EPS[184, 185]. The EPS matrix also contributes to antimicrobial resistance and allows the accumulation of nutrients from the host environment as well as the release of post-death cellular substances and bacterial end products[186, 187]. Furthermore, quorum sensing during biofilm allows communication between the bacteria due to sensing of molecules and eDNA that makes up part of the extracellular matrix. Such communication can provide a perfect environment for genetic transformation to enhance bacterial survival[188-190].

Scanning electron microscopy (SEM) images (**Figure 3.2** [[12, 183]]) of mature pneumococcal biofilms grown on epithelial cell substratum showed a high growth rate, as more biomass and density of pneumococcal biofilm can be observed in comparison to biofilms grown on an abiotic surface (glass substratum). This fact indicates that interactions with epithelial cells influence biofilm development. In addition, biofilm development is also influenced by different environmental factors such as temperature and nutrient availability. The physiological

temperature of 34°C in the nasopharynx and chemically-defined media results in a more dense biofilms[12, 191].
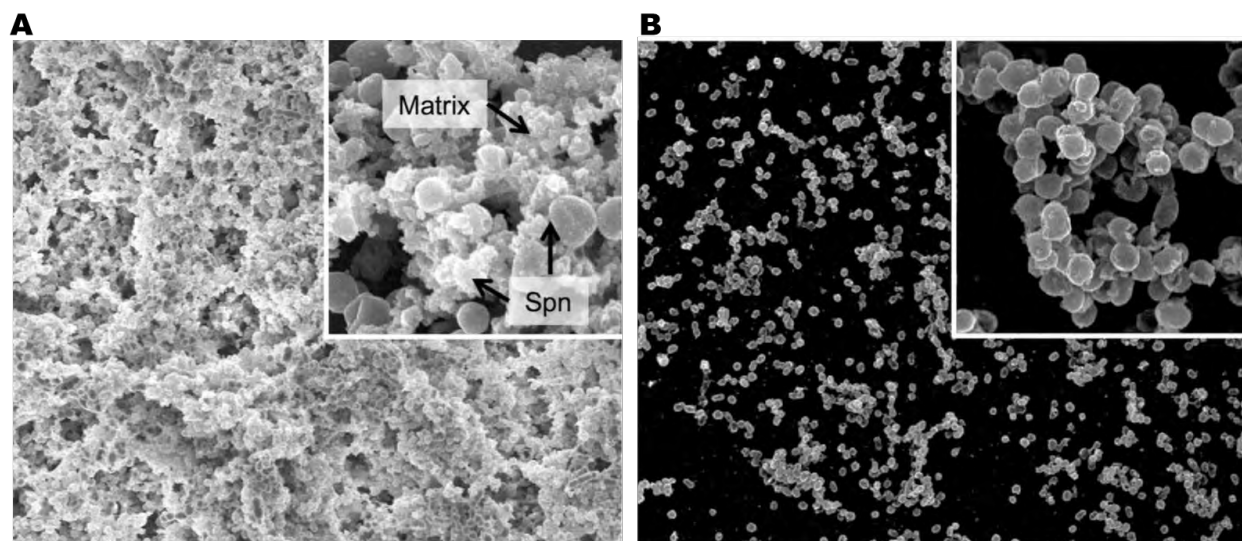
Figure 3.2: Examples of mature biofilm formation by *S. pneumoniae*.

Scanning electron microscopy of pneumococcal biofilm formation on epithelial cells (**A**) or an abiotic surface (**B**). The image shows the aggregation and organization of bacteria (diplococci) in clusters with empty spaces and the formation of an extracellular matrix encasing the bacteria. The matrix will end up holding the biofilm together and protecting it from the harsh environment[183].

Pneumococci in biofilm constitute a different community, with many bacteria in a more sessile state referred to as the "persister" phenotype[159]. The sessile bacterial population differs from planktonic pneumococci in their physiology[159]. Studies on pneumococcal populations have shown the differential gene expression by transcriptome and proteome analyses[192]. These communities can produce a specific EPS matrix in comparison to planktonic pneumococci[186]. Biofilm formation increases the expression of different genes involved in EPS matrices protein synthesis, amino acid metabolism, carbohydrate metabolism, and non-glycolytic carbohydrate metabolism at the protein level[13, 193, 194]. In contrast, planktonic pneumococci show an upregulation of genes associated with virulence factors, cell-cell communication, metabolic activity, motility, and chemotaxis[192]. It has been shown that biofilm bacteria were more pathogenic than planktonic bacteria in pneumonia and meningitis models[177]. This suggests that the pneumococcus responds to host environmental changes[194, 195]. In support of this notion, it was shown in humans and mice that pneumococci develop biofilms to persist in the harsh nasopharyngeal environment[196 197]. Moreover, some pneumococcal proteins within the biofilm matrix are enzymes, such as serine proteases, that are involved in the degradation of the EPS

matrix and the initiation of a new biofilm life-cycle[198, 199]. These enzymes might be a crucial step for preparing dispersed bacteria to adapt to a new environment. Because colonization and the subsequent event of invasive infection appears to be two different states in pneumococcal pathogenesis, this might be an important consideration for future treatments[183].

Several *S. pneumoniae* proteins have potentially a role in biofilm development[190]. For instance, the analysis of the pneumococcal TIGR4 strain showed, among other things, the increased expression of putative adhesins, choline-binding proteins, and pili proteins (RrgA, RrgB, and RrgC)[186], membrane synthases and cell wall components during biofilm formation. The deletion of these factors in pneumococci led to reduced nasopharynx colonization in a murine model[179]. Similar results were found during adherence and biofilm formation, demonstrating the interaction of both mechanisms during colonization of the host. Therefore, these findings establish a link between biofilm formation and colonization[179]. Furthermore, the inactivation of virulence factors such as autolysin, pneumolysin, and PspC formed less structured and more antibiotic-sensitive biofilms in strain D39. Interestingly, pneumococci lacking PspA, which is not correlated with early colonization, formed the same biofilm as the wild-type[12].

Last but not least, enzymes like proteases, glycosidases, and DNase can inhibit pneumococcal biofilm formation[78, 184, 200-204]. For example, biofilm treatment with DNase I inhibits biofilm development[205]. Serine proteases were thereby the most effective enzymes against biofilms among these enzymes[184]. Studies on *Enterococcus faecalis* biofilms have shown that two secreted serine protease (SprE) and gelatinase (GelE) contribute directly to biofilm formation due to an increase in the autolytic rate and higher production of eDNA[205]. A SprE mutant developed high biofilm biomass with more production of eDNA as a component of the biofilm matrix. In contrast, the absence of GelE appeared to reduce the biofilm biomass compared to the parental strain. Furthermore, serine protease Esp (also known as SspA) from *Staphylococcus epidermidis* inhibits colonization, biofilm formation and disrupts preformed biofilms of *S. aureus*[206, 207]. Recently, pneumococcal serine protease HtrA has been shown to influence bacterial release (biofilm dispersal) from heat-induced biofilm[199]. HtrA, unlike the other serine proteases, increases its activity with increasing temperature[208]. The study suggests that serine protease activity is essential and associated with changes in the colonization environment during a virus co-infection, such as higher temperatures[199]. This finding turns the pneumococcal serine protease to be necessary for bacterial release during disease transition, and may also alter the biofilm development[199].

3.4 Extracellular proteases

3.4.1 Overview

Proteolytic activity is an essential biological process for survival of many organisms and it is not surprising that all living organisms have proteases[209]. In general, extracellular and intracellular proteases break down polypeptides or proteins at their amino- or carboxyl-terminal ends[209]. They are members of several families such as the subtilisin family, signal peptidase, Lon-A peptidase, Clp protease, amino- and carboxypeptidase[209]. However, proteases that have a serine residue as part of the catalytic center are ubiquitous in viruses, bacteria and eukaryotes. Almost one-third of all proteases can be classified as serine proteases. Many protease families have been recognized, and they have been classified into clans, based on structural similarities and other functional evidence[210]. The reaction mechanism of various serine proteases is similar. The catalytic triad of serine, aspartate, and histidine is shared by trypsin-like proteases (TLPs), or chymotrypsin, and subtilisin. Serine functions as a nucleophilic component, aspartate as an electrophile, and histidine as a basic amino acid to cleave the peptide bond[210].

In bacteria, proteases are important during growth, and they can serve bacteria in different ways, including penetration of host barriers, proliferation, protein transport, degradation of improperly folded proteins, and structural integrity of the cell wall[211]. In the context of pathogenesis, serine proteases can directly affect a wide range of host immune system functions such as degrading host cell membrane proteins and damaging tissues, inhibiting immune cell activation, blocking the classical, and alternative complement pathway by preventing opsonophagocytosis[212, 213]. However, many pathogenic bacterial species express extracellular proteases that play a significant role in pathogenesis such as *Haemophilus influenzae*, *Mycobacterium tuberculosis*, *Pseudomonas aeruginosa*, and other streptococcal species like *Streptococcus agalactiae* (group B streptococcus, GBS) and *S. pneumoniae*[214-216]. Most of these proteases have been shown to play a vital role in pathogenesis. To name a few, *P. aeruginosa* is the causative agent of burn wound infections, pneumonia, and eye infections and expresses many proteases involved in host tissue damage. For instance, the protease elastase B causes extensive damage to the cornea during ocular infection[215]. Similarly, *M. tuberculosis* serine protease increases bacilli entrance into human macrophages as shown by *in vitro* macrophage-based studies[217]. Recent studies show that the dissemination of pneumococci to the lungs, blood and then spread to the brain can provoke a reaction by the bacteria that result in their proteases serving multifunction roles such as self-proteins processors and virulence factors for host protein targets[211].

3.4.2 Pneumococcal proteases and peptidases

Serine proteases have been postulated as part of the invasion strategies for many pathogenic bacteria. Pneumococci express a wide range of hydrolase proteases and peptidases, including cysteine proteases, zinc-metalloproteases, and serine proteases[211, 218, 219]. The proteases are important surface-exposed proteins of *S. pneumoniae* and are likewise featured with a broad repertoire of different functions on pneumococcal pathogenesis[211]. More than 34 proteases known in *S. pneumoniae* were identified due to development of molecular tools[211, 220]. These proteases were annotated and deposited based on the pneumococcal TIGR4 strain, whose genome has been completely sequenced, and pneumococcal experts have manually assigned more than 2240 genes[54]. Pneumococcal peptidases and proteases with their known functions in colonization, transmission, pneumonia and septicemia are shown in **Figure 3.3**. These proteases have different functions including involvement in nutrient acquisition, protein quality control, signal cleavage for pre-proteins secretion and cleavage of host ECM proteins[211, 221]. It was shown that some proteases play a significant role in virulence[13, 222]. For instance, the zinc-metalloprotease A (ZmpA, also known as IgA1-protease) is a cell wall-associated metalloprotease, which can be released into the environment during the late stationary phase[218, 223]. IgA1-protease interacts with the host immune system by cleaving human IgA1 into inactive components at the hinges region of Fc fragments[224]. The left Fab fragment promotes phagocytosis by accumulation on the bacterial surface and enhances their attachment to the epithelial cells via the receptor for the platelet-activating factor (rPAF)[152]. Moreover, the zinc-metalloprotease ZmpB is important for the modification of pneumococcal surface proteins[225, 226]. It has been reported as one of the important virulence factors in mice models of pneumonia and bacteremia. Additionally, ZmpB cleaves human collagen IV via proteolytic activity, thereby compromising collagen to promote pneumococcal invasion into the blood[225]. The most critical pneumococcal proteases with specific functions on pathogenesis are summarized in **Table 3.1**.

Apart from interactions at the host cell membrane, the complement pathway is one of the important molecular events that affect pneumococcal pathogenesis[211, 227]. The opsonophagocytosis of pneumococci by phagocytic cells (macrophages and neutrophils) is one of the main outcomes of the classical complement activation[227]. The outcome of this pathway is dependent on the effective deposition of human complement component C3b on the bacterium. Complement deposition can be inhibited by pneumococcal proteases such as pneumococcal histidine triad protein A (PhtA), C3-degrading protease (CppA) and neutral

endopeptidase O (PepO) by the classical complement pathway[228, 229]. PhtA and CppA proteases can degrade the component factor C3, which is required for pathogen recognition by phagocytes[211]. PepO binds to complement component 1q (C1q), and can also serve as complement activator and enables premature consumption by blocking the regulatory inhibitor C4 binding protein (C4BP)[230]. In the presence of a urokinase-type plasminogen activator (uPA), PepO binds and cleaves plasminogen into its active form plasmin, resulting in the cleavage of C3b. This also leads to the inhibition of C3 deposition on pneumococci, which is required for phagocytosis by binding to the complement receptor during opsonization[211, 231].
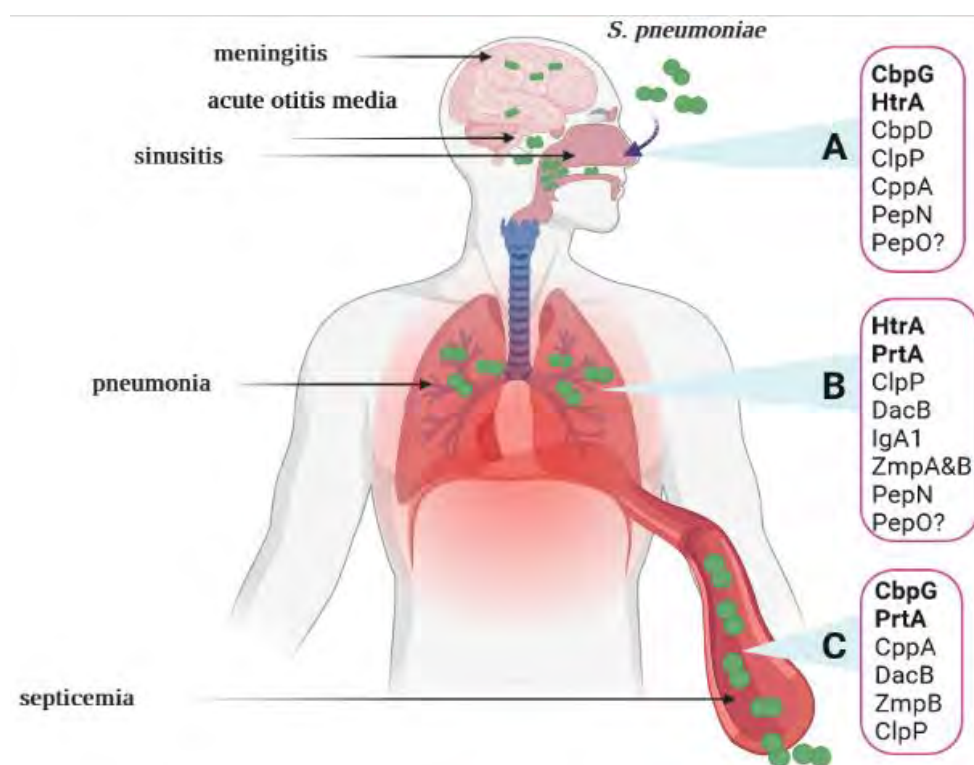
Figure 3.3: Pneumococcal proteases and peptidases.
Pneumococci can cause life-threatening invasive diseases such as pneumonia, bacteremia, and meningitis, or non-invasive diseases such as acute otitis media, sinusitis, and pneumonia without bacteremia. Pneumococcal peptidase and proteases are involved in these infections. (**A**) At least seven proteases are involved in adherence to epithelial cells during colonization. After colonization, (**B**) pneumonia and (**C**) septicemia are two pneumococcal diseases where several proteases are involved. Figure adapted from Marquart et al.,[211] and created with **BioRender**.com.

As mentioned before, pneumococci express different surface proteins grouped into several classes[88, 103, 108]. Several sortase-anchored proteins act as membrane-localized proteases and transpeptidases and are covalently anchored to the PGN via the sortase A, which cleaves the C-terminal LPxTG motif[87, 232, 233]. In addition, the pneumococcal cell wall is decorated with up to 16 CBPs, which are non-covalently bound to the phosphorylcholine of teichoic acids[100]. All pneumococcal serine proteases can be secreted and exposed on the bacteria cell surfaces as shown in **Figure 3.4**, depending on the presence of a signal peptide. Thus, this extracellular localization enables a direct or indirect cleavage of peptide bound, leading to the degradation of specific ECM components[87]. In fact, pneumococcal serine proteases are reported to play an important role on pneumococcal pathogenesis such as adhesion, colonization, and establishment of pneumococcal diseases, biofilms dispersal, and immune subversion of host cells[88, 89, 92, 106, 199, 234].

The first description of various pneumococcal serine proteases with their susceptibilities to different inhibitors was reported in 1991 by Courtney et al[235]. Recent studies were performed to evaluate the functional role of different serine proteases. Pneumococci express up to four surface-anchored or surface-associated serine proteases depending on the serotype, namely the high-temperature requirement A protein (HtrA), the subtilase family protein (SFP), the cell wall-associated serine proteinase A (PrtA), and the choline-binding protein G (CbpG). These enzymes are encoded by genes of the core genome and are highly conserved and present among a wide range of pneumococcal serotypes[236, 237]. Importantly, the proteolytic activity is characterized by three conserved amino acid (aa) residues, Ser-His-Asp, which form the so-called catalytic triad[236, 237]. Some of these proteases have already been reported to have an impact on pneumococcal pathogenesis[106, 234, 238].

Table 3.1: Pneumococcal proteases involved in pneumococcal pathogenesis[211, 239].

Protein	Locus tag (TIGR4)	Name/characteristic (s)	Pathogenic function	Human host-target/or molecular function	Ref.
PrtA	*sp_0641*	Serine protease, cell wall-associated serine protease	Colonization, and pneumonia	Cleaves human apolactoferrin, interact with collagen IV and plasminogen cleavage of leader peptides from lantibiotics possible adhesion	13, 236, 239, 240
CbpG	*sp_0390*	Serine protease, choline-binding protein G	Colonization, and pneumonia	Cell-attached form promotes adherence extracellular form degrades fibronectin important for mucosal and invasive disease	13, 100, 106, 239
SFP	*sp_1954*	Subtilisin-like/ or S8-family serine protease	Colonization, and pneumonia	Cleavage of leader peptides from lantibiotics	211, 234, 239
HtrA (DegP)	*sp_2239*	Serine protease; serine protease/ heat shock protein	Adherence, colonization pneumonia	Quality control of secreted proteins, denatured proteins	155, 234, 239, 241-245
CbpD	*sp_0341*	Choline-binding protein D, murein hydrolase	Colonization, and adherence	Lysis of the pneumococcal peptidoglycan	100, 246
ClpP	*sp_0746*	Caseinolytic protease P	Adherence, colonization, systemic infection and phagocytosis by DC	Cleaves proteins tagged for degradation (complex with ClpA or ClpX)	247-254
CppA	*sp_1449*	C3-degrading Protease	Colonization and invasion	Cleaves the complement factor C3	228, 255-258
PepN	*sp_0797*	Aminopeptidase N	Adherence, colonization, invasion	Inhibit T-cell receptor Metabolism pathway	259, 260
PepO	*sp_1647*	Neutral endopeptidase O	Adherence and host invasion, and enhance phagocytosis	Cleaves plasminogen to plasmin Cleaves the complement C3b	230, 231, 261-263
DacB	*sp_0629*	Murein peptidoglycan B peptidase	Pneumonia and septicemia	L,D-carboxypeptidase, peptidoglycan hydrolysis	264-266
PhtA	*sp_1175*	Pneumococcal histidine triad protein A	Immune evasion, cellular level	Degrades the complement factor C3	228, 258
IgA1-protease (ZmpA)	*sp_1154*	Immunoglobulin A1 protease	Adherence, and phagocytosis	Cleaves the human IgA1	216, 267-271
ZmpB	*sp_0664*	Zinc metalloprotease B	Pneumonia, septicemia, and inflammation	Cleaves human collagen IV	225, 226, 272-276
ZmpC	*sp_0071*	Zinc metalloprotease C	Pneumonia,	Cleaves the human matrix metalloproteinase-9 Cleaves the P-Selectin Glycoprotein-1 on neutrophils	277-282

3.4.3 Extracellular pneumococcal serine proteases

3.4.3.1 High-temperature requirement protein A (HtrA)

The serine protease HtrA is one of the best-studied proteases, which acts on misfolded proteins as chaperones e.g., after heat shock. HtrA is widely distributed and was found in eukaryotes and Gram-negative\positive bacteria. HtrA displays trypsin-like protease activity and belongs to the peptidase S1C family and the subfamily of DO proteases with more than 180 members[283, 284]. The family of these proteases exhibits a central catalytic domain with at least one or more C-terminal PDZ domains[285]. The protease is highly conserved in both, pathogenic and nonpathogenic bacteria[286, 287]. Furthermore, HtrA was reported to be homologous to the degradation of periplasmic proteins (DegP) or DO protease in *E. coli*[285].

Bacterial HtrA is a heat-shock-induced serine protease that displays a multifunctional role such as protein quality control and bacterial survival[287-291]. Quality control protein function serves by preventing cellular malfunctions due to the accumulation of fragmented proteins. Bacterial survival is granted due to the chaperone capacity under different stress conditions such as oxidative and heat stress[155, 292]. For instance, HtrA protease in *Lactococcus lactis* is considered as a housekeeping protease[293], while in other bacteria, HtrA prevents the cell from cytotoxicity of misfolded proteins by refolding or degrading them[288, 290]. In *E. coli*, unlike other quality control proteins, such as matrix protease ClpXP, ClpAP, and HslUV, which need ATP for their chaperone function, HtrA is functional without ATP as an additional energy source[294, 295]. More importantly, the function of HtrA can be switched from chaperone to protease function, depending on the temperature[286]. The protease effect is most apparent at higher temperatures between 38-42°C, whereas the chaperone functions at lower temperatures between 30-37°C[286]. Interestingly, the amino acid sequence of HtrA serine proteases of other pathogenic bacteria revealed a high sequence similarity, especially in the functional protease and PDZ domains[287, 292, 296]. HtrA is widely distributed in among bacterial species such as *Escherichia coli*, *Legionella fallonii*, *Thermotoga maritima*, *Mycobacterium tuberculosis* and *Streptococcus pneumoniae*[292, 295, 297-299].

In *S. pneumoniae*, HtrA was described for the first time 20 years ago, when it was found to reduce expression in a strain deficient in competence induction and alter cefotaxime susceptibility response regulator/histidine kinase (CiaRH), which regulates HtrA[155]. Since then, HtrA has been extensively studied and shown to significantly influence various functions such as bacterial fitness, adaptation to environmental stress, or enhancing pneumococcal virulence[244, 287, 300].

Pneumococcal HtrA contains no specific anchoring motif and can be found on the surface or is secreted from *S. pneumoniae* as predicted by the presence of a putative amino-terminal signal peptide[234, 239]. The surface-exposed HtrA promotes nasopharyngeal colonization, whereas secreted HtrA facilitates the subsequent invasion of host tissue by degrading ECM components[287]. The influence of HtrA on pneumococcal pathogenesis has been addressed in many studies. For example, HtrA is considered as one of the most critical serine proteases in pneumococcal virulence because HtrA degrades the competence stimulating peptides (CSPs), which directly impacts pneumococcal competence and late competence genes which in turn affects virulence[241-243]. The involvement of HtrA in pneumococcal competence is still challenging to understand due to conflicting results. Ibrahim *et al*. have shown that HtrA protease is important for competence, as the pneumococcal transformation efficiency was highly reduced in the *htrA*-mutant[242]. In another study, the proteolytic activity analysis performed with purified recombinant pneumococcal HtrA revealed that HtrA cleaves the pneumococcal CSP *in vitro*[243]. Since CSP has a significant effect on pneumococcal transformation[301], this fact suggests that HtrA has a considerable effect on pneumococcal transformation efficiency and is needed to stimulate competence. This study has also shown that the deletion of *htrA* or catalytic residues did not affect natural DNA competence[243]. In addition, the regulation of HtrA seems to be dependent on bacterial culture conditions. It has been shown that HtrA inhibits competence in a complex medium but not in a chemically defined medium[111]. These findings show that HtrA acts as a competence regulator at the protein level, and that environmental factors influence its regulation[211].

Apart from the involvement of HtrA in competence, HtrA is necessary for nasopharyngeal colonization and virulence[155, 234]. The pneumococcal two-component (TCS) CiaRH system is required for pneumococcal nasopharyngeal colonization. The expression of *S. pneumoniae* HtrA is regulated and controlled by the CiaRH system[155]. In this study, transcriptome analysis identified *htrA* and *spoJ* to be highly downregulated in the *ciaRH*-mutant[155]. In a competitive colonization model with infant rats, the *htrA*-mutant also showed reduced fitness. These findings support that HtrA mediates a portion of the carriage deficit observed with the Δ*ciaRH* strain. Therefore, HtrA was observed to be a CiaRH controlled protein essential for nasopharyngeal colonization[155]. A recent study showed that HtrA regulated by CiaRH is responsible for penicillin-binding protein 2x (PBP2x) degradation[302].[92]

Furthermore, HtrA protease is required for pneumococcal virulence in a murine pneumonia model. Mice infection studies with *S. pneumoniae* D39 demonstrated that the deficiency of the serine protease HtrA decreases the bacterial load and inflammation in the lung after intranasal

infection[234]. It has been demonstrated, that pneumococci can form well-structured biofilms associated with different pathophysiological events in bacterial colonization, persistence, chronic infections and even invasion during co-infection[186]. HtrA has been shown to modulate bacterial release (biofilm dispersal) from heat-induced biofilms, mimicking fever conditions in co-infections[199]. During influenza-pneumococcal co-infections, HtrA also highly induced the inflammation, thereby enhancing the bacterial load in a mouse pneumonia model[245].

Finally, as part of pneumococcal therapy, HtrA is immunogenic and protective in mice against invasive pneumococcal diseases[303]. Indeed, targeting bacterial HtrA offers some potential advantages. These include first the localization as HtrA can be surface-exposed or secreted, thus it is easily accessible for potential inhibitory substances or drug compounds. Second, HtrA has a defined enzymatic active site and substrate recognition domain. These two features make HtrA a potentially attractive candidate for desirable drug targets to prevent pneumococcal diseases[152, 304].

3.4.3.2 Choline-binding protein G (CbpG)

As mentioned above CbpG is a member of the CBP family (**3.2.3.1**), consisting of two conserved domains: the enzymatically active domain and the choline-binding module (CBM). CbpG can non-covalently bind to the phosphorylcholine moieties of pneumococcal LTA and WTA via its CBM[92, 99] (**Figure 3.4**). The CbpG protease belongs to the peptidase S1, PA clan superfamily of peptidases, and is a trypsin-like serine protease[305]. The protein sequence indicates that this protein possesses a chymotrypsin-like fold and double beta-barrel structure with a carboxyl-terminal choline-binding domain[283, 306].

Pneumococcal CbpG is considered to be a multifunctional surface-exposed serine protease with both proteolytic and adhesive functions[100, 106, 307]. Like other CBPs, CbpG also plays a significant role in pneumococcal colonization and sepsis[100, 308]. When CbpG is mutated, a significant reduction in the ability of pneumococci to colonize the rat nasopharynx was observed. Therefore, high reduction in mortality was observed in septicemia during the systemic infection with infant rat[106]. This fact suggested that CbpG is required for pneumococcal adhesion and colonization[106]. This turns the protein into an important virulence factor of *S. pneumoniae*. Such multifunctional proteinases can be found in many pathogenic bacterial species. The C5a peptidase of group B streptococcus[309, 310] and the well-characterized Pla surface protease of *Yersinia pestis*[311] are striking examples.

Based on the pneumococcal serotype, there are at least two variants of the CbpG expressed by *S. pneumoniae*. In strain D39, CbpG is truncated due to a premature stop codon after the N-

terminal catalytic domain. This variant is probably extracellularly secreted and released then into the environment. On contrary, the full-length CbpG variant in TIGR4 strain is cell wall-associated[106]. The modular characterization of CbpG shows that the catalytic residues are present independent of expressing a full-length protein containing a CBM domain[106]. In both configurations, the protein exhibits proteolytic activity as confirmed by Mann et al[106]. The secreted CbpG is even able to degrade ECM deposited fibronectin via its trypsin-like protease activity. The functional portion that already shows activity to degrade the substrates casein and fibronectin, independently whether the CBM is present or absent[106]. In addition, it has been proposed that CbpG can play a role in the pneumococcal transition to the blood, which may be due to its fibronectin-cleaving potential[106]. Therefore, CbpG might be a moonlighting protein with different functions or alternately expressed depending on host niche conditions[211].

Importantly, pneumococci can penetrate the extracellular matrix for disseminating human host tissue and cause invasive disease successfully. This requires the serine protease activity of pneumococcal bacteria to be cell surface-associated. To date, many of the pneumococcal proteases have been shown to have a multifunctional role, including peptidoglycan cleavage[312], proteolytic activity and virulence such as ClpP and CbpG[100]. In several mouse models, ClpP was shown to be required for nasopharyngeal colonization, lung infection, and systemic infection[250, 313] (**Figure 3.3**). Unlike ClpP, CbpG has not been ascribed the role of a bacterial protein processor. The protease ClpP might be expected to be involved at all stages of pneumococcal infection by its regulatory function as the ClpP repressed the competence for the bacteria[211]. It is not surprising that CbpG has an important role during invasive pneumococcal disease, because it has been shown that it is upregulated in all *in vivo* niches[314]. However, the molecular mechanism of proteolytic activity in pathogenesis remains to be elusive. The functional CBM in the C-terminal part of CbpG is needed to contribute to pneumococcal adherence and colonization[106].

On the other hand, like many other CPBs pneumococcal antigens[237], CbpG has significant immunogenicity as shown in different mouse models. All infected mice developed anti-CbpG antibodies that protected mice against colonization and systemic infection[106]. Similarly, another study reported that the recombinant proteins CbpM and CbpG of serotype 19F effectively induced a protective antibody response[307]. In addition, anti-CbpG antibodies enhanced neutrophil-mediated opsonophagocytosis. Therefore, CbpG is considered as one of the vaccine candidates against pneumococcal diseases[307, 315].

3.4.3.3 Cell wall-associated serine protein A (PrtA)

Pneumococcal PrtA is a serine protease enzyme that belongs to the subtilisin-like family (also known as subtilases), which is part of the S8 family peptidase[211, 236]. The PrtA protease in pneumococci is related to serine proteases like PrtP in lactobacilli, cleaving the amino-terminal leader sequences from lantibiotics. However, both, streptococci and lactococci have a wide range of endopeptidase enzymes that have physiological and pathogenic functions[316].

In pneumococci, PrtA is an important surface serine protease involved in pneumococcal pathogenicity[236, 317]. The role of PrtA in colonization and subsequent host invasion seems to be strain-specific[318]. The first report on pneumococcal PrtA emphasized immunogenicity due to the fact that it was identified using convalescent-phase serum[317]. Interestingly, a previous study showed that the *prtA* gene is a prevalent and highly conserved virulence factor in pneumococci and is found almost in all analyzed strains[236]. Both *in silico* analysis and flow cytometry confirmed that PrtA is surface localized[319]. The first molecular characterization of PrtA was published in 2001[236]. Bethe and co-workers[236] showed that *S. pneumoniae* express PrtA with different molecular weights. The full-length variant produced by pneumococci has a molecular weight of 240 kDa, whereas a shortened form has only a molecular weight of 215 kDa, which cannot be explained by signal peptide cleavage only. The same observation for maturation was also found in related proteases such as PrtP of *Lactobacillus paracasei* and *Lactococcus lactis*[320, 321]. PrtA consists of two domains the catalytic domain and DUF-1034 domain (Domain of Unknown Function). Of interest, the multisequence alignment of PrtA catalytic triad residues (Asp^{232}-His^{299}-Ser^{690}) were highly homologous to other related bacterial species of subtilisin-like serine proteases[236]. These catalytic triads showed a high degree of similarity and identity to the cell wall-associated proteases of *Streptococci*, *Lactococci*, and *Lactobacilli*[236, 322]. The DUF-1034 domain is probably involved in substrate recognition as predicted by de Stoppelaar et al.,[234].

As mentioned above, PrtA has been reported as a major surface-associated serine protease involved in pneumococcal virulence[234, 236, 317]. First, PrtA helps to cleave human apolactoferrin into an active lactoferricin-like peptide, which is a cationic antimicrobial peptide that facilitates the killing of pneumococci. This function is rather surprising because it decreases the pathogenicity potential of pneumococci[240]. Second, with a molecular weight of 240 kDa, PrtA is one of the biggest pneumococcal surface proteins and is thought to have adhesive functions comparable to other sortase-anchored pneumococcal proteins[225].

In a systemic mouse infection model, where mice were infected with the *prtA*-mutant of strain D39, extended survival times could be observed in comparison to D39 wild-type infected

mice[236]. The *prtA*-negative strain is furthermore significantly attenuated in an intranasal mouse infection model. Transcriptome studies showed an increase of PrtA expression in pneumococci isolated from blood[236].

Furthermore, PrtA degrades collagen IV and plasminogen from the ECM and serum components. This indicates that these substrates cleavage can lead to pneumococcal transcytosis of the mucosal barrier and dissemination into the bloodstream[225, 318]. Last but not least, PrtA has also been demonstrated to stimulate the production of IL-17A, which is a key mediator of tissue inflammation[323]. Interestingly, PrtA seems to be highly immunogenic in humans and mice[317, 319]. Nevertheless, it failed to effectively protect mice against pneumonia and septicemia.

3.4.3.4 Subtilase family protein (SFP)

The SFP serine proteases (known as serine peptidase) is another enzyme that can cleave leader peptides from lantibiotics[211]. SFP was identified as EpiP (epidermin leader peptide processing serine protease) which belongs to the subtilisin-like/or S8-family serine protease similar to PrtA[211, 234].

In 2013, de Stoppelaar et al.,[234] identified SFP in the *S. pneumoniae* D39 strain for the first time and showed that this protein has a catalytic domain and a sortase A anchoring motif. In their study, they also showed the involvement of SFP in the pathogenesis of pneumococcal infections, which is still not apparent due to the minor effect in virulence of the *sfp*-negative strain, tested in mice[234]. In a pneumonia model, mice infected with a low dose of a *sfp*-mutant revealed a lower bacterial load in the lungs, and the immune cell influx during pneumonia was modest in comparison to an infection with the D39 wild-type. In contrast, after a high dose infection the number of the reisolated *sfp*-mutant bacteria was comparable to the parental strain[234]. However, this was shown in the presence of all other serine proteases produced by D39 strain. Hence, this study suggested further investigation on the functional role of SFP in different strains backgrounds.

SFP protein reveals a high homology to the cell surface serine endopeptidase CspA[234]. CspA is one of the important virulence factors of the human pathogen *Streptococcus agalactiae*[324]. Opsonophagocytosis of bacteria by host immune cells is one of the critical outcomes of classical complement activation[325]. The complement component C3b deposited on the *S. agalactiae* cell surface can be cleaved by CspA, indicating the importance of CspA for immune evasion[324]. So far, the impact of complement inactivation by its pneumococcal orthologue SFP is not known.
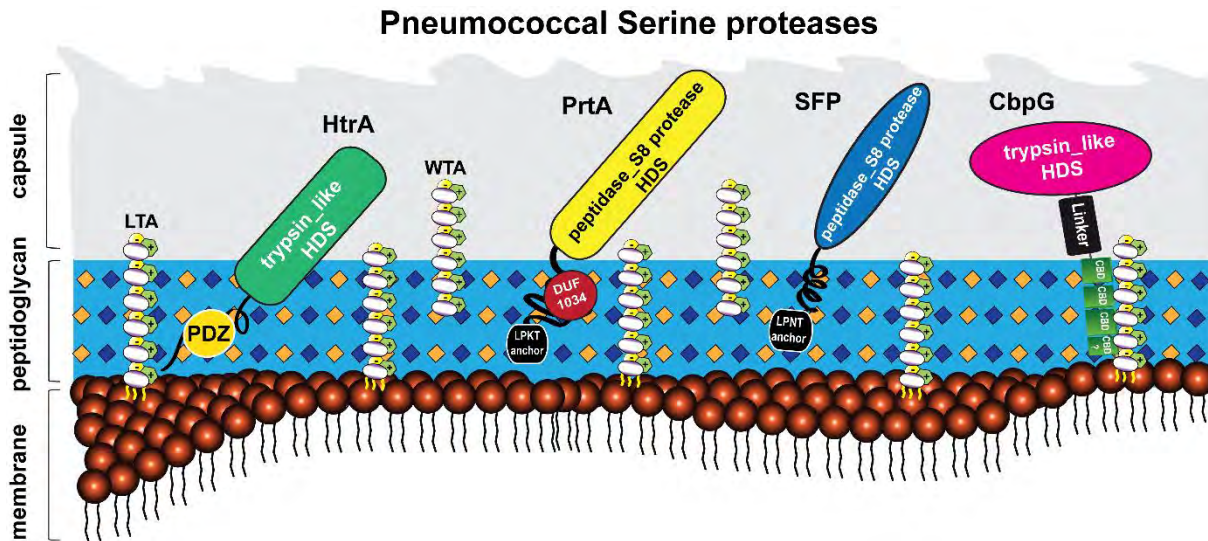
Figure 3.4: Localization of pneumococcal serine proteases on the surface of bacteria.

The pneumococcal cell wall of *S. pneumoniae* contains four different classes of surface-exposed proteins: choline-binding proteins (CBPs), sortase-anchored LPXTG proteins, lipoproteins, and non-classical surface proteins[103, 108, 326]. These proteins are associated with different structures of the cell wall, consisting of peptidoglycan (light blue), wall teichoic acids (WTA) and lipoteichoic acids (LTA) (carbohydrates repeating units in white circles for WTA and LTA). Wall teichoic acids are directly linked to the peptidoglycan (PGN), lipoteichoic acids are anchored to the phospholipid bilayer (membrane) *via* a lipid anchor[327]. Pneumococcal teichoic acids are decorated with phosphorylcholine (*PCho*) residues[328]. The pneumococcus displays four serine proteases on the bacterial surface[106, 234]. Choline-binding protein G (CbpG, pink); bound non-covalently, via the conserved choline-binding repeats (CBRs green) to the phosphorylcholine residues of WTA or LTA. Subtilase family protein (SFP, blue) and cell wall-associated serine proteinase (PrtA, yellow) belong to the subtilisin-like proteases. Both contain an N-terminal signal peptide and a C-terminal LPXTG motif. The latter is necessary to bind SFP and PrtA to the peptidoglycan, catalyzed by the transpeptidase sortase A. High-temperature requirement A (HtrA, green) belongs to the family of trypsin-like proteases and contains no specific cell wall anchoring motif. All of these proteins contain the catalytic active protease domain with the Asp-His-Ser triad (HDS)[329].

3.5 Objectives and outline of this thesis

Pneumococci are opportunistic human pathogens causing life-threatening diseases, including pneumonia, sepsis and meningitis. A plethora of virulence factors is crucial for *S. pneumoniae* as a commensal and as a pathogen. *S. pneumoniae* can secrete numerous proteins, which are contributing to adhesion, colonization, and infection. Among these secreted proteins are the extracellular pneumococcal serine proteases. The individual role or the synergistic effects of pneumococcal extracellular serine proteases on colonization and subsequent dissemination into the lung or blood have not been analyzed systematically. The redundancy of serine proteases may compensate for the deficiency of a single enzyme. Therefore, serine proteases have the potential to facilitate pneumococcal colonization and binding to their host targets. This function suggests that they could be promising targets for developing specific antimicrobials to reduce pneumococcal colonization and transmission.

The main goal of this work was to investigate the impact of secreted or cell-bound pneumococcal serine proteases HtrA, PrtA, SFP, and CbpG on pneumococcal pathogenesis. Bioinformatics analysis tools for pneumococcal serine proteases genes are available in the databases and were employed to study their molecular characteristics and genetic organization. This study is aiming to understand the individual role of each serine protease in double and triple mutants. Mutants deficient in the target serine proteases were generated by insertion-deletion mutagenesis in the non-invasive serotype 19F strain EF3030 and the invasive serotype 4 strain TIGR4. Furthermore, bacterial fitness of constructed mutants was assessed in *in vitro* growth experiments using complex and chemically-defined media. To study the impact of serine proteases, *in vitro* adhesion assays and an *in vivo* murine nasopharyngeal colonization model were applied using the non-invasive serotype 19F strain EF3030. In addition, the role of serine proteases on virulence and dissemination was assessed in an acute murine pneumonia model, and in *in vitro* phagocytosis assays using the invasive TIGR4 strain and isogenic triple *serine protease* mutants. Nasopharyngeal colonization is associated with biofilm formation, and hence biofilms are considered as a critical event in pneumococcal pathogenesis. The role of extracellular serine proteases contributing to pneumococcal biofilm formation were evaluated using static and substratum biofilm models with serotype 19F strain EF3030. Finally, serine protease activity during pneumococcal infection can contribute to the destruction of the epithelial barrier or degradation of ECM components. Therefore, after heterologous serine protease expression in *E. coli* and subsequent protein purification, recombinant serine proteases were applied in proteases activity assays.

Chapter **4**

RESULTS

4. Results[1]

4.1 General overview

S. pneumoniae expresses up to four different extracellular serine proteases containing the typical Asp-Ser-His triad essential for catalytic activity. However, the role of these bacterial enzymes in colonization or epithelial barrier disruption is insufficiently explored. Previous studies focused on the individual effect of a single-knock out of these proteases. Thus, pneumococcal serine proteases were studied here in different experimental approaches, because of their possible compensatory impact on virulence during the interaction between *S. pneumoniae* and host tissue. The main focuses are to investigate the role of pneumococcal serine proteases in the host-pathogen interaction, barrier disruption of the epithelial cell layer, and pathophysiology of this human pathogen. Therefore, single, double and triple *serine protease* mutants were generated in the background of the invasive serotype 4 strain TIGR4 and colonizing strains of serotypes 19F and 35A. To that end, adherence studies were performed using human nasopharyngeal epithelial cells. *In vivo* infection experiments were carried out by intranasal infection of CD-1 outbred mice with bioluminescent TIGR4*lux* and 19F_EF3030 in murine models of acute pneumonia and colonization. Pneumococcal colonization is associated with biofilm formation, which can lead to invasion in the long term. Therefore, epithelial cells were incubated for different time periods with 19F_EF3030 wild-type or their isogenic serine proteases mutants to study the impact of the different serine proteases on biofilm formation. Furthermore, all serine proteases were heterologously expressed, and the recombinant proteases were isolated to analyze the protease activity *in vitro*.

4.1.1 Bioinformatic analysis

Many pneumococcal peptidases and proteases have been identified recently due to improved genetic methods, such as next-generation sequencing (NGS) and new tools for bioinformatics analysis[211]. In particular, up to four different extracellular serine proteases are present in pneumococci, PrtA, SFP, CbpG, and HtrA. Bioinformatics analysis was performed using different databases tools (*see* **heading 7.8**), to demonstrate the variability of pneumococcal serine proteases on the molecular level, which also has consequences for the functionality of

---

these important enzymes. These analyses include **i)** identifying the different domains of the pneumococcal serine proteases, **ii)** demonstrating the genome organization and the distribution in different pneumococcal serotypes and other bacteria species, and **iii)** comparing the hypothetical 3D structural models.

4.1.2 Gene sequence conservation of serine proteases among different pneumococcal strains

To study the homology of pneumococcal serine proteases, the genomes of 10 clinically relevant *S. pneumoniae* strains were analyzed on the DNA and protein level with BlastN for gene alignments and BlastP for comparisons of the deduced amino acid sequences[330]. The comparative analysis revealed a maximum of four different serine proteases present in pneumococci. Comparison on the protein level revealed high identities and similarities, indicating highly conserved sequences among the different pneumococcal strains. In general, these results indicated that the pneumococcal serine proteases CbpG, HtrA, and PrtA are highly conserved with at least ≥ 87% identity. In particular, the deduced amino acid sequence of CbpG in different strains is highly conserved with a range between 87 to 99% identity. Serine protease HtrA is well conserved in all analyzed serotypes D39, Hungary 19A, serotype 19F_EF3030, and R6 strain, resulting in high protein sequence identity (up to 100%). Serine protease PrtA has an identity ranging from 95.8% in these strains. The SFP protease is not present in t serotype 19F or 19A strains used for this study. In all other analyzed serotypes, the identity is 100%. (**Table 4.1**). Therefore, pneumococci have only three serine proteases serotype 19A and 19F strains (**Figure 4.8A**), and four proteases in most strains of the other, such as D39 (serotype 2), TIGR4 (serotype 4), ST81 (serotype 23F), JJA (serotype 14), and R6 (serotype 2).

Table 4.1: Protein sequence homology [%] of serine proteases among different selected pneumococcal strains based on the protein sequences from *S. pneumoniae* TIGR4[54], and D39 for SFP.

S. p. Strain (serotype)		gene no.	CbpG	Gene no.	HtrA	Gene no.	PrtA	gene no.	SFP
TIGR4 (4)	%ID	*sp_0390*	**100.0**	*sp_2239*	**100.0**	*sp_0641*	**100.0**	*sp_1954*	**100.0**
	%SIM		**100.0**		**100.0**		**100.0**		**100.0**
D39 (2)	%ID	*spd_0356*	**99.0**	*spd_2068*	**100.0**	*spd_0558*	**95.8**	*spd_1753*	**100.0**
	%SIM		**99.5**		**100.0**		**97.8**		**100.0**
EF3030 (19F)	%ID	*EF3030_01920*	**99.6**	*EF3030_11105*	**99.7**	*EF3030_03025*	**97.5**	----	---
	%SIM		**99.6**		**100.0**		**98.7**		
ST556 (19F)	%ID	*snd:MYY_0470*	**87.0**	*snd:MYY_2162*	**100.0**	*snd:MYY_0688*	**95.9**	----	---
	%SIM		**95.7**		**100.0**		**97.9**		
ST81 (23F)	%ID	*spn23F03640*	**97.8**	*spn23F22720*	**99.7**	*spn23F05790*	**97.4**	*spn23F9760*	**100.0**
	%SIM		**98.5**		**100.0**		**98.8**		**100.0**
JJA (14)	%ID	*spj_0378*	**99.3**	*spj2269*	**99.7**	*spj_0592*	**97.4**	*spj_1948*	**100.0**
	%SIM		**100.0**		**100.0**		**98.6**		**100.0**
R6 (2)	%ID	*spr0349*	**99.0**	*spr2045*	**100.00**	*spr0561*	**95.8**	*spr1771*	**100.0**
	%SIM		**99.5**		**100.00**		**97.8**		**100.0**
G54 (19F)	%ID	*spg_0356*	**100.0**	*spg_2188*	**98.4**	*spg_0584*	**96.1**	----	----
	%SIM		**100.0**		**99.0**		**97.8**		
Hungary 19A-6 (19A)	%ID	*sph_0499*	**96.8**	*sph_2438*	**99.5**	*sph_0733*	**96.2**	----	----
	%SIM		**96.8**		**100.0**		**98.1**		
R6_CIB17 (2)	%ID	----	----	*E5Q10_10910*	**100.0**	-----	----	*E5Q10_09305*	**100.0**
	%SIM		----		**100.0**				**100.0**

*Analysis of the proteins was performed with tool databases BlastP[330], and EMBOSS[331]. Protein sequences derived from TIGR4 strain were used as reference ID: Identity and SIM: Similarity. The meaning of the bold values are % ID, Identity percentage; % SIM, Similarity percentage.

4.1.3 Alignment of serine proteases to other bacterial species

The serine protease HtrA is widely distributed. Thus, it was of interest to determine whether serine proteases in *S. pneumoniae* are homologs to other bacteria. The comparison was performed with the BlastP tool and more than 70% sequence similarity of pneumococcal HtrA was detected to HtrA proteins of *S. pyogens* (group A streptococci), *S. agalactiae* (group B streptococcal), and *S. mutans* (**Table 4.2**). The PrtA protease shared a range between 35-38% similarities with other streptococcal species and lactobacilli subtilisin-like proteases. The protein sequence homology of the CbpG to other Gram-positive bacterial species was analyzed as well. Significant homologies of CbpG in a range between of 40-56% were found to serine proteinases of different bacterial species such as *Enterococcus faecalis*, *Staphylococcus aureus*, and *Salmonella enterica*. Finally, the complete SFP protein sequence homology to orthologues of other bacterial species was found only with *Streptococcus agalactiae* (38.5%) and with *Lactococcus lactis* (56.6%).

Table 4.2: Comparison of pneumococcal serine proteases to other bacterial species

Protein* (locus tag)	Bacterial species	Protein accession no.	Similarity [%]
HtrA	*S. pneumoniae*	AAK76286.1	**100.0**
(*sp_2239*)	*S.pyogenes*	BAQ53883.1	**71.7**
	S. agalactiae	Q8DWP1	**73.6**
	S. mutans	VEI61035.1	**72.8**
	Campylobacter jejuni	WP_061099826.1	**52.6**
	Helicobacter pylori	AHC56659.1	**54.5**
	Mycobacterium tuberculosis	5ZVJ_A	**52.7**
PrtA	*S. pneumoniae*	AAK74791.1	**100.0**
(*sp_0641*)	*S. pyogenes*	P15926	**38.1**
	S. agalactiae	U56908	**37.5**
	Lactobacillus delbrueckii bulgaricus	L48487	**37.4**
	Lactobacillus helveticus	AF133727	**36.5**
	Lactobacillus paracasei	M83946	**35.8**
	Lactobacillus lactis subsp.cremoris	J04962, M26310	**35.4**
	Streptococcus thermophilus	AAG09771	**35.3**
CbpG	*S. pneumoniae*	AAK74556.1	**100.0**
(*sp_0390*)	*Enterococcus faecalis*	GEJ60330.1	**56.0**
	Bacillus subtilis	P39790.1	**42.9**
	Bacillus intermedius	1P3C	**42.2**
	Staphylococcus aureus	WP_010922847.1	**40.4**
	Salmonella enterica subsp	NP_460444.1	**39.2**
SFP	*S. pneumoniae*	ABC75782.1	**100.0**
(*spd_1753*)	*Lactococcus Lactis*	4MZD_A	**56.6**
	Streptococcus agalactiae	CNG97209.1	**38.5**

*Analysis of the proteins was performed with tool databases BlastP and EMBOSS. Protein sequences derived from the TIGR4 strain were used to reference HtrA, PrtA and CbpG. D39 strain for SFP.

4.1.4 Domain organization of pneumococcal serine proteases

Pneumococcal serine protease gene organization was analyzed *in silico* to illustrate genomic arrangement. For this reason, serine protease gene sequences for *prtA* (*sp_0641*), *htrA* (*sp_2239*), *sfp* (D39: *spd_1753*) and *cbpG* (*sp_0390*) of *S. pneumoniae* strain TIGR4 and D39 were retrieved from the KEGG databases[332]. For strain 19F_EF3030 the gene sequences were obtained from the National Center for Biotechnology Information (NCBI)[330].

The pneumococcal HtrA protein (**Figure 4.1A**) has a molecular weight of 42 kDa (393 aa) including an amino-terminal signal peptide (31 aa), which is cleaved by signal peptidase I for secretion. The terminal signal peptide is followed by a single transmembrane helix domain (12-34 aa). Additionally, HtrA contains two highly conserved unique domains, a serine protease domain and a PSD-95/Dlg/ZO-1 (PDZ) domain[234]. The trypsin-like serine protease domain contains the typical triad His^{112}-Asp^{152}-Ser^{234} (HDS) in the catalytic center, which was identified previously[234] using Interproscan IPR009003 and IPR001940[333]. Finally, the PDZ domain (abbreviation combining letters of the first three proteins discovered to share this domain, **p**ostsynaptic density protein, **D**rosophila discs large tumor suppressor, and **z**onula occludens-1) is located at the C-terminal end. HtrA in other bacterial species contains one or more PDZ domain(s)[287, 292, 334]. The HtrA-PDZ domain acts as a protein folding stress sensor and controls the pyrolytic activity[300, 335, 336]. Thus, the PDZ domains are responsible for recognizing or binding substrate proteins[211]. Fan *et al.* solved the pneumococcal HtrA-PDZ structure[334], which contains three α-helices and five β-strands (amino acid residues 262-386). Moreover, a comparison of the amino acid sequences of HtrA-PDZ domains in different bacterial species showed that the pneumococcal PDZ domain, which is most likely involved in the ligand recognition, has only a moderate sequence similarity and conserved secondary structure[337].

The molecular analysis of full-length CbpG in TIGR4 comprises 285 aa with a molecular weight of 32 kDa as shown in **Figure 4.1B**. According to our SignalP 4.0 analysis[338], a leader peptide (secretion signal peptide) is not present in all analyzed serotypes except for the serotype 19F strain ST556 (*see* **8.4.1**). Therefore, it is still unknown whether and how CbpG is translocated from the cytoplasm to the bacterial cell surface. The functional domain is the trypsin-like domain with 184 aa, spanning from aa 14-197, containing the catalytic triad His^{34}-Asp^{87}-Ser^{159}. Furthermore, this domain is linked to the CBM by a short linker region (^{aa}Lys-Pro-Phe-Ile^{aa}) that provides flexibility to the protein and may provide stability to the catalytic

domain. This catalytic functional module exhibits sequence similarities to trypsin-like serine proteases present in all CbpG variants[100, 106].

Moreover, it has been mentioned that the CBM, for example in strain TIGR4, exhibits only three choline-binding repeats (CBRs), which are located at position aa 207-265. This represents the shortest identified CBM among all choline-binding proteins. It has been proposed that at least four repeats are needed to attach the protein non-covalently to the teichoic acids of the cell wall[339]. Therefore, it is still unknown if CbpG can bind to the bacterial cell surface when only three CBRs are present. It has been suggested by our genome re-analysis that a fourth CBR at position aa 267-285 might attach CbpG to teichoic acids and allows the protein to be surface-associated. This repeat includes the aromatic residues YW and fulfills the number of aromatic residues involved in choline-binding[340] to teichoic acids.

The PrtA protein sequence was analyzed using InterPro and the analysis indicates that this protein is part of the S8 family peptidase. The full-length form of PrtA has a molecular weight of 234 kDa (2140 aa) after cleavage of the leader peptides and integration into the peptidoglycan via sortase A. PrtA contains a typical sortase A recognition LPKTG motif followed by a hydrophobic region at the carboxy-terminus. The sortase A catalyzes covalent anchoring to the bacterial PGN[234, 239]. PrtA consists of two domains, the active peptidase-S8 domain, which contains the typical catalytic triad (Asp^{232}-His^{299}- Ser^{690}), spanning the region between aa 223-764[236, 239], and the second domain DUF-1034 (domain of unknown function), which consists of 140 amino acids and is localized between aa residues 795-934. The modular organization of PrtA is illustrated in (**Figure 4.1C**).

Based on the molecular characterization of SFP[234], it has been predicted to have protease activity and to be secreted on the surface. The full-length SFP has a molecular weight of 64.9 kDa and consists of an N-terminal signal peptide (1-22 aa) and a LPNTG anchoring motif at the C-terminal region. This motif is thought to be functional as a target site for the sortase A and anchoring the protein to PGN. In addition, the peptidase domain spanning the aa residues 167-461 contains the catalytic triad (Asp^{176}-His^{223}-Ser^{280}) (**Figure 4.1D**).

Modular organization of pneumococcal serine proteases

Serine proteases	N-terminal SP	Catalytic domain	Other domain architecture	TM helix	Length (amino acids)
A HtrA_*sp_2239*	1-32 N—	chymotrypsin-like serine protease 96-277 **HDS** 182 aa	289-375 PDZ —C	12-34	42 kDa 393 a.a
B CbpG_*sp_0390*	? N—	trypsin-like serine protease 14-197 **HDS** 184 aa	198 206 207 265 Linker CBD CBD CBD CBD? —C	0	32 kDa 285 a.a
C PrtA_*sp_0641*	1-27 N—	peptidase_S8 protease like 223-764 **HDS** 542 aa	795-934 **DUF 1034** 140 aa 2099-2140 LPKTG anchor —C 42 aa	2112-2134	240 kDa 2140 a.a
D SFP_*spd_1753*	1-22 N—	peptidase serine protease 167-461 **HDS** 296 aa	543-550 LPNTS anchor —C 37 aa	552-574	64.9 kDa 579 a.a

Figure 4.1: Protein sequence domain organization of pneumococcal serine proteases. (**A**) HtrA (AAK76286.1), the signal peptide sequences (aa 1-32) are illustrated in red. The serine protease catalytic domain is shown in light green, PDZ domain is labeled in yellow. (**B**) CbpG (AAK74556.1), most likely has no signal peptide. The trypsin-like serine protease catalytic domain is shown in pink, the repeats of the choline-binding domains (CBDs) are marked in green, connected by short linker region aa 198-206. The C-terminal region aa 267-285 is probably also involved in binding to choline residues of teichoic acids. (**C**) PrtA (AAK74791.1), the signal peptide sequence aa 1-27 is depicted in red, the serine protease catalytic domain is illustrated in yellow, the DUF 1034 domain is shown in red. (**D**) SFP (ABJ54257.1), the signal peptide sequence (1-22 aa) is shown in red, the serine protease catalytic domain is marked in blue. The C-terminal anchoring motif for PrtA and SFP is labeled in black. The length of each serine protease is given as the number of amino acids (aa). HDS: histidine, aspartate and serine; TM: transmembrane domain; SP: signal peptide sequences.

4.2 Serine proteases genetic organization and mutants generation

Pneumococcal serine proteases single knockouts for the genes *cbpG*, *htrA*, *prtA,* and *sfp* were generated by insertion-deletion mutagenesis.Previous studies focusing on the deletion of single pneumococcal serine proteases are difficult to interpret due to the compensatory effects of the other serine proteases. Therefore, to study the role of pneumococcal serine proteases *in vitro* and *in vivo*, the double and triple deficient *serine protease* mutants were generated. For example, mutants were generated in bioluminescent D39*lux*, TIGR4*lux*, and in strain *S.p.*35A, which produces lower amounts of CPS and is a model strain for the adherence and transmigration studies. Furthermore, the strain 19F_EF3030, which was isolated from a patient with otitis media[341, 342], was used to generate *serine protease* mutants for nasopharyngeal colonization and biofilm formation models. This strain 19F_EF3030 was kindly provided by Anders P. Håkansson (Lund University, Sweden). All of the mutants were verified by analytical PCR. As an example, a description of one of the generated triple mutants is presented in the following sections.

4.2.1 CbpG gene organization and molecular analysis

CbpG has a multifunctional role as an adhesive protein *via* binding to the epithelial cells and is involved in the bacteria virulence[100, 106].

The CbpG gene region organization was analyzed *in silico* in strain TIGR4 (*sp_0390*) (AAK74556.1) and 19F (*EF3030_01920*) (QBF69943.1). The analysis of the pneumococcal target gene using the online sequence databases (NCBI and KEGG) showed that *cbpG* is highly conserved and identical in both strains. The *cbpG* gene comprises 858 nt in TIGR4 or 822 nt in 19F_EF3030 wild-type. Downstream to *cbpG*, another gene encoding choline-binding protein F (*sp_0391*) was found in both strains. Upstream of *cbpG*, the gene *sp_0389*, which i*s* annotated a hypothetical protein, and *sp_0386* and *sp_0387*, encoding a two-component system , are localized as described by Mann et al.[106]. All genes are highly homologous among pneumococci, indicating a conserved localization as shown in **Figure 4.2A**.

Furthermore, to investigate the role of CbpG in pneumococcal colonization and virulence, double and triple *serine protease* mutants with only active CbpG+ were generated. The *cbpG*-mutant was generated as described previously (Wüst et al.,[343] AG Hammerschmidt) by insertion-deletion mutagenesis. The insertion of the erythromycin (*ermB*) and spectinomycin (*add9*) gene cassettes are shown in **Figure 4.2B**. The homologous recombination of the flanking

regions with the corresponding regions in the genomic DNA of the parental strain led to the deletion of the *cbpG* gene (allelic replacement) in the genome of *S. pneumoniae*.

Figure 4.2: *S. pneumoniae* TIGR4 and 19F EF3030 *cbpG* genetic organization.
(**A**) Chromosomal localization of the *cbpG* gene (*sp_0390* and *EF3030_01920*) shows the same gene organization with 99.6% homology on a DNA level. Program BlastP was used for the homology search. Arrowheads indicate the orientation of the genes. The transcription termination signals are shown as loops. (**B**) Schematic model of Δ*cbpG::ermB* and Δ*cbpG::aad9* mutants constructed by insertion-deletion mutagenesis carrying the *erm* or *aad9* gene cassette.

The recombinant plasmid pAW1101 (pSP72DΔ*cbpG*::Spec[r]) constructed for *cbpG* mutagenesis (*see* **material and methods7.2.7**) was used to transform the different pneumococcal strains to generate double and triple serine protease deficient mutants in 19_EF3030, TIGR4*lux*, *S.p.*35A and D39*lux* strains. The recombinant clones were selected on blood agar plates with the appropriate antibiotic (Spectinomycin). The replacement of the wild-type sequence against the mutated region was confirmed via analytical PCR by amplifying the 5'- and 3'-flanking homologous regions of the corresponding gene (**Figure 4.3A** and **B**). Because the amplification products are different in size, two verification steps were used: first, PCR amplification with

the primer pairs P1467/P1422, and the second with the primer pairs P1467/P118 (**Figure 4.3A**). In a second step, the generated mutants were used to generate double mutants.
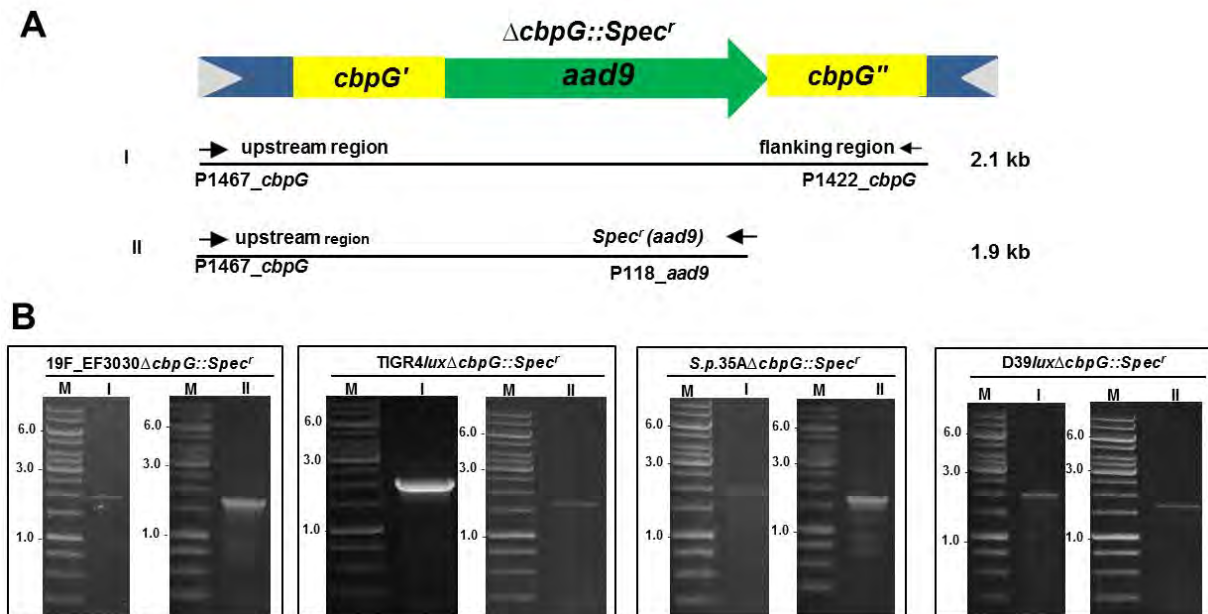
Figure 4.3: Molecular analysis of isogenic *cbpG*-mutant in *S. pneumoniae* strains.

(**A**) Schematic representation of the isogenic *cbpG*-mutant in *S. pneumoniae* and the expected fragments by PCR verification. The flanking chromosomal DNA sections are shown in dark blue. The 5′and 3′regions of the *cbpG* gene are shown in yellow. The green arrow (*aad9*) shows the location and orientation of the inserted antibiotic cassette. (**B**) Verification of *cbpG*-mutants in strain 19F_EF3030, TIGR4*lux*, *S.p.*35A, and D39*lux S. pneumoniae* strains by PCR analysis. **I:** PCR determination with the primer pairs 1467/1422; **II:** PCR with primer pairs 1467/118. Marker: Gene Ruler 1 kb DNA Ladder; *cbpG*�\' and *cbpG*ʺ the rest of the gene 5′and 3′region of *cbpG*; Spec^r, Spectinomycin; r, resistance.

4.2.2 Localization of the *htrA* gene and generation of double mutants

The pneumococcal HtrA belongs to the trypsin-like proteases and is encoded by *sp_2239* in TIGR4 (1182 nt) and *EF3030_11105* in 19F_EF3030 (1194 nt). The *htrA* gene is located upstream of the *spoJ* gene (*sp_2240*; 759 nt. and 252 aa) and encodes for a protein of unknown function, whereas the gene upstream of *htrA* (*sp_2238*; 480 nt and 159 aa) encodes for a encodes for an rRNA subunit methyl-transferase (**Figure 4.4A**). The plasmid pNM991 (pSP72DΔ*htrA*::Cm^r) was constructed for *htrA* mutagenesis (**Figure 4.4B**) as described previously by N. Mehlitz,[344] (Hammerschmidt lab).
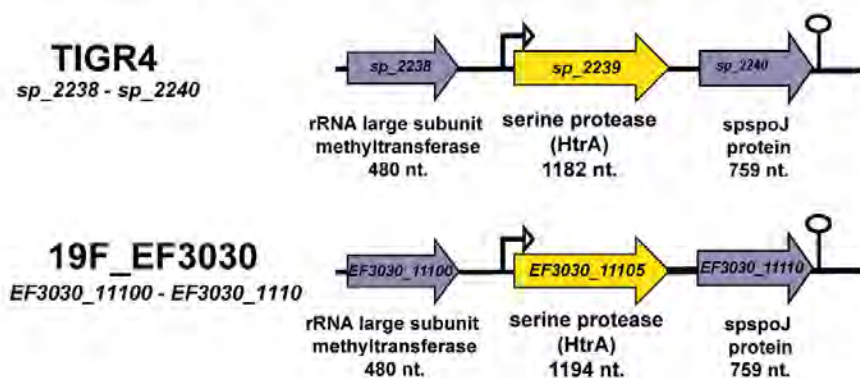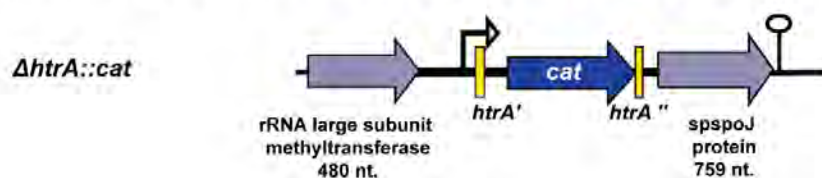
Figure 4.4: *S. pneumoniae htrA* gene organization in TIGR4 and 19F_EF3030.

(**A**) Chromosomal localization of the *htrA* gene (high-temperature requirement gene: *sp_2239* in TIGR4 and *EF3030_11105* in 19F). The gene upstream of *htrA* encodes for an rRNA subunit methyl-transferase. The downstream gene is annotated as ParB/RepB/Spo0J family partition protein. Arrowheads indicate the orientation, the predicted putative promoters (black arrowheads) and the transcription termination signals are shown as loops. (**B**) Schematic model of *htrA*-mutants constructed by insertion-deletion mutagenesis in TIGR4 and 19F shows the deletion of an internal *htrA* sequence and the replacement by the *cat* gene cassette inserted in the *htrA* gene sequence.

To study the role of HtrA in double and triple serine proteases deletion mutants, the constructed plasmid pNM991 (pSP72DΔ*htrA*::Cm^r) was used to transform different mutated *S. pneumoniae* strains. The generated single mutant with *cbpG* deletion in 19_EF3030, TIGR4*lux*, *S.p.*35A and D39*lux* strains (**Figure 4.3**) were transformed with pNM991 to receive the double mutants. Afterward, the generated <u>*htrA*</u> mutation was verified by PCR using different primer combinations (**Figure 4.5A**). The PCR amplifications were performed with the primer pair P1089/P158 and primer pair P1088/P159 and the expected 1,9 kb amplicons were observed (**Figure 4.5B**).
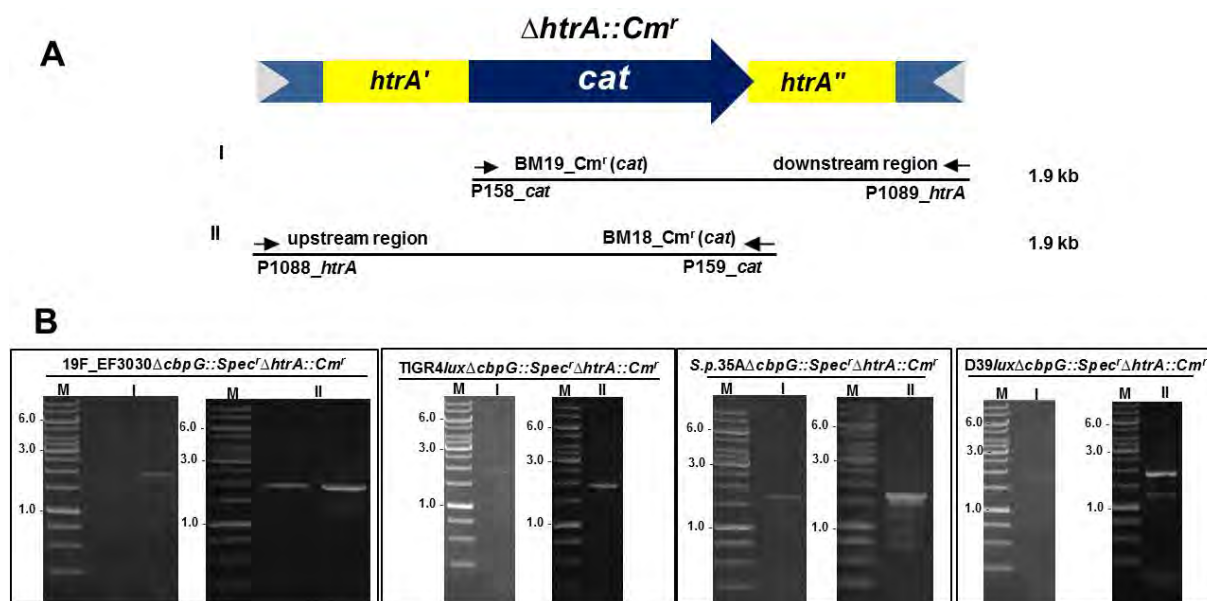
Figure 4.5: Double mutant molecular analysis of isogenic *htrA*-mutant in *S. pneumoniae* strains.

(**A**) Schematic representation of the *htrA*-mutant in *S. pneumoniae* and the expected fragments by PCR verification. The flanking chromosomal DNA sections are shown in dark blue. The regions of the *htrA* gene are shown in yellow. The dark blue arrow Cm^r (*cat*) shows the location and orientation of the inserted antibiotic cassette. (**B**) Verification of *htrA*-mutants in 19F_EF3030, TIGR4*lux*, *S.p.*35A, and D39*lux S. pneumoniae* strains by PCR analysis. **I:** PCR determination with the primer pairs P158/1089; **II:** PCR with primer pairs P1088/159. Marker: Gene Ruler 1 kb DNA Ladder; *htrA*' and *htrA*" rest of the gene 5´and 3´region of *htrA*; Cm^r, Chloramphenicol; R, resistance.

4.2.3 PrtA gene organization and triple serine protease mutant generation

The cell wall-associated serine protease A (PrtA) is a highly conserved virulence factor in pneumococci and can be found in almost all strains. The *prtA* gene is encoded by *sp_0641*, 6423 nt (AAK74791.1) in TIGR4 and by *EF3030_03025*, 6420 nt (QBF68585.1) in 19F_EF3030. The gene organization was analyzed and compared between strain TIGR4 and 19F as shown in **Figure 4.6A**. Up-and downstream of the *prtA* gene are different genes localized in the two genomes. In TIGR4, the upstream gene is transcribed in the same orientation as *prtA,* whereas the downstream gene is transcribed counterclockwise. In strain 19F_EF3030, the upstream gene is transcribed counterclockwise to *prtA* and the downstream located gene is transcribed clockwise.

The recombinant plasmid with the mutated *prtA* gene was constructed using vector pUC18 where part of *prtA* is deleted and placed by the *ermB* antibiotic resistance gene cassette to generate pGB1019 (pUC18Δ*prtA*::Erm^r) **Figure 4.6B**.
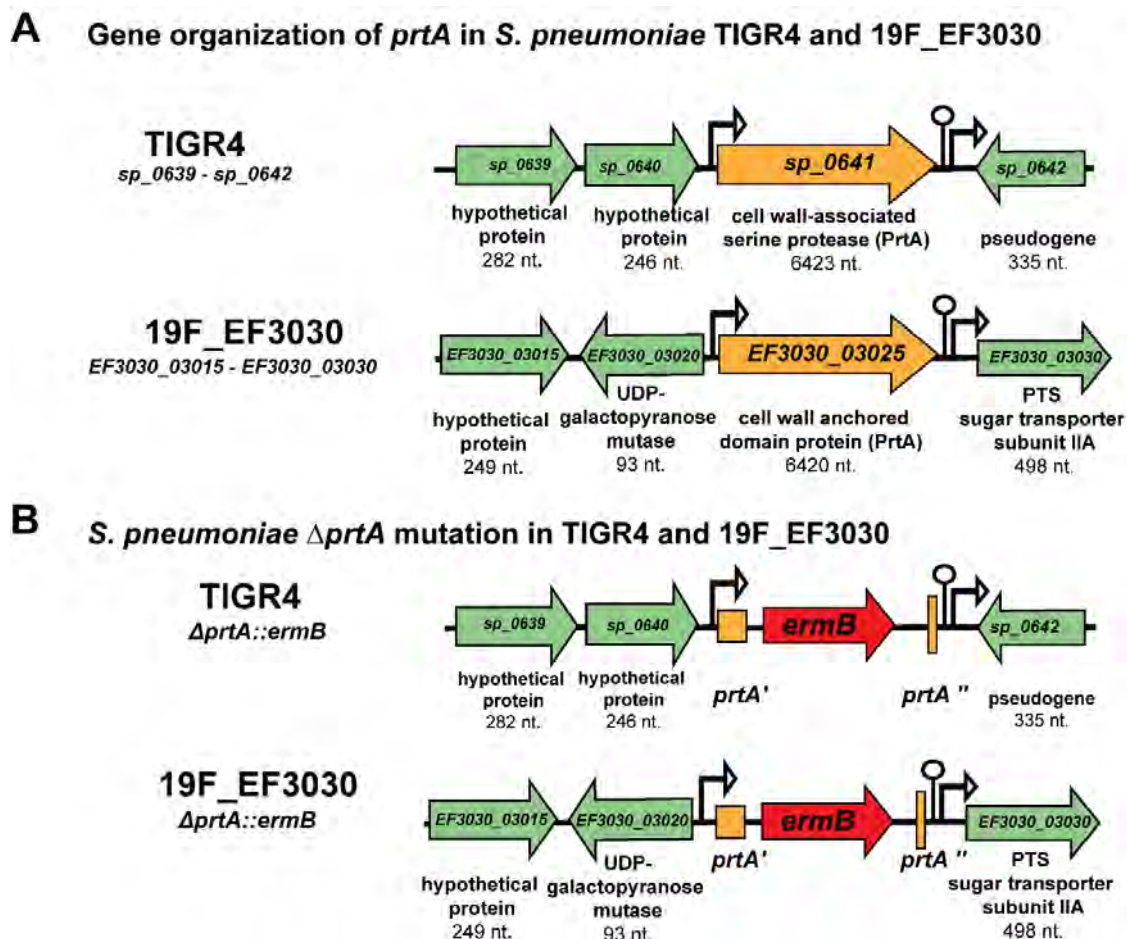
Figure 4.6: Genetic organization of *prtA* in *S. pneumoniae* iTIGR4 and 19F_EF3030.

(**A**) Chromosomal localization of the *prtA* gene encoding for the cell wall-associated serine protease (*sp_0641, EF3030_03025*). In TIGR4, the gene downstream of *prtA* is a pseudogene, while in 19F*_EF3030,* the gene encodes for a PTS sugar transporter subunit. The genes upstream of *prtA* (*sp_0640,* 246 nt and *sp_0639,* 282 nt) in TIGR4 encode for hypothetical proteins. 19F_EF3030 shows different upstream genes compared to TIGR4; these genes are *EF3030_03015* and *EF3030_03020*, which encode for a hypothetical protein or UDP-galactopyranose mutase. Arrowheads indicate the gene's orientation, the predicted putative promoters (black arrowheads), and the transcription termination signals are shown as loops.

 (**B**) Schematic model of *prtA*-mutants constructed by insertion-deletion mutagenesis in TIGR4 and 19F_EF3030 shows the deletion of an internal part of *prtA* and insertion of the *ermB* gene cassette in the *prtA* gene sequence.

The constructed plasmid pGB1019 was transformed into pneumococcal serine protease double mutants to generate the triple serine protease mutation in strains TIGR4, D39, and *S.p.* 35A. Verification of the *prtA* mutation was carried out by colony PCR with the primer combination P99/P1139 to amplify the expected 1.9 kb fragment. As control primer, P99/P100 amplified only the erm*B* gene cassette of 1.1 kb. The PCR amplification of genomic DNA from mutant strains was performed with two specific primer combinations: primer pair P99/P1139 and the second with P99/P100 for the erythromycin resistance gene cassette (**Figure 4.7**).
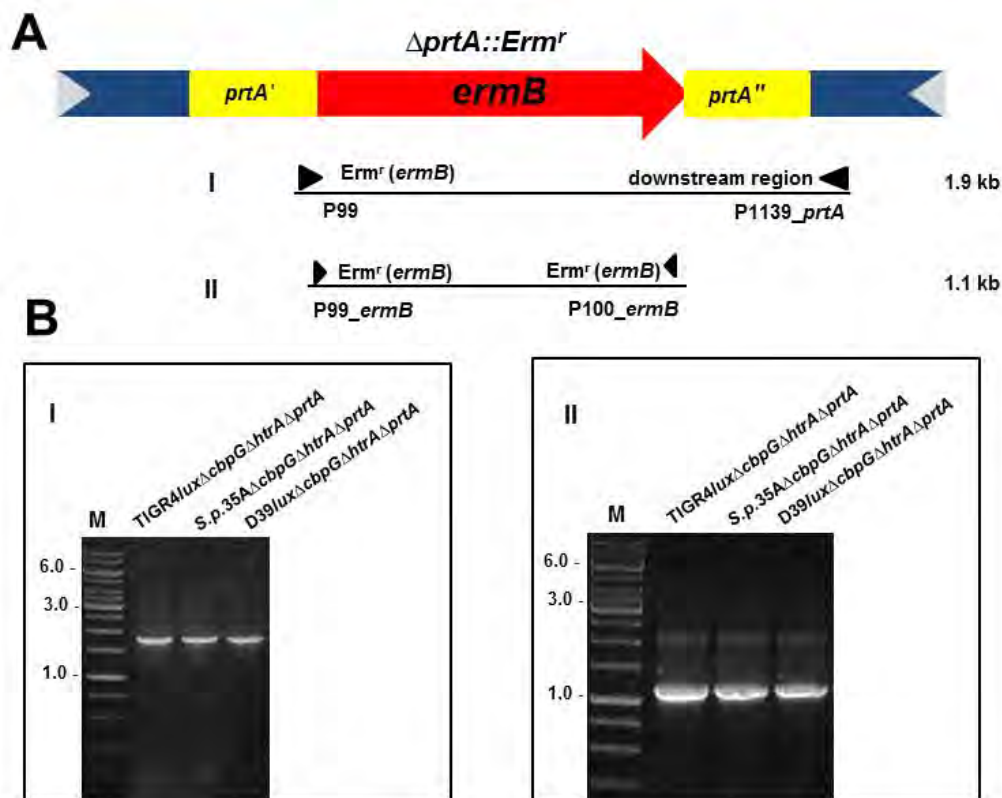
Figure 4.7: Verification of triple isogenic mutant in *S. pneumoniae* strains. (**A**), Schematic representation of the isogenic *prtA*-mutant in *S. pneumoniae* and the expected fragments by PCR verification. The flanking chromosomal DNA sections are shown in dark blue. The flanking regions of the *prtA* gene are shown in yellow. The dark blue arrow Erm^r (*ermB*) shows the location of the inserted antibiotic cassette. (**B**), Verification of *prtA* mutants in TIGR4*lux*, *S.p.*35A, and D39*lux S. pneumoniae* strains by PCR analysis. **I:** PCR determination with the primer pair P99/1139; **II:** PCR with primer pair P99/100. Marker: Gene Ruler 1 kb DNA Ladder; *prtA'* and *prtA"* rest of the gene 5´and 3´region of *prtA*; Erm^r, Erythromycin.

4.2.4 Gene organization of SFP protease and molecular analysis

The first characterization of SFP was published 2013 by de Stoppelaar et al.[234]. The *sfp* gene is annotated in strain TIGR4 as *sp_1954*, and contains 1404 nt (467 aa), with a calculated molecular mass of 52 kDa.

The comparative analyses of the *sfp* gene region was analyzed *in silico*, in strain TIGR4 and 19F_EF3030. Interestingly, the genomic organization of the *sfp* locus between 19F_EF3030 and TIGR4 strains in *S. pneumoniae* are different. The *sfp* gene and six upstream and three downstream genes are not present in the whole genome of the 19F strain EF3030 (NCBI, CP035897.1), as demonstrated in **Figure 4.8A**. Furthermore, another comparison between *sfp* genes in *S. pneumoniae* strain D39 *spd_1753* (1740 nt, 579 aa), and TIGR4 *sp_1954* (1404 nt, 467 aa) was performed using the SYBIL database[345]. The multiple sequence alignment (MSA) analyses showed a shorter version of the *sfp* gene in TIGR4 compared to *sfp* of D39 and other strains (*see* **appendix 178**). The truncation of *sfp* in strain TIGR4 is based on the deletion of one base (A) at position 1381. Instead of 8 A bases in a row, only seven are present in strain TIGR4, which was confirmed by DNA sequencing of the TIGR4 *sfp* gene. This frameshift leads to a premature stop at position 1404. Hence, this truncated SFP of TIGR4 cannot be covalently anchored to the peptidoglycan. Instead, SFP of TIGR4 is most likely secreted into the extracellular environment. In D39 and other pneumococci, the SFP protein contains an LPNTS motif, which is a target site for the sortase A, which covalently anchors SFP to the peptidoglycan.

The *sfp*-mutant was generated to get the triple serine protease deletion mutant. The recombinant plasmid was designed with three different antibiotic resistance gene cassettes *ermB*, *aad9*, or *cat*. The recombinant plasmids used for *sfp* mutation were pRB1119 (pSP72DΔ*sfp*::Erm^r), pRB1132 (pSP72DΔ*sfp*::Spec^r), and pRB1131 (pSP72DΔ*sfp*::Cm^r), shown in **Figure 4.8B**.

Figure 4.8: *sfp* gene organization in *S. pneumoniae* TIGR4 and EF3030.

(**A**) Comparison of genome regions in TIGR4 (*sp_1946-sp_1959*) with 19F (*EF3030_09485-EF3030_09505*). The *sfp* gene in TIGR4 *sp_1954* is shown as a red arrow with 1404 nt, while in 19F_EF3030, the *sfp* gene and six genes upstream are absent. The homologous analysis was done (Clustal Omega tool) to check whether the *sfp* gene is located in a different gene locus in 19F. The genes *sp_1957* and *sp_1959* (light green) encode for an ABC transport system (ATP-binding protein) and a nucleoside-diphosphate kinase were identical to *EF3030_09500* and *EF3030_09505*, respectively. Therefore, three genes downstream of *sfp* are missing in EF3030. Arrowheads indicate the orientation of the gene, the predicted putative promoters are shown as black arrowheads, and the transcription termination signals are shown as loops. (**B**) Schematic model of *sfp*-mutants constructed by insertion-deletion mutagenesis shows the *erm*, *aad9* or *cat* gene cassettes inserted into the *sfp* gene.

The *sfp* mutation was verified by colony PCR with the primer combination P1416/P181 to check for the expected amplicon of 1.68 kb. As a control the primers of the ab cassette revealed the size of 1.1 kb (**Figure 4.9B**).
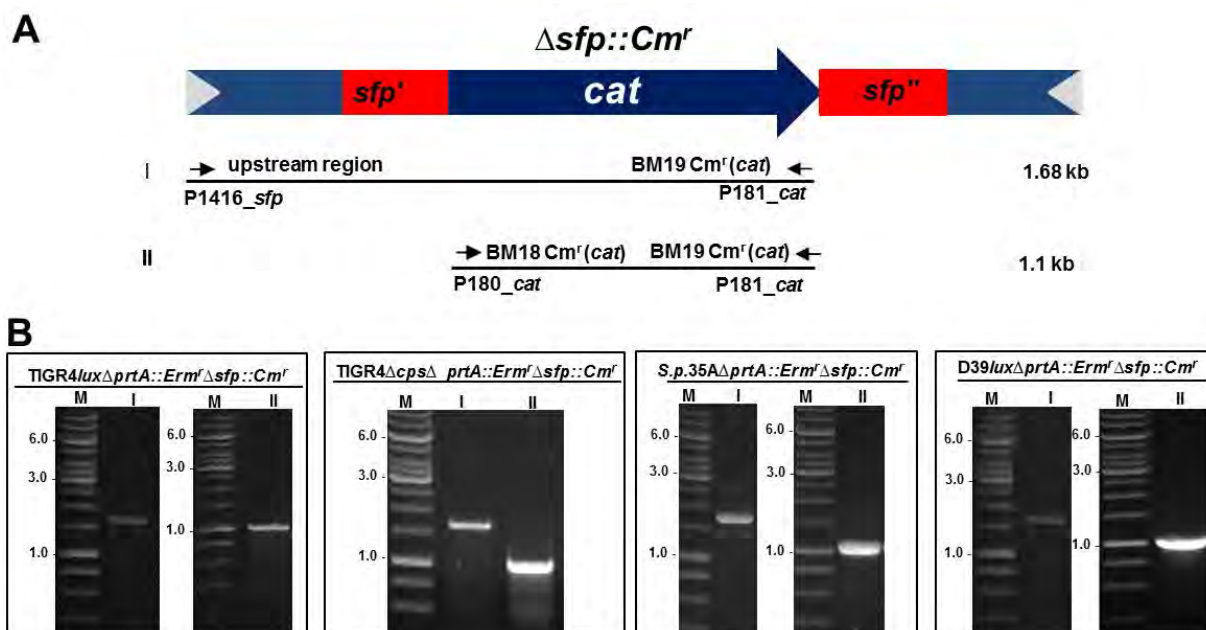
Figure 4.9: Molecular analysis of isogenic *sfp*-mutant in *S. pneumoniae* strains.

(**A**) Schematic representation of the *sfp*-mutant in *S. pneumoniae* and the expected fragments revealed by PCR amplification. The flanking chromosomal DNA sections are shown in dark blue. The regions of the *sfp* gene are shown in red. The dark blue arrow represents the chloramphenicol resistance Cm^r (*cat*). (**B**) Verification of *sfp* deletion in TIGR4*lux*Δ*prtA*, *S.p.*35AΔ*prtA*, and D39*lux*Δ*prtA S. pneumoniae* strains by colony PCR analysis. I: PCR with the primer pair P1416/181; II: PCR with primer pair P180/181. Marker: Gene Ruler 1 kb DNA Ladder; *sfp*', upstream of *sfp* area; *sfp*", downstream of *sfp* area; Cm^r, chloramphenicol.

4.3 Impact of pneumococcal serine proteases on bacterial fitness and growth behavior

Two *S. pneumoniae* strains, TIGR4 and EF3030 wild-types as well as isogenic *serine protease* mutants were studied regarding their growth and bacterial fitness to characterize the mutants. Bacterial growth was monitored in a chemically defined RPMI$_{modi}$ or complex THY medium. The growth behavior was monitored for up to 10h by measuring the optical density (OD$_{600nm}$). In addition, the generation times of the wild-type and mutants were calculated under the growth conditions (*see* **7.1.6 and Supplementary Table 8.5**).

Interestingly, the growth of 19F_EF3030 wild-type and the isogenic *serine proteases* mutants showed no growth difference in complex THY medium. Only the mutant with CbpG+ as the only functional serine protease started to lyse shortly after 7h of growth (**Figure 4.10A**). Therefore, statistically significant differences were only monitored in the late stationary phase for the CbpG+ mutants at time points 8, 9, and 10h of growth compared to the wild-type. Growth in a chemically-defined medium of mutants and wild-type of strain 19F_EF3030 was similar with an extended generation time compared to the complex THY medium. The lowest growth rate was observed with the *cbpG* positive mutant. In addition, the CbpG+ and the HtrA+ mutant had an extended lag phase. However, these mutants reached a similar optical density compared to the wild-type. No significant differences in the growth of the wild-type and isogenic mutants were observed in RPMI$_{modi}$ (**Figure 4.10B**).

In comparison to 19F_EF3030, the TIGR4Δ*cps* strain and its isogenic *serine protease* mutants showed similar growth behavior in complex and chemically defined medium. The strain TIGR4 and the isogenic *serine protease* mutants showed also a similar growth behavior in THY medium (**Figure 4.10C**). After five hours, the optical density started to decrease in all cultures, indicating lysis of bacteria. No differences between the wild-type and the isogenic mutants were observed when grown in the minimal medium RPMI$_{Modi}$, as shown in **Figure 4.10D**. In contrast to strain 19F, the TIGR4 wild-type and the isogenic mutants reached an optical density of 1.3, whereas, for strain 19F, only an optical density of 0.3 could be determined.

The encapsulated TIGR4*lux* and isogenic *serine protease* mutants were cultured in THY and RPMI medium. The mutants expressing CbpG+ or SFP+ showed a reduced growth rate compared to the parental strain. In contrast to the TIGR4Δ*cps* mutants, an extended stationary growth was determined for the encapsulated TIGR4*lux* and derivatives. The mutants entered the stationary phase already at a lower optical density (**Figure 4.10E**). In the RPMI medium, moderate growth of the wild-type and the isogenic mutants was observed, reaching an optical

density up to 0.4 as shown in **Figure 4.10F**, which is remarkably less compared to the non-encapsulated TIGR4 strain (**Figure 4.10D**).
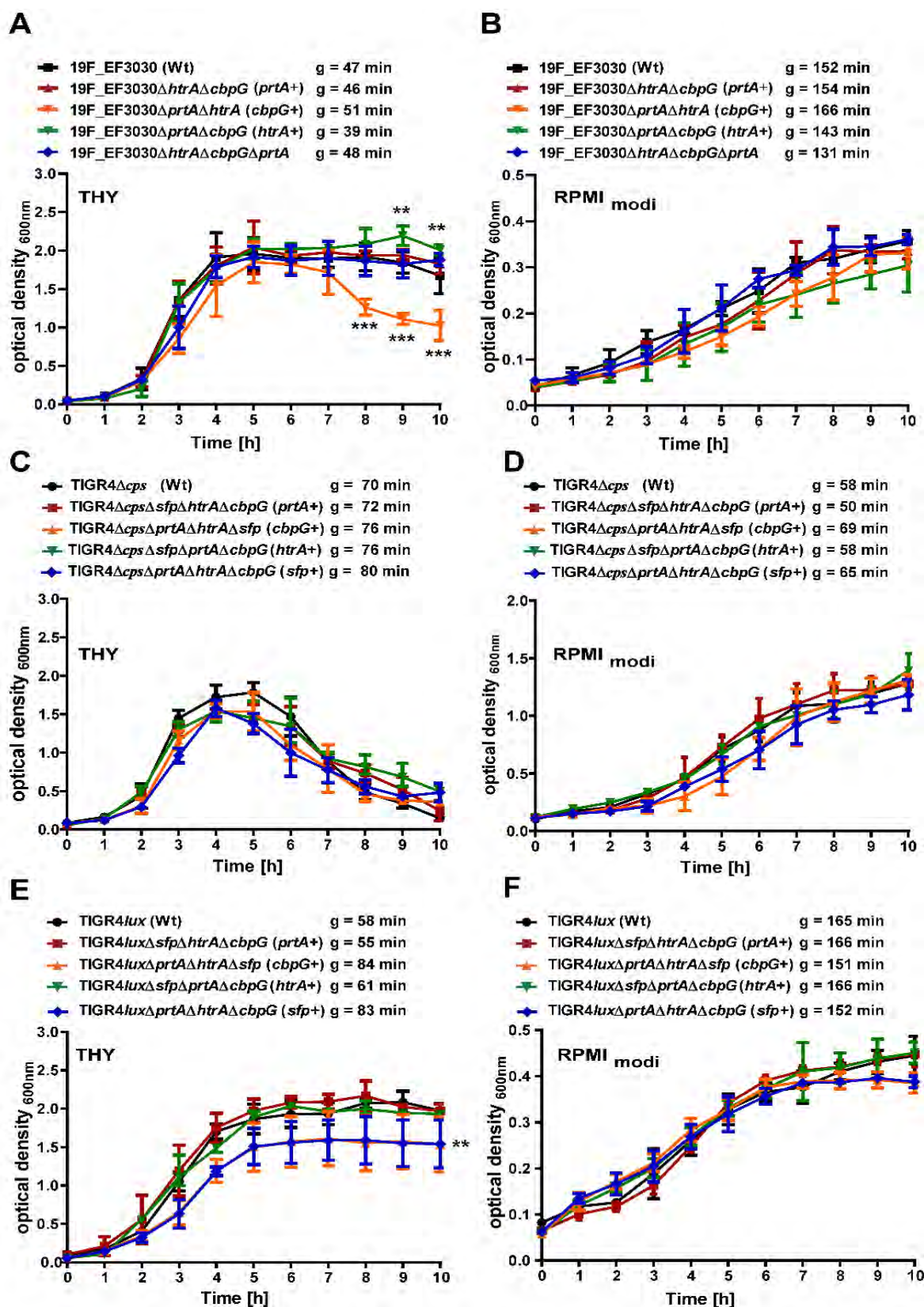
Figure 4.10: Growth behavior of *S. pneumoniae* wild-type strains and isogenic triple *serine protease* mutants.

The wild-type and isogenic mutants were cultured at 37°C in a complex THY or chemically defined medium (RPMI$_{modi}$). Pneumococcal growth was monitored for up to 10h by measuring the optical density (OD$_{600nm}$). The mean of four individual growth experiments is shown for 19F_EF3030 (**A and B**), TIGR4Δ*cps* (**C, and D**), and TIGR4*lux* (**E, and F**). Error bars represent SD (n = 4). The symbol "g" indicates the generation time, calculated from four biological replicates. The data were statistically analyzed using a two-way ANOVA analysis *P < 0.05.*

The different growth behavior of the parental 19F strain E3030 and their single isogenic mutants were further investigated in complex TYH and RPMI$_{modi}$ medium. In THY, growth reached a maximum optical density of 2.5 and after 6h of cultivation, the optical density decreased below 2.0. No growth differences were observed in comparison to the wild-type (**Figure 4.11A**). In the RPMI$_{modi}$ medium, similar growth of the single mutants was observed compared to the wild-type (**Figure 4.11B**).

Finally, to determine whether the deletion of serine proteases have an effect on nutrient acquisition, growth of strains 19F_EF3030, TIGR4 and isogenic mutants expressing no serine protease (19F) or only SFP, were performed in CDM[346, 347]. The CDM medium was supplemented with 2% casein hydrolysate as the primary source of amino acids and peptides (*see* **Table 6.6** and **Table 6.7**) as described earlier by Härtel et al.,[348]. Indeed, an increase in the optical density up to 0.5 after 6 h of cultivation was observed. The same medium was also used to compare the growth of the encapsulated TIGR4*lux* strain and the triple but SFP+ mutant. Both, the wild-type and the triple serine protease mutant showed an extended lag phase and reached the maximum growth with an optical density of 3.0 after 10 h of cultivation, which is higher compared to THY medium. No significant growth differences were observed between the mutants and the corresponding parental strains (**Figure 4.11C and D**). These results indicated that serine proteases play no role in the uptake of nutrients under the tested conditions. Similarly, the lack of a single serine protease would not affect growth at all. Growth rates are listed in **Supplementary Table 8.5**.
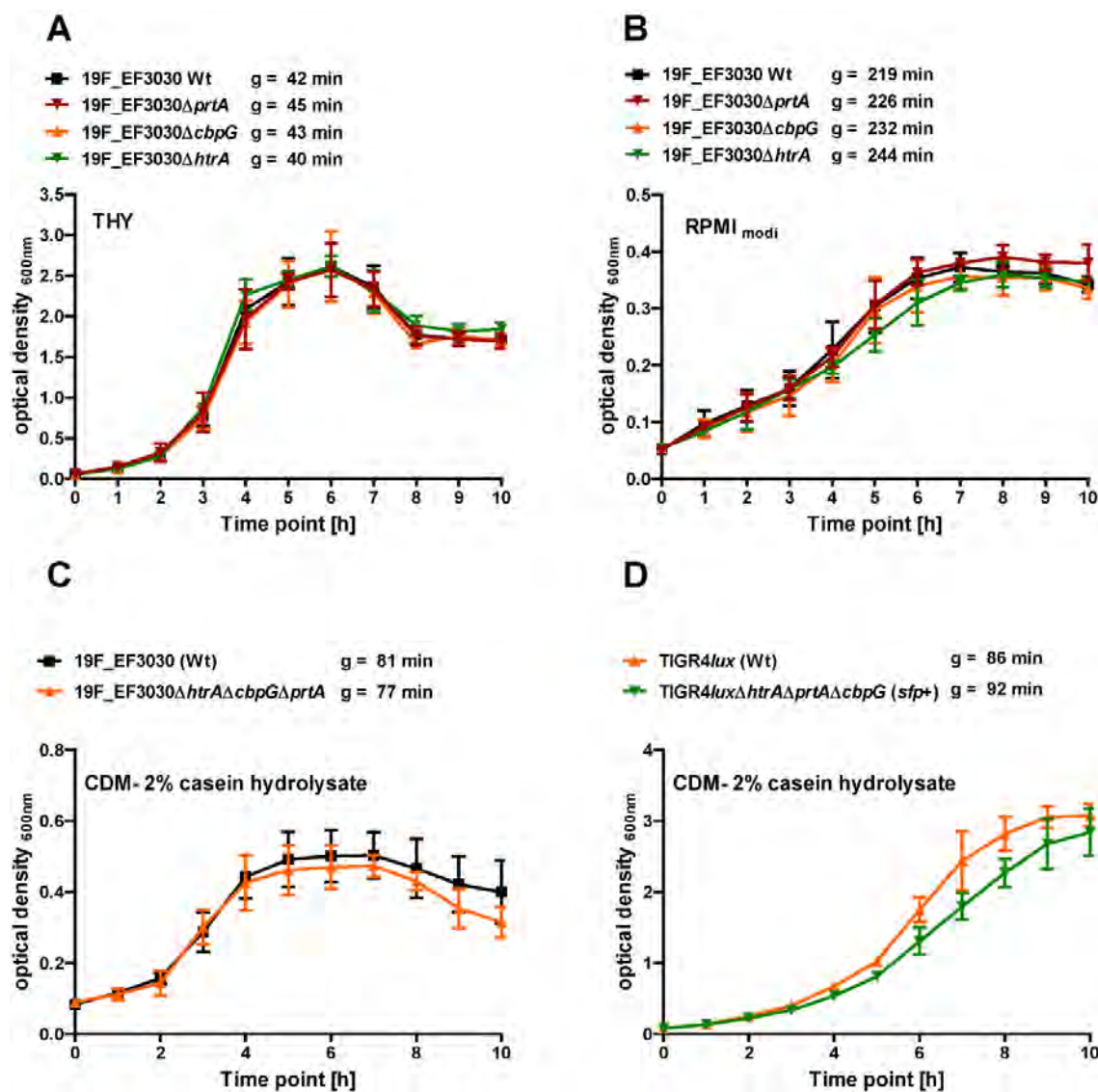
Figure 4.11: Growth behavior of *S. pneumoniae* wild-type and their isogenic single *serine protease* mutants.

(**A, B**) *S. pneumoniae* 19F_EF3030 wild-type and single isogenic mutants were cultured at 37°C in THY and a chemically defined medium (RPMI$_{modi}$). (**C, D**) 19F_EF3030 and the isogenic mutant lacking all *serine proteases* as well as TIGR4 wild-type and its SFP positive triple mutant were grown in CDM without the addition of the essential amino acid solution I and II, supplemented instead with 2% casein hydrolysate. Error bars represent SD (n=4). The symbol "g" indicates the generation time, calculated from four biological replicates. The data were statistically analyzed using a two-way ANOVA analysis.

4.4 Impact of serine proteases on host-pathogen interaction

Host-pathogen interactions of serine proteases were characterized in *in vitro* cell culture assays such as adherence and phagocytosis experiments, respectively.

4.4.1 Serine proteases deficiency decreased pneumococcal adherence to human nasopharyngeal epithelial cells

The specific adherence of pneumococci to epithelial cells of the upper respiratory tract (URT), which facilitates pneumococcal colonization, is an essential step in pneumococcal infection[13]. Therefore, the role of extracellular serine proteases in adherence to host epithelial cells was investigated using the human nasopharyngeal epithelial cell line Detroit-562. The adherence to cellular receptors is impaired by capsule expression of the bacteria[76, 77], and non-encapsulated TIGR4Δ*cps* showed a much stronger *in vitro* adherence to epithelial cells compared to the encapsulated TIGR4 strain[349]. Thus, the impact of serine proteases on pneumococcal adherence was studied by infecting Detroit-562 cells with TIGR4Δ*cps* (serotype 4) or 19F EF3030. Detroit-562 cells were infected with wild-type or isogenic *serine protease* mutants with a multiplicity of 50 (MOI 50) for 4 h.

The wild-type bacteria adhered to Detroit-562 was set to 100%. The 19F strain with the serine protease deletions showed thereby only 8-13% of adhered bacteria in comparison to the wild-type. The lowest counts were observed for the mutant with no functional serine protease. Therefore, double *serine protease* mutants expressing only one functional serine protease (CbpG+, PrtA+ or SFP+) or a mutant lacking all serine proteases revealed a significant reduction of 19F adherence to Detroit-562 cells in comparison to the parental strain (*P*<0.05, and *P*<0.01) (**Figure 4.12A**) and (**Supplementary Table 8.6**) in the appendix.

Similarly, the TIGR4Δ*cps* triple *serine protease* mutants with only a single functional protease showed a substantial reduction of adherent bacteria compared to the parental strain (*P*<0.05). Only the PrtA+ positive triple mutant reached up to 50% of adhered pneumococci compared to the wild-type. The lowest amount of 17% adhered pneumococci was estimated for the TIGR4Δ*cps* mutant with only functional SFP (**Figure 4.12B**). The reduction of pneumococci with deficiency for two-or three serine proteases is in accordance with the immunofluorescence microscopy images (**Figure 4.12C**). Taken together, these data confirm the substantial role of extracellular serine proteases in adherence to epithelial cells.
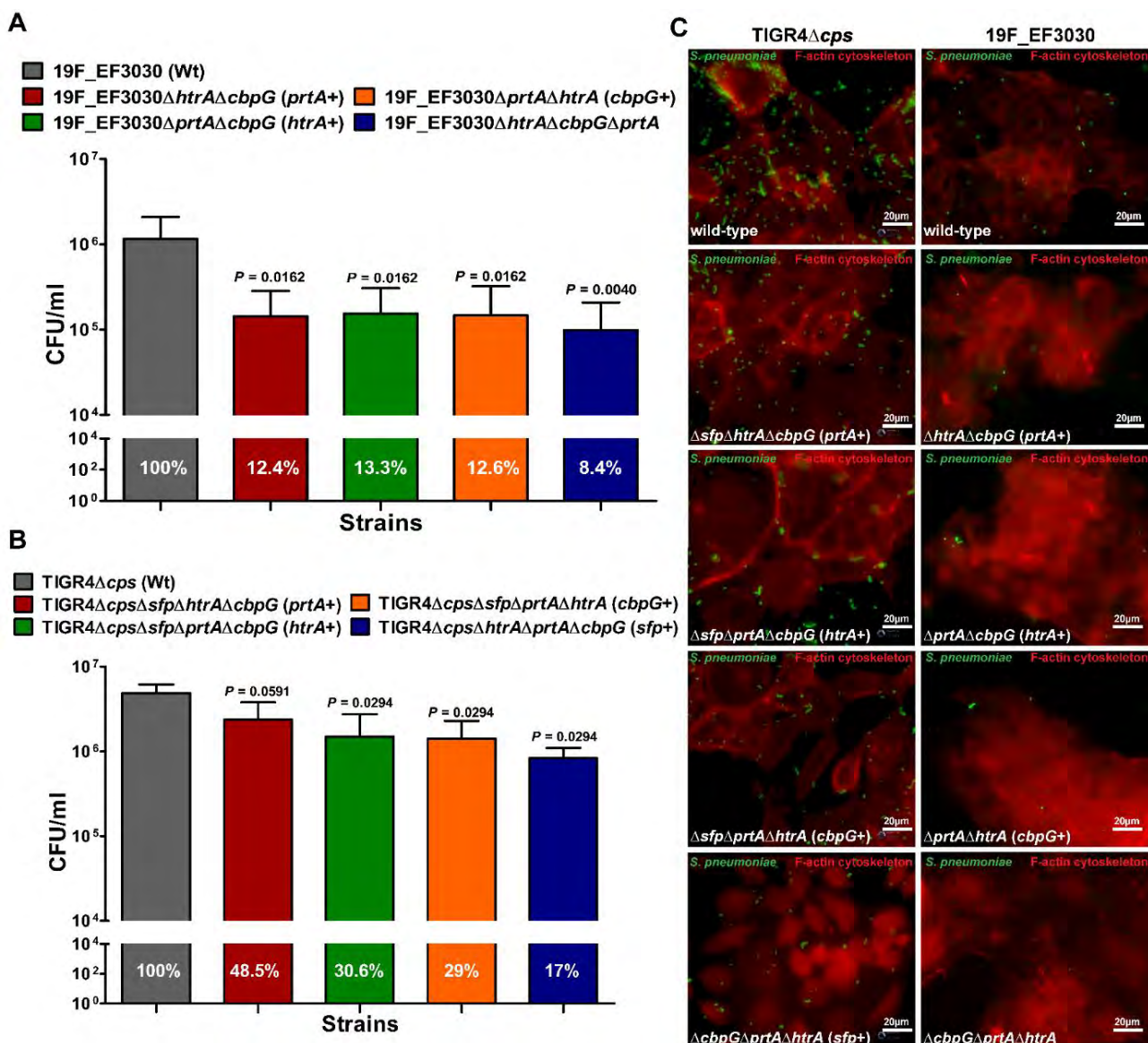
Figure 4.12: Pneumococcal serine proteases affect pneumococcal adherences to host epithelial cells.

(**A, B**) Pneumococcal adherence to human Detroit-562 cells was quantified after 4h of infection with an MOI of 50 pneumococci of 19E_EF3030, or TIGR4Δ*cps* wild-type and isogenic mutant strains. Colony-forming units were determined post-infection by plating host cell adherent pneumococci on blood agar plates. Results are presented as the mean ± SD for at least four independent experiments, performed in triplicates. *n.s.* **P < 0.05*, and ** *P ≤ 0.01* relative to the parental 19F_EF3030 pneumococcal strain. (**C**) Immunofluorescence microscopy of pneumococci attached to Detroit-562 cells after 4h post-infection. Adherent pneumococci were visualized using anti-pneumococcal antiserum followed by a secondary Alexa-488 conjugated anti-IgG antibody (green). The epithelial F-actin was stained with Phalloidin-iFlour-594 conjugate (red). Each bar in the images represents 20μm.

4.4.2 Impact of serine proteases on pneumococcal phagocytosis by murine macrophages

Phagocytosis experiments with murine macrophages (J774A.1) were conducted to investigate whether the extracellular serine proteases protect pneumococci against phagocytosis by macrophages or improve the intercellular survival. Therefore, we characterized the contribution of pneumococcal serine proteases on intracellular survival in professional murine phagocytes by using the invasive non-encapsulated strain TIGR4Δcps. Murine macrophage J744A.1 cells (2×10^5 cells/well) were infected with wild-type or isogenic triple *serine protease* mutants for 30 minutes at 37°C with an MOI of 50 pneumococci per cell. Subsequently, extracellular pneumococci were killed by 1h of antibiotic treatment. The number of intracellular living pneumococci was determined by plating the permeabilized macrophages on blood agar plates. In parallel, double immunofluorescence microscopy was conducted to visualize and identify internalized pneumococci.

Interestingly, the number of recovered *serine protease* mutants where only PrtA+, SFP+ and HtrA+ functional were significantly decreased ($P<0.05$) with up to 26-30% in comparison to the parental wild-type (100%). Contrarily, the mutant with only CbpG+ function (93%) did not show a significant difference compared to the wild-type strain (**Figure 4.13A**). In addition, the immunofluorescence microscopy showed a remarkable reduction of internalized pneumococci when only SFP+ and HtrA+ serine protease is functional (**Figure 4.13B**). Only 28-36% of the SFP+ and HtrA+ mutant were recovered whereas up to 54-71% of intracellular TIGR4 mutants (PrtA+ and CbpG+) were determined. This result point to a more important impact of SFP and HtrA in preventing the uptake of pneumococci by macrophages. Importantly, immunofluorescence images that show the internalized pneumococci in **Figure 4.13C** correlate with the calculated percentage of intracellular determined pneumococci. The results of the antibiotic protection assay and double immunofluorescence confirmed that loss of serine proteases has a significant effect on pneumococcal uptake by professional phagocytes.
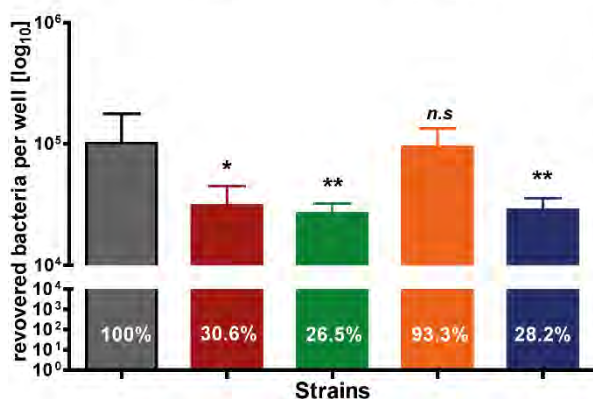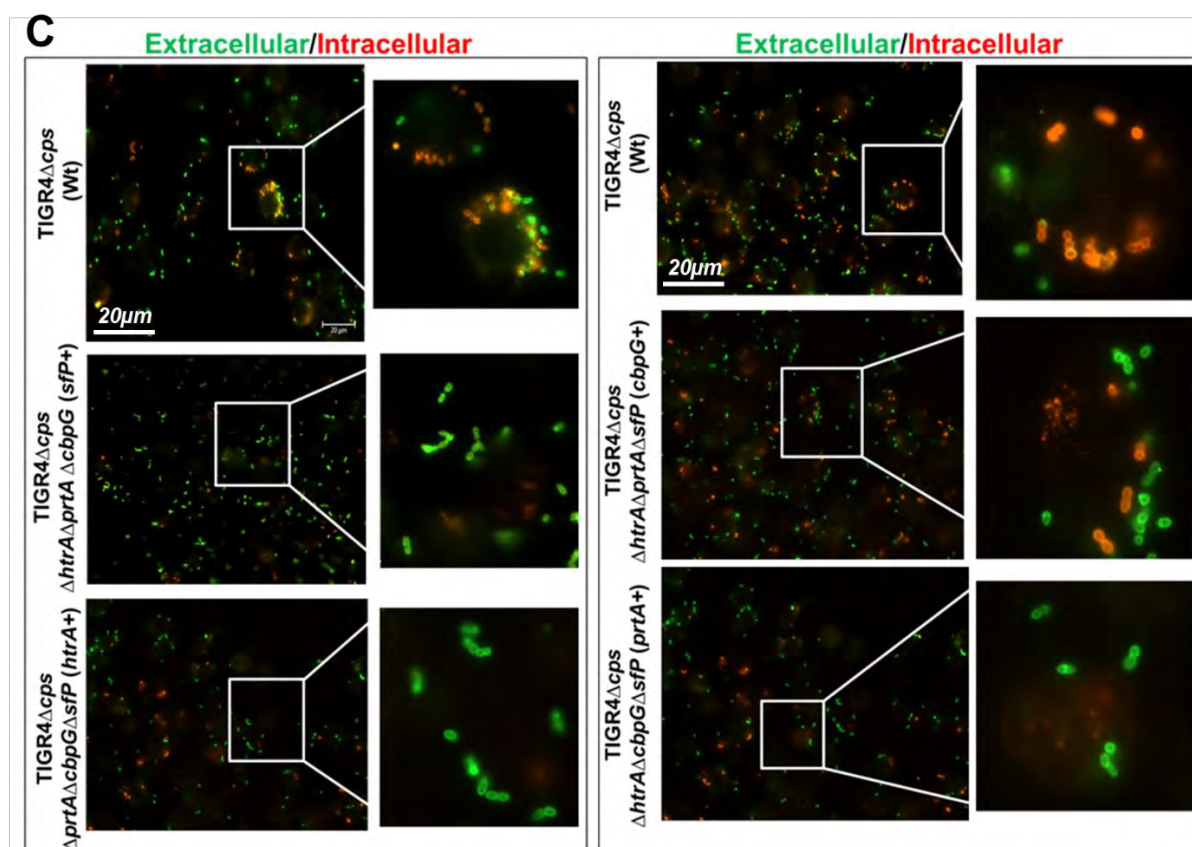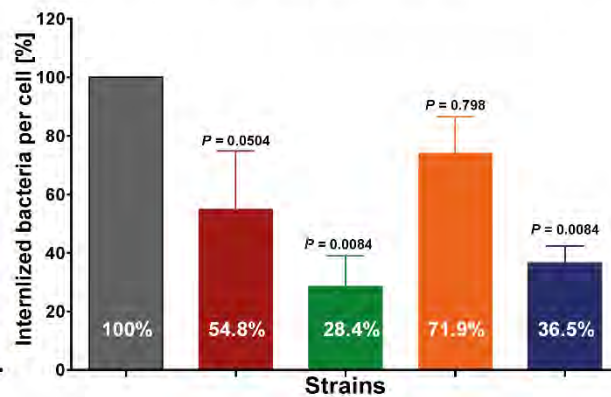
Figure 4.13: Effect of serine proteases on phagocytosis by professional murine macrophages.

Murine macrophages J774A.1 were infected with TIGR4Δ*cps* or isogenic triple mutants with a MOI of 50 bacteria per cell for 30 minutes at 37°C. (**A**) The recovery of intracellular pneumococci was determined by plating the bacteria from permeabilized macrophages on blood agar plates. Extracellular pneumococci were killed *via* gentamicin and penicillin treatment. The results represent the mean ± SD of four independent experiments performed in triplicate. (**B**)

The percentage of intracellular pneumococci was determined by immunofluorescence microscopy. At least pneumococci from 50 phagocytes were counted. The number of intracellular wild-type pneumococci was set to 100%, (n≥4, *P<0.05; Mann- Whitney U-test). (**C**) Immunofluorescence microscopy of pneumococci attached (green) to J774A.1 and phagocytized intracellular pneumococci (red) after 30 min post-infection. Extracellular bacteria were stained with anti-pneumococcal antiserum followed by secondary Alexa-488 conjugated anti-IgG antibody (green), while intracellular bacteria were stained with Alexa-568 conjugated anti-IgG antibody (red). Each bar in the images represents 20µm.

4.5 Impact of serine proteases on colonization, pneumonia and invasive infection

The influence of serine proteases on nasopharyngeal colonization and pneumonia was analyzed *in vivo* by applying two different mouse infection models, namely the nasopharyngeal colonization model and the acute pneumonia model. These *in vivo* infection models are well established in our laboratory.

4.5.1 Serine proteases have only a moderate effect on *S. pneumoniae* TIGR4 virulence in an acute pneumonia model

Pneumococci can colonize the upper respiratory tract, subsequently spreading to lower parts of the respiratory track causing pneumonia or sepsis[350]. It has been hypothesized that the development of an acute lung infection would be prevented or attenuated in the absence of serine proteases. Therefore, the murine acute pneumonia model was used to investigate the role of serine protease deficiency on pneumococcal virulence. CD-1 outbred mice (14 mice/group) were intranasally infected with 9×10^7 bioluminescent TIGR4*lux* or corresponding isogenic triple *serine protease* mutants expressing only one out of four functional serine proteases. The influence of serine proteases on pneumococcal dissemination into the lungs and bloodstream was monitored *in vivo* via the IVIS®Spectrum bioimaging system (Caliper, USA). Already 24h post-infection, mice infected with the parental TIGR4*lux* strain showed first signs of bacterial infection in the lungs.

In contrast, the triple mutants, lacking the expression of three out of four serine proteases, developed first signs of the lung infection after 40h post-infection (**Figure 4.14A, and B**). Furthermore, all mice infected with serine proteases mutants expressing only CbpG+, SFP+, or PrtA+ showed a substantial reduction in the bioluminescent intensity (**Figure 4.14C and D**),

which correlates with the mount of bacteria. On the other hand, in mice infected with the mutant expressing only HtrA+, the bioluminescent flux did not show a significant difference compared to the parental strain, as demonstrated in **Figure 4.14D**.

Importantly, the survival rate of mice infected with triple serine proteases mutants expressing only one functional serine protease (CbpG+, SFP+, and HtrA+) showed no significant differences compared to the wild-type except for the PrtA+ mutant **Figure 4.14 E and F**. The TIGR4*lux* PrtA+ mutant had a significantly prolonged survival rate (*p-value* 0.0414) compared to the parental strain. Bioimaging revealed that only one out of fourteen mice developed severe pneumonia 64h post-infection (**Figure 4.15F**). Interestingly, the results of the real-time bioimaging with the TIGR4 mutant expressing only PrtA+ showing the lowest bioluminescence correlates with the survival time of the mice (**Figure 4.14C and F**). Collectively, the results indicate that the deletion of serine proteases did not dramatically affect the virulence of TIGR4 in the acute mouse pneumonia model. Only the mutant with functional PrtA+ showed a significant attenuation, indicating that the deleted proteases may play an important role in the acute pneumonia model.
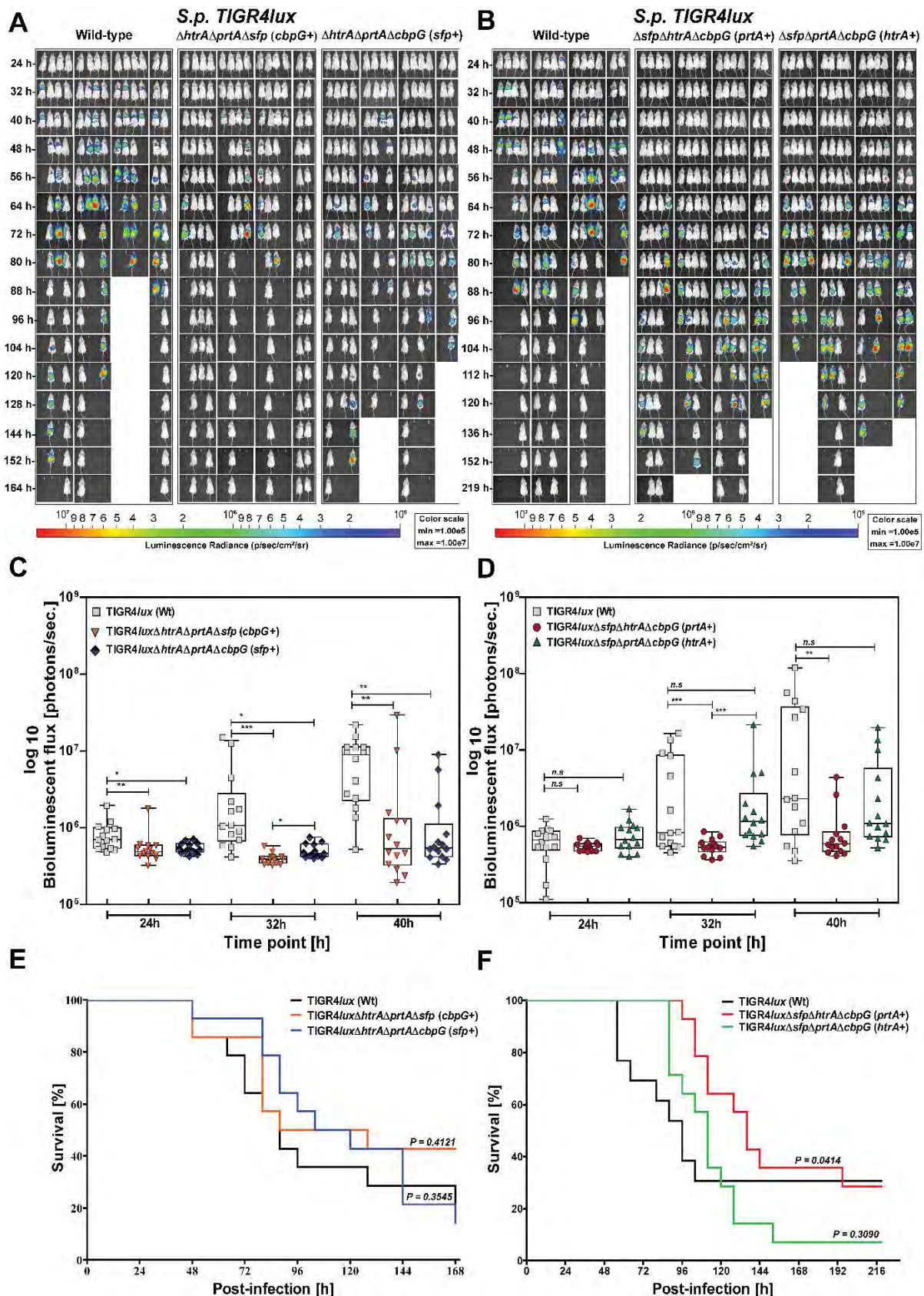
Figure 4.14: Impact of serine proteases on virulence acute pneumonia model.

Female CD-1 mice (8–10-week-old, 14 mice/group) were challenged with *S. pneumoniae* TIGR4*lux* wild-type or isogenic triple *serine protease* mutants using an infection dose of 9×10^7

CFU. **(A, B)** The course of infection was monitored in real-time using the IVIS®-Spectrum *in vivo* bioimaging system. The results are shown as **(C, D)**, the multiplication and dissemination of bioluminescent pneumococci in the infected mice quantified at indicated time points by measuring the luminescence intensity (photons/second). Data are shown as Box-Whisker graph showing the values for each mouse. Kruskal-Wallis test was used for statistical analysis. **(E, F)** Kaplan-Meier survival curves of mice infected with *S. pneumoniae* TIGR4*lux* or *serine protease* deficient mutants. The log-rank (Mantel-Cox) test was used for statistical analysis. *, *p<0.05*; **, *p<0.01* and ****p<0.001*.

4.5.2 Extracellular serine proteases are involved in pneumococcal colonization

Pneumococcal colonization of the upper respiratory tract is a prerequisite for invasive diseases such as pneumonia, meningitis, and septicemia[350]. Furthermore, it affects pneumococcal transmission from host to host, enabling pneumococci to spread within a population[351]. Importantly, the results of adherence to human nasopharyngeal epithelial cells have indicated a high reduction by the loss of functional serine proteases (**4.4.1**). Thus, we postulated that in the absence of serine proteases nasopharyngeal colonization significantly decreased *in vivo*. Hence, the impact of pneumococcal serine proteases on nasopharyngeal colonization was investigated in a murine colonization model. Female CD-1 mice (7 mice/group) were intranasally challenged with 10^7 CFUs of either 19F_EF3030 (wild type) or isogenic *serine protease* mutants. Pneumococci were recovered from the nasopharyngeal cavity and lung tissue after 2, 3, 7 and 14 days post-infection. Interestingly, the loss of serine proteases resulted in a substantial decrease (*p<0.01*, and *p<0.001*) of pneumococcal CFU in the nasopharyngeal cavity two days after infection compared to the isogenic wild-type. After days 2, 3, 7, and 14 post-infection, ten-fold less pneumococcal mutants were determined in comparison to the wild-type (**Figure 4.15A**). Moreover, the pneumococcal load of 19F_EF3030Δ*prtA*Δ*cbpG* mutant expressing only HtrA+ in the bronchoalveolar lavages was dramatically decreased. However, the other mutants did not show a significant reduction in comparison to the parental 19F strain, despite a trend toward decreased CFU in the lower respiratory tract (**Figure 4.15B**). Overall, mice colonized with mutants lacking serine proteases eradicate the pneumococci faster from the lower respiratory tract on days 7 and 14 compared to the 19F_EF3030 wild-type infected mice. These data confirm the low invasive potential of strain 19F in the lung host compartment. Altogether, pneumococcal adhesion to the epithelial cells of the nasopharynx is strongly affected by the

loss of different serine proteases, according to the results of the *in vitro* adherence results and the experimental mouse colonization model.
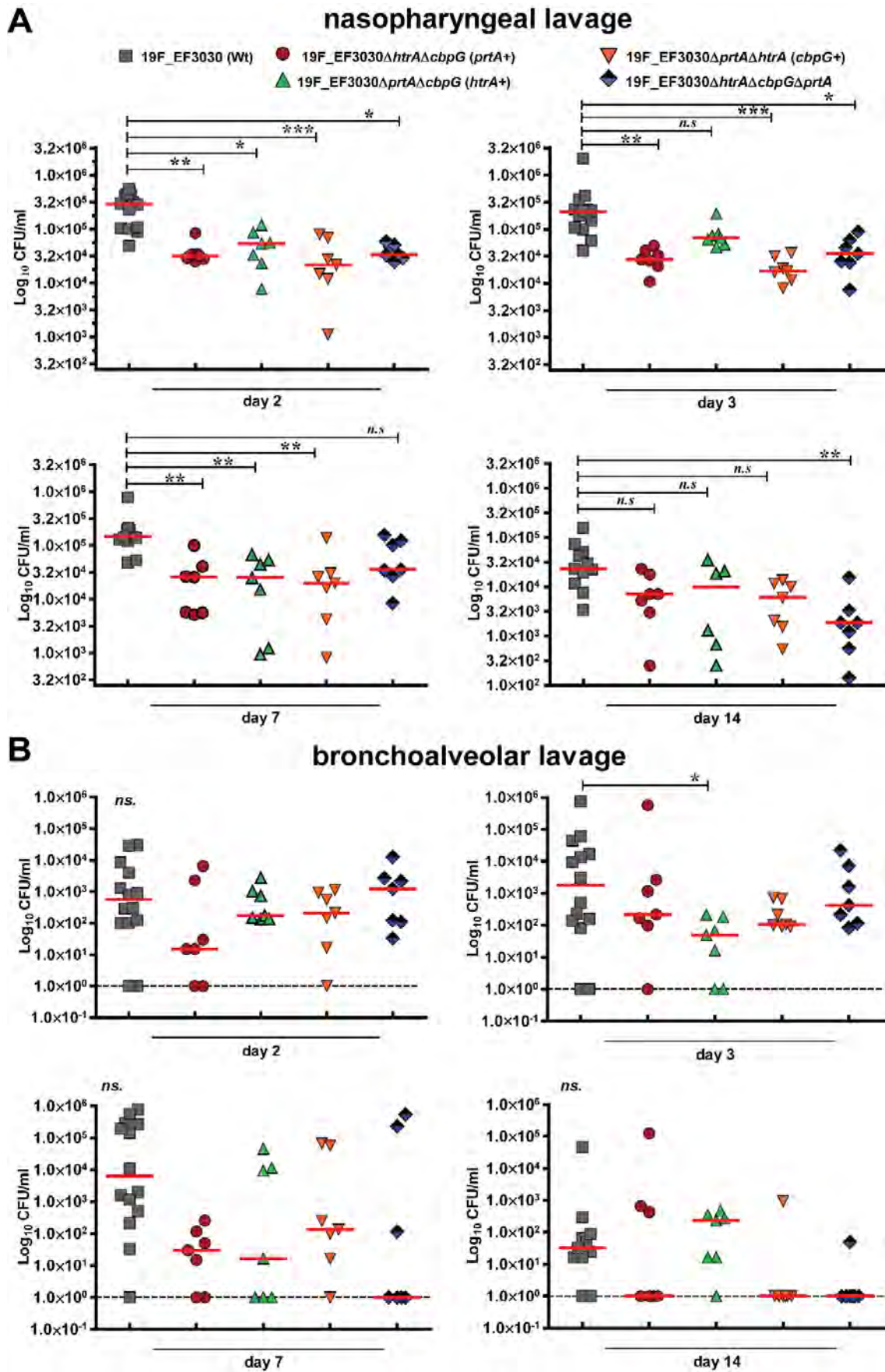
Figure 4.15: Nasopharyngeal colonization in a murine infection model.

Eight to ten-week-old female CD-1 outbred mice (n=7) were intranasally infected with a CFU of 1×10^7 pneumococci of serotype 19F_EF3030 or isogenic *serine protease* mutants: 19F_EF3030Δ*htrA*Δ*cbpG* (*prtA+*), 19F_EF3030Δ*prtA*Δ*cbpG* (*htrA+*), 19F_EF3030Δ*htrA*Δ*prtA* (*cbpG+*) or 19F_EF3030Δ*htrA*Δ*prtA*Δ*cbpG*. Mice were sacrificed at day 2, 3, 7 or 14 post-infections, and pneumococci recovered by a nasopharyngeal (**A**) or bronchoalveolar lavage (**B**) and plated on blood agar plates for quantification. Results are shown as scatter plots, where each dot represents one individual mouse. The data were statistically analyzed using a Kruskal-Wallis test *, *p<0.05*; **, *p<0.01* and ****p<0.001*. The dashed line represents the limit of detection.

4.6 The influence of serine proteases on biofilm formation

Biofilm formation of pneumococci has been described particularly during chronic otitis media (OM) and chronic rhinosinusitis[175, 352]. It has been speculated to occur also during colonization of mucosal surfaces[175, 186]. Pneumococcal biofilms have been recognized as a critical event in the pathogenesis of pneumococcal diseases[12, 177]. It is more difficult to eradicate colonization than invasive diseases due to the increased resistance of pneumococcal biofilms to antimicrobial agents[196, 353]. However, pneumococci in biofilms are a sessile heterogeneous bacterial population and are different from planktonic pneumococci (free-floating) regarding their physiology. Sessile pneumococci attached in a continuous-flow-through system to human surfaces were highly attenuated in experimental invasive disease models, suggesting a better correlation to bacteria during colonization[186]. The initial step of bacterial biofilm formation depends on the primary attachment to biotic or abiotic surfaces via proteinaceous structures. Because in addition to eDNA also proteins contribute to the integrity of EPS, it was of interest if serine proteases may alter biofilm development. Therefore, the impact of the serine proteases on the biofilm formation of *S. pneumoniae* was studied by using biological and abiotic surfaces in the following section.

4.6.1 Pneumococcal biofilm formation on biological versus abiotic surfaces.

To investigate the impact of extracellular serine proteases on pneumococcal biofilm formation, we have established two experimental biofilm models. The mucosal surface of the nasopharynx is the most important niche for pneumococcal colonization. Therefore, we used the substratum biofilm model to evaluate the biofilm formation *in vitro* and to mimic the *in vivo* environmental conditions. The pneumococcal biofilm in this model was grown on living epithelial cells using the human epithelial cell line Detroit-562 as described by Chao et al.[198]. During growth,

pneumococci produce large amounts of lactic acid, leading to acidification of the medium and, therefore, causing cell death[354]. To avoid this, the culture medium was changed every 4-6h to protect the epithelial cells from the low pH. In addition, a static biofilm model was also used as control without cells on sterile round coverslips as described by Marks, *et al.*,[12].

Both surfaces were inoculated with *S. pneumoniae* (10^7 CFU). Grown biofilm was visualized by scanning electron microscopy (SEM) at 24, 48 and 72h of incubation at 34°C and 5% CO_2. The ability of the 19F_EF3030 strains to form biofilms was first validated on both surfaces over time. On the epithelium substratum, an increase of biofilm mass was observed between 24h and 48h. After 72h, the epithelial cells became more visible and aggregates containing the pneumococci were detected. In contrast, strain 19F grown on glass coverslips showed an increase of pneumococcal growth over time and biofilm formation started after 48-h. The biofilm density of the 19F strain was higher on the epithelial substratum compared to that on glass cover slips (**Figure 4.16**).

Furthermore, to study the impact of the serine proteases on biofilm formation, mutants of strain 19F without or only one functional serine protease (CbpG+, HtrA and PrtA+) were cultured on the epithelium substratum. An overview SEM images is shown in **Supplementary Figure 8.10** for the biofilm on living cells and on a glass surface (**Supplementary Figure 8.11**). The scanning electron microscopy analysis revealed that less biofilm density was formed in mutant strains compared to the wild-type over 24h on both surfaces. Biofilm development on living epithelial cells was increased and a higher density after 48 and 72h was observed compared to the biofilm formation on a glass surface. These results suggest that the growth surface directly affects the morphology and structure of the biofilms. Therefore, the experiments were carried out by using only the biological surface and two time points were chosen to check the biofilm density by counting the viable pneumococci from the biofilm and the supernatant. In addition, an antibiotic resistance test with gentamicin was performed.
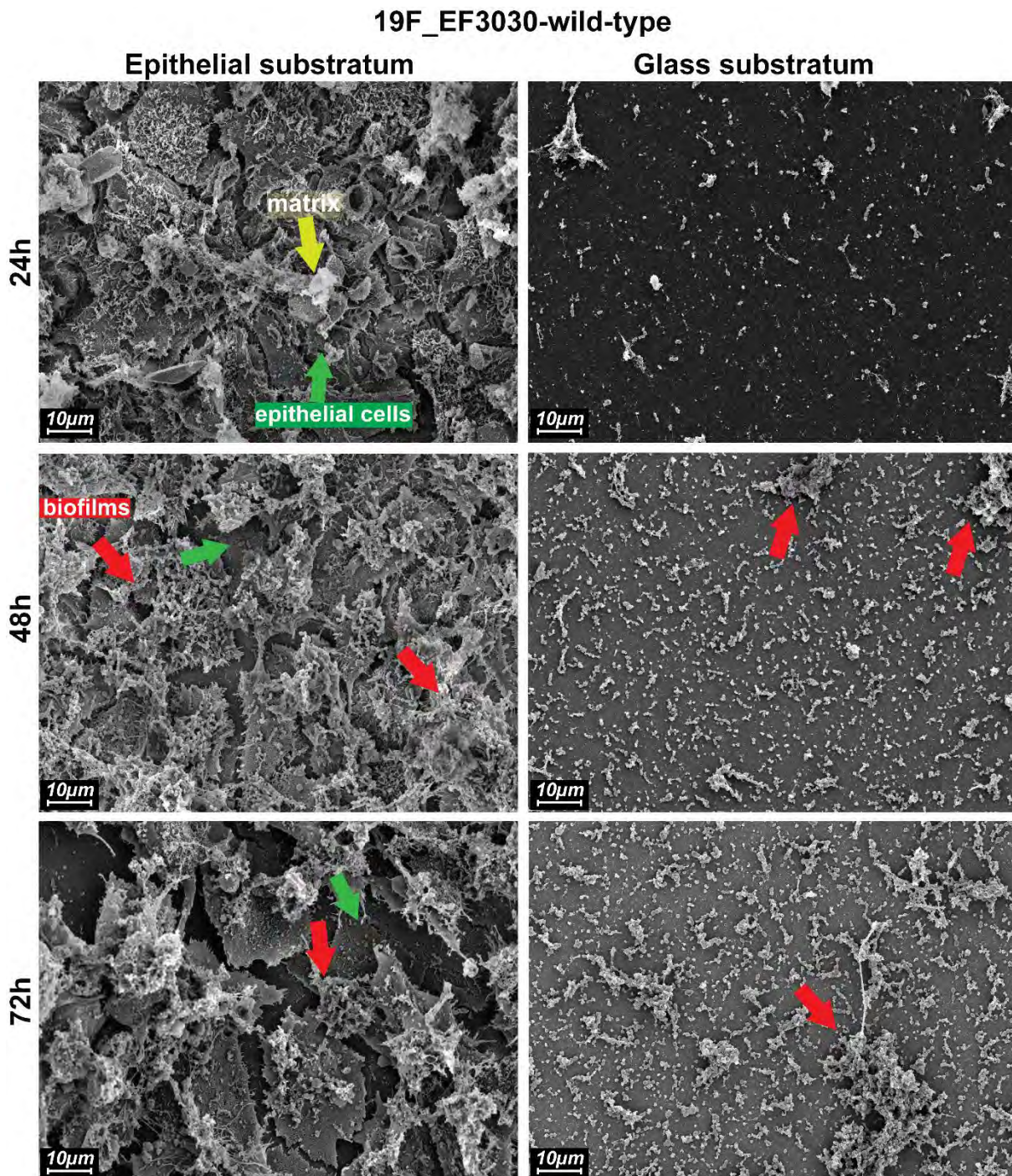
Figure 4.16: Comparison of *S. pneumoniae* biofilms formed on living epithelial cells or an abiotic surface.

Scanning electron microscopy images show the morphology of *S. pneumoniae* 19_EF3030 wild-type seeded on living nasopharyngeal epithelial cells (Detroit-562) (**A**), or on (**B**) glass substratum (round coverslips, diameter 12 mm) placed in 24-well tissue culture plates for 24h, 48h and 72h at 34°C incubation.

4.6.2   Pneumococcal biofilm on living epithelial cells

To investigate the influence of serine proteases on biofilm formation, living epithelial cells were used as a substratum mimicking the natural conditions in the nasopharynx. 19F_EF3030 wild-type or mutants were seeded on living human epithelial cells (Detroit-562) and incubated for 48h and 72h at 34°C (*see* **subheading 7.4.5**). After indicated time points, the matured biofilms were resuspended and plated on blood agar plates to enumerate the amount of bacteria in the biofilm. In addition, invasive pneumococcal infections appear to be facilitated by the dispersal of bacteria from colonizing bacterial cells in biofilms[355]. Therefore, pneumococci in biofilm disperse passively as the biofilm. The supernatant was also collected without disturbing the biofilm and was plated on blood agar plates to count the released bacteria. Bacteria within biofilms are much more resistant to antimicrobial agents than planktonic bacteria (free-floating). Gentamicin inhibits bacterial protein synthesis[12]. Therefore, gentamicin resistance during biofilm formation was further assessed in this model. Matured biofilms were incubated for 3h at 34°C in the absence or presence of gentamicin (*see* **subheading 7.4.5.6**).

After 48h, the 19F wild-type formed a well-structured biofilm on the epithelial surface. A similar biofilm phenotype was also observed in the mutant, which had only a functional HtrA protease. Mutants with only functional PrtA+ or only functional HtrA+ showed more cell gaps in the biofilm, and the density was increased in the functional CbpG+ mutant. The highest density of biofilm was observed in the mutant expressing with only CbpG+ and in the mutant where all serine proteases were inactivated. In addition, the mutant without serine proteases showed higher cell disruption compared to the 19F_EF3030 wild-type (**Figure 4.17**). As a control, the epithelial cell layer was incubated without pneumococci, demonstrating the integrity of the layer over time (**Supplementary Figure 8.9**). After 72h of incubation the biofilms were structurally less dense and less organized, although some morphological differences were detected compared to the 48h biofilm. The biofilm density was increased in the wild-type and the HtrA+ mutant, whereas the density was not changed in the CbpG+ mutant. The mutant with only HtrA+ showed a more dense biofilm and matrix formation compared to the wild-type. In contrast, in the mutant without functional serine protease, the epithelial cells were detached from the surface and biofilm was dissolved. The epithelial cell disruption after 48h biofilm led to cell detachment after 72h, and less dissolved biofilm was also detected in the mutant with only functional PrtA+ (**Figure 4.18**).

Biofilm structure after 48-h
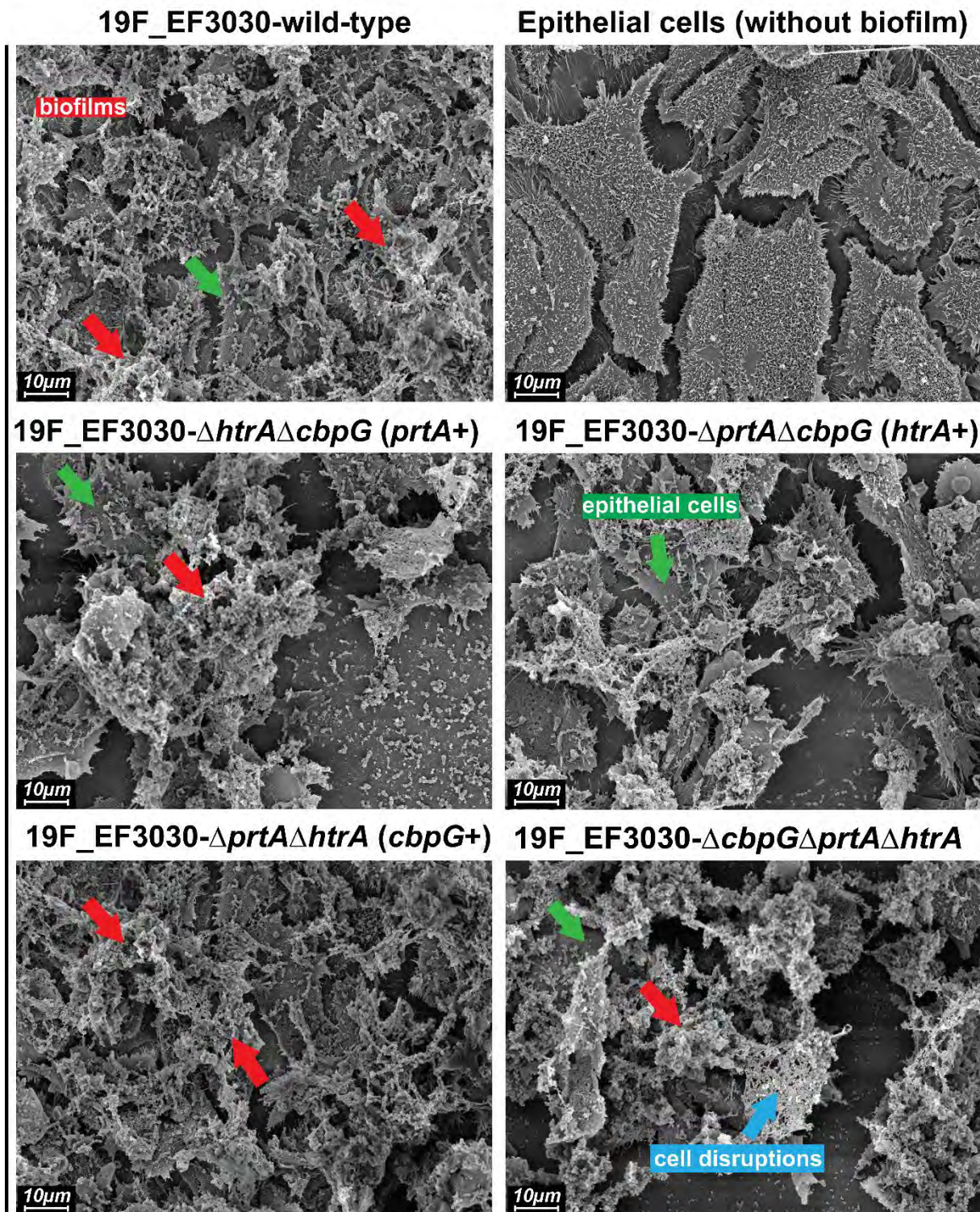
Figure 4.17: *S. pneumoniae* biofilm development of *serine protease* mutants after 48h. Developed biofilms by 19F_EF3030 wild-type and *serine protease* mutants illustrated by scanning electron microscopy. The epithelial cells are indicated by the green arrows, biofilm by red arrow and cell disruption in light blue. The white line represents the scale bar $=10\mu m$.

Biofilm structure after 72-h
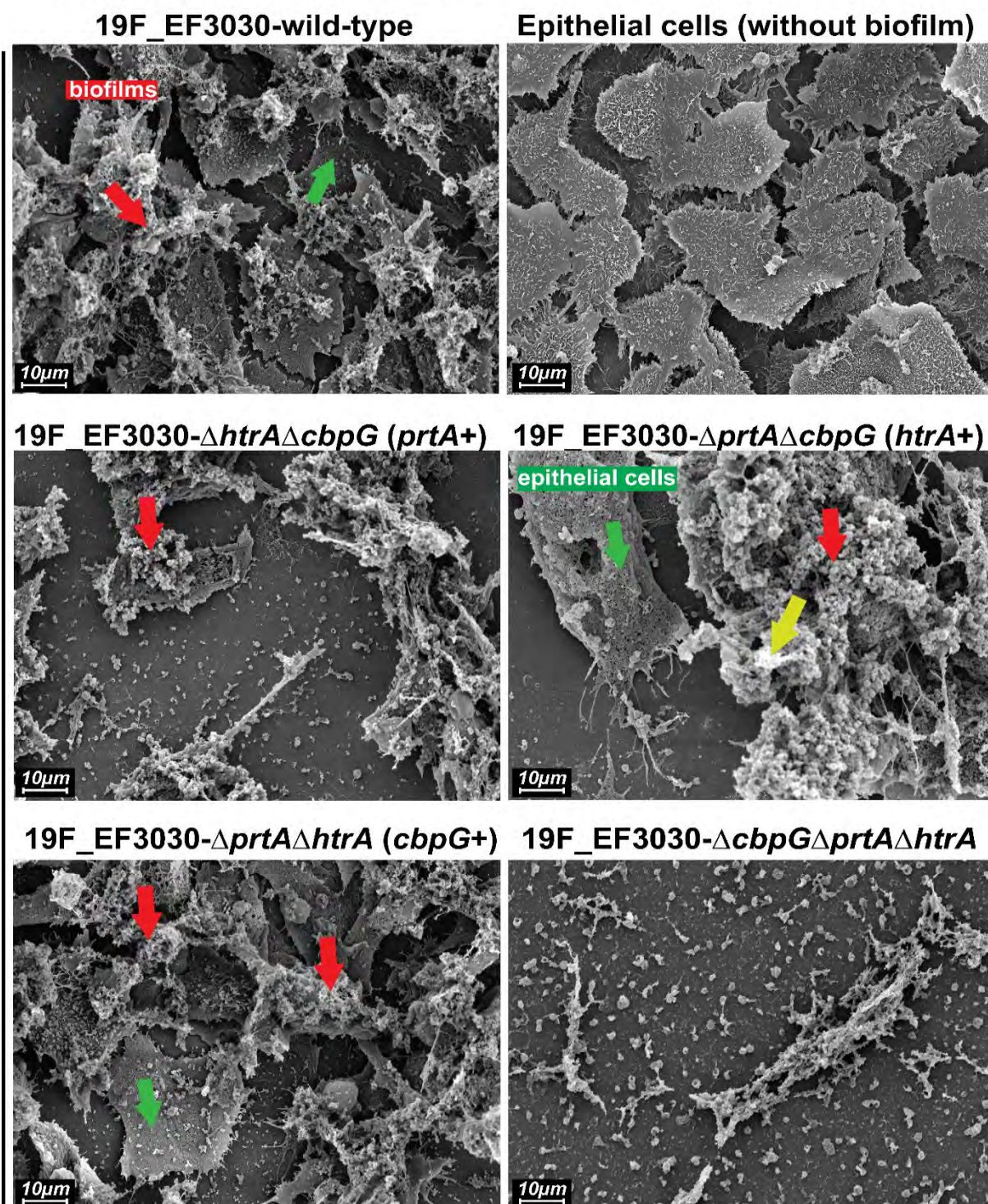
Figure 4.18: Development of biofilm morphology by *S. pneumoniae* 19F and *serine protease* mutants after 72h on epithelial cells.

19F_EF3030 wild-type or isogenic serine proteases mutant were incubated at 34°C on living epithelial cells in RPMI medium. Biofilm structures were detected by scanning electron microscopy. The white line represents the scale bar =10μm. The epithelial cells are indicated by the green arrows, biofilm by red arrow and biofilms matrix in yellow.

In addition, the viable bacteria and the viable bacteria released from the biofilm were counted to compare the wild-type with the *serine protease* mutants. After incubation for 48h and 72h. More viable pneumococci from strain 19F_EF3030 were obtained of the mutant with functional CbpG+ compared to the wild-type at both time points as shown in **Figure 4.19A and B**. The same results were received by counting the passively released pneumococci from the biofilm supernatant (**Figure 4.19C and D**). This is in accordance with the counted pneumococci after gentamicin treatment demonstrating that the biofilm lifestyle protects the bacteria against the antibiotic (**Figure 4.19E and F**). The other two *serine protease* mutants with a functional PrtA or HtrA showed indifferent results and were quite similar compared to the wild-type. The mutant without functional serine proteases showed a significant reduced number of recovered viable bacteria from the biofilm after 48h of incubation, but after 72h the numbers were at least the same as for the wild-type.

Furthermore, the significant difference of the biofilm formation between the wild-type and the mutant with only functional CbpG+ is represented in **Figure 4.20** with a higher resolution. The wild-type 19F showed more filamentary structures and more extracellular matrix. The CbpG+ mutant showed more bacteria and round shaped structures, like vesicles were located on the surface. However, these images show the importance of bacterial-host cell interactions which are essential to form functional and well-structured biofilm. The higher amount of bacteria in this picture of the CbpG+ mutant is in accordance with the counted bacteria from the biofilm (**Figure 4.19A**).

In conclusion, the scanning electron microscope images were not in line with the counted CFU. This may be due to biofilms, which were mechanically disrupted and resuspended before plated on blood agar plates. Therefore, the total CFU count reflects only viable pneumococci from the biofilms disregarding the other biofilm materials. As mentioned previously, the total biofilm matrix consists of three primary components: extracellular polysaccharides, nucleic acids, and proteins, as well as a high percentage of water[164]. The biofilms formed on living cells showed an increase in biomass after 48h compared to 72h, which correlated with increased resistance to gentamicin, indicating the development of mature and functional biofilms. Furthermore, reduced sensitivity to gentamicin killing of biofilm grown bacteria can be a result from altered metabolic activity or from forming a physical barrier (i.e., by the biofilms matrix) to the antibiotic, suggesting that the physical barrier is intact in biofilms to protect against the antibiotic.
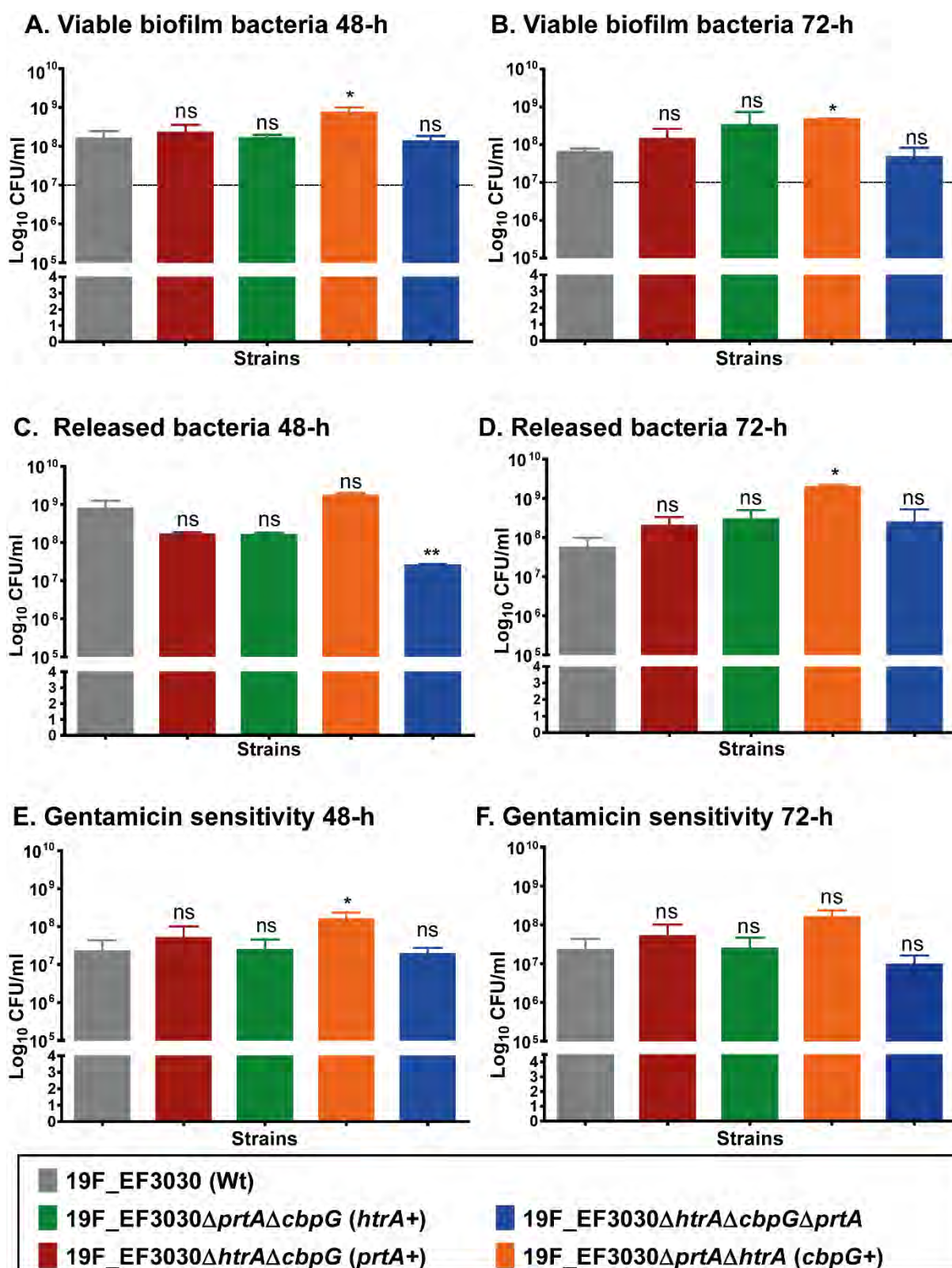
Figure 4.19: Role of serine proteases on biofilm formation was formed over 48-h and 72-h at 34°C on living epithelial cells.

(A, B) Viable bacteria from biofilms (CFU\ml) of wild-type and *serine protease* mutants was evaluated by plating the viable bacteria after 48-h and 72-h. Supernatants were removed, and biofilms were resuspended, serially diluted, and plated on blood agar plates. **(C, D)** Supernatants containing released bacteria from the matured biofilm were determined by plating. **(E, F)** Mature biofilms were washed with RPMI medium and exposed to 500 µg/ml gentamicin for 3-h at 34°C, to determine gentamicin sensitivity. Data represents duplicates or triplicates from three independent experiments, with the mean and SD displayed. Statistical analysis was performed using a Kruskal-Wallis test; *P<0.05* and **, *p<0.01*.
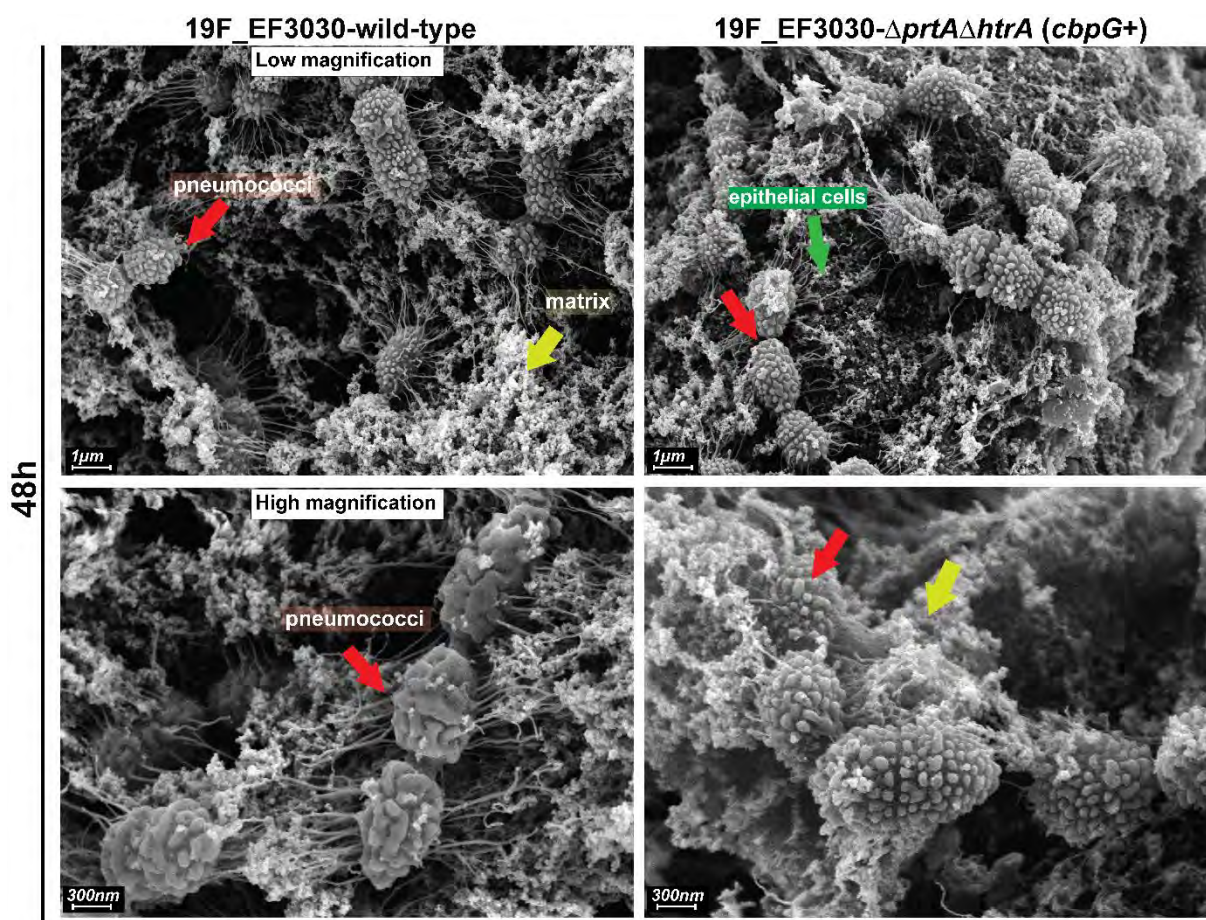
Figure 4.20: Characteristics of pneumococcal biofilms grown on living epithelial cells after 48h.

S. pneumoniae 19F_EF3030 wild-type or *serine protease* mutants with functional only CbpG+ were seeded on epithelial cells (Detroit-562). After 48h, biofilms were transplanted onto living epithelial cells: upper panel low magnification and lower panel high magnification. The white scale bar represents 1µm and 300 nm. The epithelial cells are indicated by green arrows, biofilms matrix in yellow, and pneumococci (diplococcus) aggregations by red arrow.

4.7 Heterologous expression and purification of pneumococcal serine proteases

The recombinant serine proteases were expressed in the heterologous host *E. coli* BL21 or *E. coli* ArcticExpress (DE3). The constructs of the expression vectors pTP1 or pDB-His$_6$-MBP containing the serine proteases target genes HtrA, CbpG, PrtA and SFP were produced in line with former MSc. theses by Nadine Henck, Robert Bolsmann, and BSc. Andreas Wüst. The heterologous expression and purification of all four serine proteases were optimized during this work.

Expression of serine proteases in *E. coli* was induced with 1 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside) using different conditions. Purification was performed by affinity chromatography using the "ÄKTA purifier liquid chromatography system" (GE Healthcare GmbH) according to the instructions of the manufacturer. After SDS-PAGE, the quality and purity of the recombinant serine proteases were confirmed by Coomassie Brilliant Blue (CBB) staining or silver staining. The purification was performed under native conditions in the absence of protease inhibitors enabling us to conduct the functional protease assays.

4.7.1 Heterologous expression and purification of HtrA

Serine protease HtrA (*sp_2239* in TIGR4; nucleotide 94 to 1182; amino acid 32 to 393) was fused to maltose-binding protein by cloning into the vector pDB-His$_6$-MBP. The recombinant plasmid was transformed into *E. coli* BL21. Purification of His$_6$-MBP-HtrA fusion was performed by affinity chromatography with the MBPTrap™ column. HtrA protein expression was induced with 1mM IPTG for 2h at 27°C incubation. Samples of the growing culture were collected before and after induction and analyzed by SDS-PAGE and CBB staining. The highest protein expression of His$_6$-MBP-tagged HtrA protein (~80 kDa) was detected after 2h of induction (**Figure 4.21A**). For further analysis of HtrA, the recombinant His$_6$-MBP-tagged HtrA (Nadine Henck[344]) was purified by affinity chromatography. Two proteins were identified by SDS-PAGE gel with a size of 42 kDa and 38 kDa (**Figure 4.21B**) due to the autocatalytic cleavage of the fusion protein. The elution fraction from the first purification containing His$_6$-MBP and rHtrA was loaded onto the HisTrap™ HP Ni-NTA column for the second step purification to separate rHtrA, which appeared in the flow-through with a molecular weight of ~38 kDa and displaced the His$_6$-MBP in the elution fraction (**Figure 4.21C**). SDS-PAGE and Coomassie Brilliant Blue staining revealed that the HtrA protein was significantly higher than 95% with a final concentration of 200µg/ml. The HtrA protein had a molecular weight of 42

kDa. After cleaving off the His$_6$-tag, the molecular weight of the protein was at 38 kDa. The purified protein was used for the protease activity assay with casein.
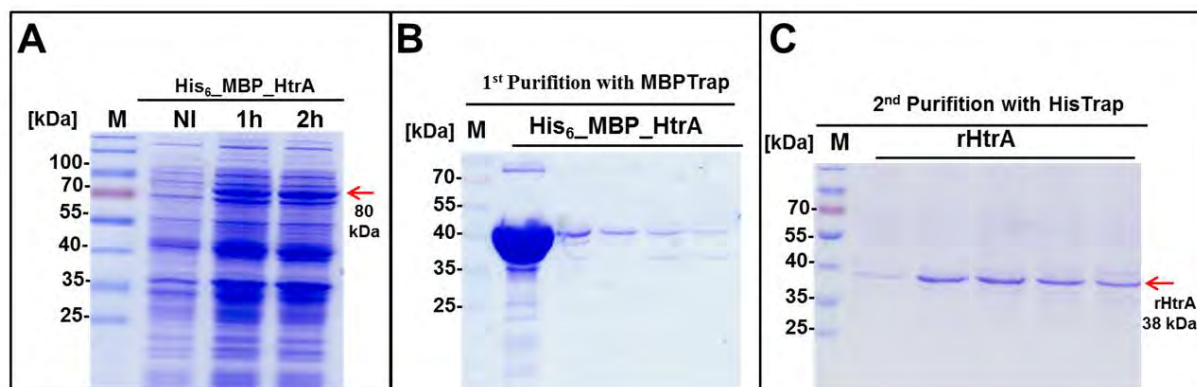
Figure 4.21: Heterologous expression and purification of His$_6$-MBP-HtrA in *E. coli* BL21.

(**A**) Recombinant *E. coli* (BL21 DE3) clone expression pDB-His$_6$MBP-*htrA* was cultivated in 500 ml LB medium at 27°C and induced with 1 mM IPTG. Samples were taken before and after induction and loaded onto 12% SDS-PAGE. Protein expression was detected by staining the gel with CBB. (**B**, **C**) Purified proteins were collected and separated by SDS-PAGE gel (15%) followed by CBB staining. The expected molecular weight of rHtrA is ~38 kDa. M: Prestained Protein Ladder; NI, not induced; IN, induced.

4.7.2 Purification of the recombinant CbpG and rCbpG$^{32-184aa}$ in *E. coli* ArticExpress

The rCbpG full-length and rCbpG$^{32-184aa}$ gene sequence encoding the functional domain was cloned in expression vector pTP1 to generate a His$_6$-tag fused protein (Andreas Wüst[343]). The expression was performed with *E. coli* BL21 (DE3) ArticExpress after overnight incubation with 1 mM IPTG for 16h at 12°C in LB medium. This strain is genetically modified to prevent the generation of protein aggregates known as inclusion bodies. High expression of the His$_6$-CbpG and CbpG$^{32-184aa}$ fused protein is shown at 16h post-induction, while no expression was detected without induction (**Figure 4.22A**). The bacteria were harvested after induction and lysed by ultra-sonification. Purification of the His$_6$-tagged serine proteases rCbpG and rCbpG$^{32-184aa}$ was carried out using HisTrap™ HP Ni-NTA column and the ÄKTA purifier liquid chromatography system with a Tris-HCL buffer system under native conditions. Final concentrations of 140µg/ml for CbpG and 380µg/ml were obtained. The purity of the CbpG proteins was checked by SDS-PAGE and CBB staining. The full length (without leader peptide sequence) His$_6$-tagged CbpG protein had a molecular weight of 34 kDa. While the His$_6$-tagged rCbpG$^{32-184aa}$ had a molecular weight of 19 kDa (**Figure 4.22B and C**). The His$_6$-tag was

removed via TEV-protease cleavage and CbpG protein was further used to check the thermos-stability and enzyme activity.
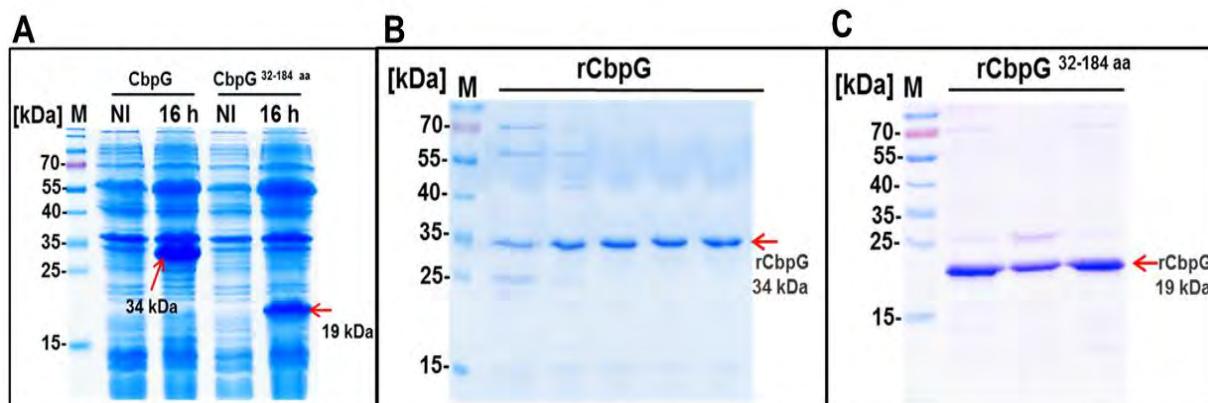
Figure 4.22: CbpG heterologous expression and protein purification by Ni-affinity chromatography.

(**A**) show the heterogeneous expression for pTP1-*cbpG* and pTP1-*cbpG*$^{32-184aa}$ in *E. coli* BL21 ArcticExpress were cultivated in 500 ml LB medium and induced with a final concentration of 1mM IPTG for 16-h overnight at 12°C. Pre- and post-induction samples were collected, and proteins were separated on 12% SDS-PAGE and stained with CBB. Purification was performed by using the HisTrap™ Ni-NTA column with the ÄKTA purifier system. Lanes in (**B**) show samples for the rCbpG full-length purification while (**C**) His6-tagged rCbpG$^{32-184aa}$ protein-containing elution fractions (10μl) were used in SDS-PAGE (12%) followed by CBB staining, thereby confirming the purity of the proteins.

4.7.3 Heterologous expression and purification of PrtA

Former studies showed that the N-terminal part of PrtA ranging from aa 28 to 942 is catalytic active and contains protease activity (Nadine Henck[344]). This reduced protein was called PrtA-2 (PrtA-2$^{28-942aa}$) and was purified as functional protein. The heterologous expression for PrtA-2$^{28-942aa}$ protein was performed by induced the 500 ml LB culture with 1mM IPTG for 3h at 27°C. The highest protein expression was detected after 3h of induction with a molecular weight of 104 kDa, and no expression were found in non-induced samples (**Figure 4.23A**). The bacteria cells were lysed and loaded into Ni-NTA HisTrap™ column for purification by affinity chromatography. Two purifications were required to get high purified protein (**Figure 4.23B and C**). The protein concentration was determined by using the Bio-Rad Bradford protein assay, with final concentrations 250μg/ml. The His$_6$-tagged PrtA$^{28-942aa}$ protein had a molecular weight of 104 kDa. The purified proteins were used for the protease activity determination with casein as substrate further analyzed for protein stability.
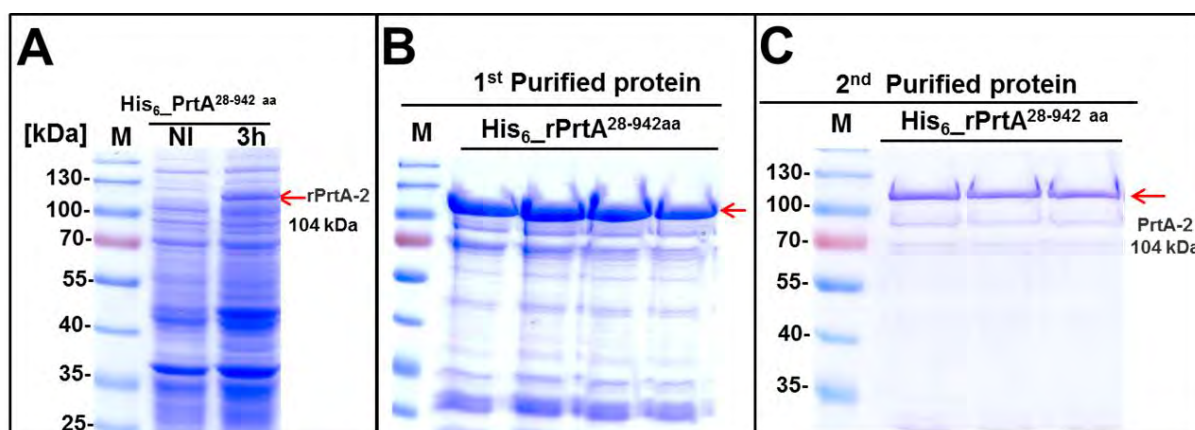
Figure 4.23: Heterologous expression and purification of His$_6$_PrtA$^{28-942 aa}$ in *E. coli* BL21.
(**A**) SDS-PAGE (10%) and CBB staining of bacterial cell lysates of PrtA$^{28-942 aa}$. The protein production of His$_6$-PrtA$^{28-942aa}$tagged was induced by 1mM IPTG, and the protein was purified by affinity chromatography. Purification was performed by a HisTrap™ HP Ni-NTA column using the ÄKTA purifier liquid chromatography system. His$_6$-PrtA$^{28-942aa}$ protein was eluted with an increased imidazole gradient. Lanes in (**B**) shows elution fraction samples from the 1st purification while lanes in (**C**) show His$_6$-tagged PrtA$^{28-942aa}$ of the 2nd purification loaded in SDS-PAGE gel (10%) and stained with CBB.

4.7.4   SFP heterologous expression and purification

The serine protease SFP (*spd_1753* in D39; nucleotide 79 to 1740) was cloned (Robert Bolsmann[356]) to generate the N-terminal His$_6$-tagged SFP for heterologous expression in *E. coli*. The recombinant *E. coli* (BL21 DE3) clone expression pTP1-His$_6$-*sfp* was cultivated in 500ml LB medium at 27 °C and induced with 1mM IPTG for 3h of incubation. The samples were taken after each hour less expression was detected even in non-induced samples as analyzed by SDS-PAGE and Coomassie Brilliant Blue staining (**Figure 4.24A**). The purification was carried out by affinity chromatography using the ÄKTA purifier system under native conditions. The purity of purified proteins was confirmed on SDS-PAGE stained with Coomassie brilliant blue. The recombinant SFP was detected with a molecular weight of 65 kDa and a final concentration of 150µ/ml (**Figure 4.24B**).
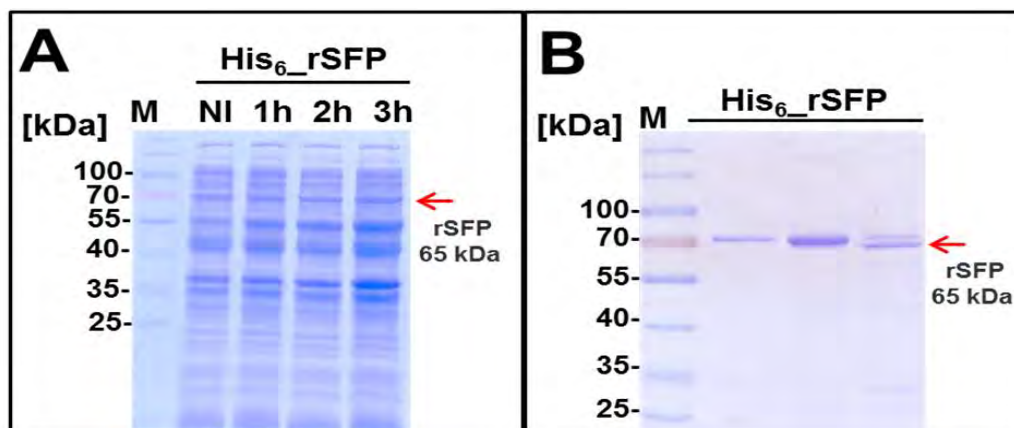
Figure 4.24: SFP heterologous expression and purification by Ni-affinity chromatography. (**A**) Heterogeneous expression for pTP1-*sfp* in *E. coli* BL21 was cultivated in LB medium at 27 °C and induced with an end concentration of 1mM IPTG for 3h. Pre-and post-induction samples were collected and proteins were separated on 12% SDS-PAGE. Purification was performed by using the HisTrap™ Ni-NTA column with the "ÄKTA purifier" system. Lanes in (**B**) shows the elution fraction samples of His$_6$-tagged SFP loaded in SDS-PAGE gel (15%) and stained with CBB.

4.8 Enzymatic activity of the recombinant serine proteases

4.8.1 Serine protease activity test with unspecific substrate casein

Previous studies could demonstrate that bacterial serine proteases can degrade several substrate molecules, including casein, albumin, hemoglobin and ECM proteins such as fibronectin, fibrinogen, and collagen[106, 209, 357]. Therefore, one of our perspectives is to look at the substrate specificity of pneumococcal proteases to understand the molecular mechanisms of these proteases on the human host. Purified pneumococcal serine proteases (**4.7**) were tested for thermos-stability and activity at 37°C by using *β*-casein as substrate control. The His$_6$-tagged serine protease was purified under native conditions (without proteases inhibitor) and dialyzed overnight against 50 mM Tris-HCl system containing 50mM Tris-HCl, 200mM NaCl, and 5mM CaCl$_2$, with pH 7.4 at 4°C. It is well recognized that serine proteases exhibit autocatalytic activity. Therefore, the thermos-stability test of serine proteases was checked to verify whether it is degraded at 37°C after 48h and that degradation of casein occurred only in the presence of serine protease. No degradation of the purified serine proteases was observed, suggesting that proteases are stable at 37°C for 48h. (First-panel **Figure 4.25A**). The following day, the serine proteases CbpG (5µg), PrtA$^{-228-942aa}$ (1.5µg), HtrA (5µg), and SFP (5µg) were incubated with 20µg *β*-casein at 37°C for 24h and 48h. After incubation time, the reaction was stopped by

freezing the samples at -20°C. Degradation of *β*-casein samples was assayed by 12% and 15% SDS-PAGE gels and followed by either CBB or silver staining. As shown in **Figure 4.25B**, the degradation of the substrate by the catalytic activity of rCbpG was detected only after 48h However, high degradation of casein was observed immediately after 24h and 48h when incubated with rPrtA-2$^{28-942aa}$ as shown by SDS-PAGE (**Figure 4.25C**). Furthermore, the casein was hydrolyzed faster by HtrA serine protease with two hours and after 48h, because the protein was completely degraded (**Figure 4.25D**). Despite SFP was showing two protein bands, the catalytic activity was detected after 48h (**Figure 4.25E**). In conclusion, serine proteases are stable at 37°C and casein can be cleaved by serine proteases and used as a control protein in our substrate assays.
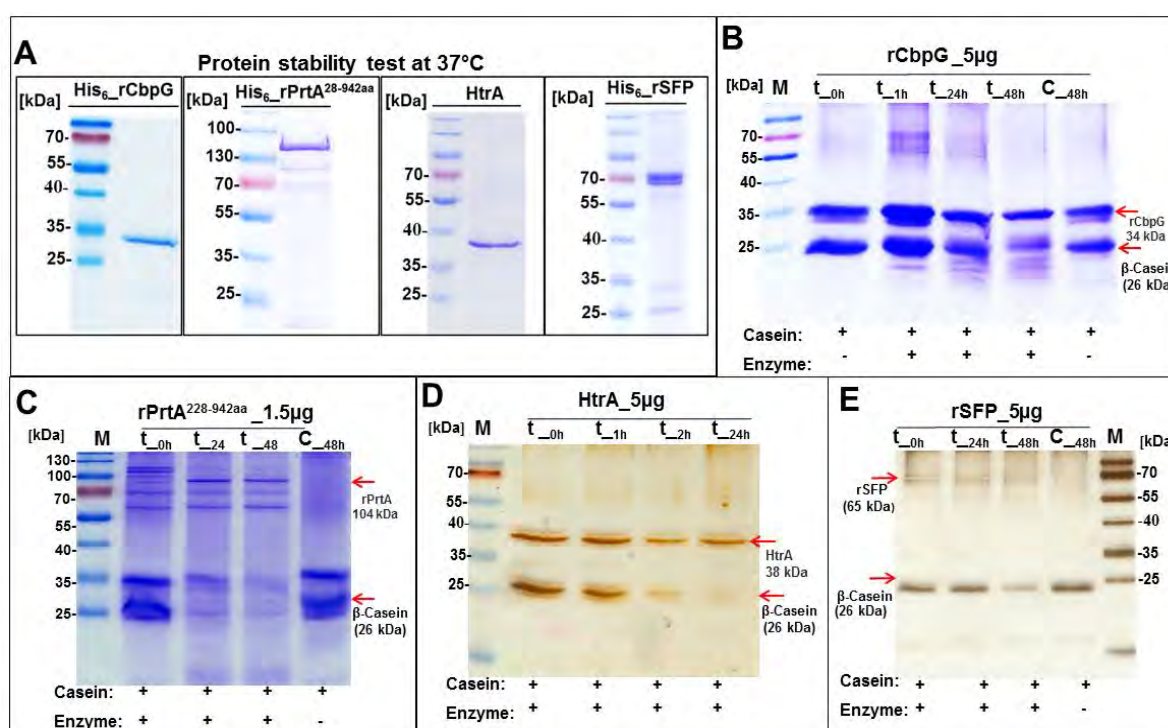
Figure 4.25: Determination of the stability of recombinant purified serine proteases and enzymatic activity.

(**A**) The purified serine protease proteins was incubated at 37°C for several days for stability test. (**B**) Enzymatic activity of 5µg rCbpG purified protein was incubated with 20µg casein as a substrate in Tris-HCL buffer system at 37°C. Samples were analyzed by SDS-PAGE (12%) and were stained with CBB. (**C**) Enzymatic activity of 1.5 µg of rPrtA$^{28-942aa}$ functional domain was incubated with 20µg casein in a Tris based buffer system for 24h and 48h. Samples were analysed by SDS-PAGE (15%) gel and stained with CBB to determine the degradation casein. (**D, E**) Enzymatic activity of 5µg rHtrA and rSFP were incubated with 20µg casein for different time points 1 h, 2 h, and 24 h at 37°C and the samples were analyzed by SDS-PAGE (15%) and

stained with silver staining, to determinate the degradation for substrate degradation by the serine protease rHtrA and rSFP. t_0: Casein (substrate) and serine protease (enzyme) at time 0, t_1-t_{48}: the reactions after different incubation times, and C48: only the casein (substrate).

Chapter 5

DISCUSSION

5. Discussion

S. pneumoniae is a clinically important human pathogen and has evolved versatile strategies to colonize the human nasopharynx[11, 358]. Nasopharyngeal colonization of the human URT by *S. pneumoniae* is the first step in an infectious process. High colonization rates facilitates pneumococcal transmission within the human population and thus increase the risk of pneumococcal diseases like pneumonia, meningitis, and sepsis[351]. Understanding the virulence factors that affect pneumococcal colonization and infection is important for the prophylaxis and treatment of pneumococcal diseases. Moreover, the identification and functional analysis of gene products involved in colonization and invasion is a key tool to understanding the host-pneumococci interactions.

Successful colonization or infection requires *S. pneumoniae* to overcome the host immune system. Pathogenic bacteria have evolved a wide range of mechanisms to circumvent the immune system. One of these mechanisms include the deployment of bacterial extracellular proteases that can efficiently degrade complement or ECM components[212]. Serine proteases in pathogenic bacteria are crucial key virulence determinants, as shown in **Table 5.1**. For example, the expression of serine proteases can cause cell disruption as shown for the HtrA protease of *Helicobacter pylori*[359] facilitating bacterial to spreading to deeper tissues. However, the enzymatic protease activity during pneumococcal infections can also contribute to the disruption of the epithelial barrier, cleavage of host proteins, including immunoglobulins and complement compounds, or degradation of extracellular matrix proteins[128, 216, 360]. Pneumococci can spread to the lungs, blood, middle ear, or the central nervous system during severe infection. Pneumococcal serine proteases are either secreted or bound to the cell surface. The extracellular localization enables a direct cleavage of peptide bonds, leading to the degradation of specific proteins[106, 225].

Despite considerable knowledge that demonstrates the impact of serine proteases on the pathophysiology of pneumococci, the individual role of single serine proteases on pathogenesis is only poorly understood. Pneumococci express up to four different serine proteases, which have in common a conserved catalytic site constituted by the serine, aspartate, and histidine triad. Therefore, it is still unknown if a single mutation of one of these proteases can be compensated by the other serine proteases. However, many studies have disregarded the redundancy of serine proteases in mutants with only a single serine protease knockout[100, 106, 234, 236, 238, 241, 318, 323]. Thus, the role of serine proteases contributing to pathogenesis is still a crucial

issue to understand. Therefore, we have used an alternative approach to investigate the influence of serine proteases on pneumococcal pathogenesis by addressing the following research questions:

1. What is the impact of each serine protease on nasopharyngeal colonization and on lung infection?

2. How does each serine protease contribute to pneumococcal biofilm formation?

3. What kind of proteins or ECM molecules can be degraded by these seine proteases?

4. Which host proteins are targeted by pneumococcal serine proteases? Alternatively, how do they cleave the ECM proteins with regards to host invasion?

The present study contributes to address the aforementioned research questions and provides detailed evidence on the importance of the combined role of the pneumococcal serine proteases HtrA, PrtA, SFP and CbpG, on pneumococcal pathogenesis. In particular, it was found that the deficiency of serine proteases impairs nasopharyngeal colonization significantly. Therefore, serine proteases have the potential to facilitate pneumococcal colonization and binding to host targets. This function turns them into promising targets for the development of specific antimicrobials to reduce pneumococcal colonization and transmission.

5.1 Impact of extracellular serine proteases on pneumococcal colonization in the nasopharynx

S. pneumoniae initiates infection by colonizing mucosal surfaces of the respiratory tract. Adherence of the pneumococci to host cells is the initial step and a prerequisite for a successful colonization. Some pneumococcal proteases were confirmed to have a possible functions in colonization and transmission[211] (**Table 5.1**). Therefore, it is evident that proteases secreted by pneumococci or bound to the cell surface play a pivotal role in pneumococcal pathogenesis[211]. Proteases have critical patho-physiological functions, including enzyme modification and cleavage of host immune proteins, which affects colonization and evasion of the host defense[361, 362]. Pneumococcal proteases are most likely involved in these functions by either degrading the host proteins or modifying the pneumococcal proteins involved in pathogenesis. This has been intriguingly shown for the zinc-metalloprotease ZmpC, which induces the removal of MUC16 from the epithelium[363], and the metalloprotease ZmpB[276], which was shown to modify pneumococcal surface proteins. In comparison, the IgA1 protease in pneumococci interacts with the human immune system by cleaving IgA into inactive components[216]. Furthermore,

several reports have also assessed the impact of serine proteases on pneumococcal virulence using different experimental infection models. In most studies, the pneumococcal mutants were deficient for only one of the mentioned serine proteases[100, 106, 234, 236, 238, 241, 318, 323].

Therefore, a potential redundancy of their mode of action cannot be finally excluded. In this study, we have analyzed the functional role of a single expressed serine protease on pneumococcal adherence and colonization in the genetic background of pneumococcal triple protease deletion mutants. For adherence and nasopharyngeal colonization studies, we have used the noninvasive serotype 19F (strain EF3030)[342]. The *S. pneumoniae* serotype 19F strain EF3030 described as the causative agent of otitis media has already been shown to be an efficient colonizer in murine model systems[364, 365]. Strain TIGR4 (serotype 4 strain) is well-known as an aggressive pneumococcal strain causing invasive diseases in humans and mice, and it is well known that encapsulated strains show only a low capacity to adhere to eukaryotic cells[82, 140]. Therefore, the invasive strain TIGR4Δ*cps* was used to test the ability of TIGR4 and their isogenic mutants to adhere to human Detroit-562 cells compared to the *in vitro* data of serotype 19F_EF3030.

The *in vitro* cell-culture and mice colonization model enabled us to characterize the impact of a single pneumococcal serine protease on nasopharyngeal colonization. Our results demonstrate that the attachment of *S. pneumoniae* to human nasopharyngeal cells was dramatically impaired in the absence of two or three serine proteases expressed by wild-type 19F_EF3030 and TIGR4Δ*cps*.

Our results with the mutant expressing only HtrA supports the finding from other studies which showed that the deletion of pneumococcal HtrA or its homologs in other bacterial species leads to a decrease in bacterial adhesion to human epithelial cells[366, 367]. It is widely accepted that HtrA plays an important role during human infection of many invasive bacterial species. As mentioned previously, HtrA in *S. pneumoniae* shares high amino acid sequence similarities to HtrA proteases of other Gram-negative and positive bacteria[287]. For instance, several reports were published characterizing various aspects of HtrA in *Campylobacter jejuni,* including interactions with host epithelial cells[368, 369]. The deletion of HtrA in *C. jejuni* showed a strong reduction in adherence and invasion of the bacteria to human epithelial cells[366]. Importantly, HtrA can be secreted to the environment where it is in close contact with host components, paving the way for bacterial spreading by degradation of the ECM matrix[370]. In contrast, membrane-associated HtrA in *H. pylori* has an important role in adherence, colonization, and persistent infections[208, 371]. Notable, our *in-silico* analysis of pneumococcal HtrA demonstrate that HtrA has no specific anchoring motif and seems to be surface-exposed, thus promotes

nasopharyngeal colonization[155]. Other studies showed that HtrA undergoes an auto-cleavage process[372] and is released to the environment. Hence, secreted HtrA facilitates the subsequent invasion of host tissue by degrading ECM components[287]. Collectively, HtrA homologs in other bacterial species were also involved in processing adhesins and thus in the activity of adhesins[287]. Whether the reduced adherence of a HtrA-deficient mutant is due to a blocking of a direct interaction with the epithelial cell or causing the HtrA-mediated effect on another unknown pneumococcal protein important for adherence, remains unclear.

CbpG is proposed to be a multifunctional protein cleaving ECM proteins and is involved in adherence as indicated by a reduced adherence of a TIGR4 *cbpG*-deficient to nasopharyngeal epithelial cells[100, 106]. Our results with *S. pneumoniae CbpG*-negative mutants indicate a dramatically reduced adherence to nasopharyngeal cells *in vitro* compared to isogenic TIGR4 and 19F_EF3030 wild-type strains. This is in accordance with other previous studies showing that CbpG has adhesive properties[106] and similar, a mutagenesis study with *cbpG*-mutants led to the loss of adherence in an *in vitro* adhesion model[100]. In these two studies infection experiments were performed with a single *cbpG* knockout in strain TIGR4 and R6[106]. In addition, the RNA transcriptome analysis for adherence-associated genes indicated that *cbpG* is upregulated during *in vitro* adherence to host epithelial cells in serotype 19F[373]. However, the specific mode of CbpG and its contribution to adhesion and colonization requires that CbpG is surface-bound by the C-terminal choline-binding domain[106]. Our *in silico* analysis results indicated that the choline-binding domain is missing in other analyzed serotypes strains such as D39 (serotype 2), Hungary19A-6 (19A), R6 (2) and ST556 (19F), due to a premature stop codon after the N-terminal catalytic functional module (**8.4.1.1**). Thus, it is not known how this truncation of the choline-binding module (CBM) has an impact on adherence in these pneumococcal strains

In contrast, both strains TIGR4 and EF3030 used here in infection experiments contain a CBM with at least four choline-binding repeats for the cell wall association of CbpG. In consequence, these findings require CbpG to be surface attached via CBM. Therefore, it is clear that the CBM of CbpG is most likely required for the CbpG-mediated pneumococcal attachment to epithelial cells, but this seems to be strain-dependent. Therefore it is not surprising that CbpG is involved in nasopharyngeal adherence, has been indicated for other pneumococcal choline-binding proteins such as PspC (CbpA/SpsA)[92, 374] and Pce (CbpE; phosphorylcholine esterase)[88], which were shown to be important for bacterial adherence to host cells[100, 315]. The underlying molecular mechanism of CbpG on the pathophysiology and how it is bound to the host epithelial cells maybe due to its fibronectin-cleaving ability. Nevertheless, studies are in progress to

further purify the protein to elucidate the structure and substrate specificity of functional CbpG for a better understanding of CbpG functions.

It was also demonstrated that PrtA is involved in adherence and colonization of pneumococci to host nasopharyngeal cells[236]. PrtA is a sortase A- anchored, surface localized proteins[319] and one of the largest pneumococcal surface proteins with a molecular weight of 240 kDa. The large C-terminal domain could be involved in adherence as described for other pneumococcal sortase-A anchored proteins such as PavB[225]. The adherence results with a PrtA+ strain in double or triple *serine protease* mutants revealed a reduced adherence to nasopharyngeal epithelial cells *in vitro*. To our knowledge, there are no studies available that have performed adherence studies with a *prtA*-mutant in *S. pneumoniae*. Nevertheless, pneumococcal PrtA is related to proteases described in lactococci[316]. There are similarities between the pneumococcal PrtA and its homologs PrtP (SK11), a proteinase in *Lactococcus lactis* that also show adhesive properties[79, 375]. However, PrtA homologs in other bacterial species were also involved in the processing of adhesins and thus in the activity of these adhesins. The N-terminal part of PrtA contains the catalytic domain followed by the DUF domain of unknown function, and it is not clear if the DUF or the C-terminal part is involved in adherence.

The molecular mechanisms of these proteases might differ and have to be explored in further studies because of the decrease in adherence of the 19F_EF3030 double deletion mutants with only one functional serine protease under *in vitro* infections condition. Based on these results, we hypothesized that the loss of serine proteases would also have an impact on pneumococcal colonization of the murine nasopharynx. Hence, the murine nasopharyngeal colonization model was used to assess whether the diminished adhesion of the 19F_EF3030 mutants to human epithelial cells correlates with the inability of serine protease deficient mutants to colonize the murine nasopharynx. The mice in the nasopharyngeal colonization model were challenged using a low bacterial dose of 1×10^7 CFU compared to the pneumonia model. This inoculum results in stable colonization of the epithelia without causing any symptoms of disease[376]. It should be mentioned that this model of nasopharyngeal colonization was limited and is similar to pneumococcal carriage in humans[376, 377]. The results of our experiment showed a highly significant decrease in the bacterial loads of the *serine protease* mutants in the murine nasopharyngeal colonization model compared to the parental strain 19F_EF3030. The small difference and an increased amount of recovered pneumococcal bacteria after 2 or 3 days of infection observed in the mutant with only functional HtrA+ might indicate that HtrA could be more important for the colonization than the other serine proteases. Another important finding

was that the use of a triple serine protease knockout in 19F (lacks all of serine proteases), showed a more pronounce reduction in colonization, indicating that serine proteases are indispensable for pneumococcal colonization. Thus, these results strongly suggest that serine proteases are indispensable for colonization.

Interestingly the effect of serine proteases are more pronounced in the nasopharynx compared to the lower respiratory tract, which is not surprising considering the low capacity of 19F_EF3030 to cause lung infections in the mouse pneumonia model[342, 355]. In addition, the double mutant expressing HtrA showed a similar behavior in the bronchoalveolar lavage, 14 days post-infection, compared to the wild-type 19F, despite early time points showing a significant reduction in colonization. Thus, each serine protease might contribute to pneumococcal adherence and the expression of single serine proteases is probably not sufficient to reach an adherence comparable to that of the wild-type.

The growth behavior of all *serine protease* mutants in complex THY medium or chemically defined medium was quite similar in comparison to the isogenic wild-type. These results were in accordance with another study conducted by de Stoppelaar and colleagues[234], in which they showed that the single deletion of a serine proteases has no effect on pneumococcal growth. We, therefore, do not believe that the inability of the serine proteases mutant to colonize the nasopharynx can be attributed to a constitutive growth defect or decreased bacterial fitness.

One interesting finding was that the strain 19F expresses no serine protease SFP, and SFP of strain TIGR4 can not be anchored to the peptidoglycan due to a premature stop codon (**8.4.1.2**). It should be mentioned that strain D39 encodes a complete SFP[234], which is anchored in a sortase A-dependent manner. Thus, in TIGR4, SFP seems to be secreted and it is not clear whether it is attached to the bacterial surface. The distribution of the *sfp* gene in different serotypes needs to be further explored. Although the studies were limited on SFP, our adherence results with only SFP expressed in TIGR4 strains showed that SFP may not be required for bacterial adhesion to the host epithelial cells.

5.2 *S. pneumoniae* serine proteases: roles in lung infection and their participation in biofilm formation

Following adherence and asymptomatic colonization, pneumococci must overcome host defense mechanisms to gain access to the lower respiratory tract and the systemic circulation, to cause invasive diseases. After nasopharyngeal colonization *S. pneumoniae* can develop pneumonia and septicemia, and it is of interest if serine proteases are involved in this process. Previous studies revealed that the caseinolytic protease P (ClpP) is involved in both diseases[248, 250, 313], and serine protease HtrA is involved in the development of acute pneumonia[211]. Studies with different mouse models using a *clpP*-mutant indicated the importance of the ClpP protease for nasopharyngeal colonization, lung infection, and systemic disease[313]. At the same time, systemic infection in mice with a *htrA*-negative strain had reduced the bacterial load in the lung, blood, spleen, and liver[155, 234]. However, we have shown that serine proteases are required for nasopharyngeal colonization. In this regard, one hypothesis could be that pneumococcal serine proteases are also important virulace factors that can avoid the recognition by professional phagocytes within the upper respiratory tract, allowing the pneumococci to efficiently colonize the nasopharynx. In addition, serine proteases appear to be required for bacterial spreading to the lung after intranasal infection. Therefore, the role of serine proteases were studied in a *in vitro* cell culture experiments using murine macrophages and *in vivo* acute pneumonia model. For that, gene deletion mutants were generated in the invasive TIGR4 background.

In the innate immune system, macrophages are sentinels and the first line of defense against bacteria in the lungs. The efficiency of host phagocytes to identify and destroy *S. pneumoniae* in the lungs and blood is dependent on their ability to recognize and eliminate the pathogen. Hence, the impact of serine proteases on interplay of the pathogen with phagocytes was analyzed by using J774A.1 macrophages. Consequently, these interactions were investigated in our studies using the standard phagocytosis assay in combination with DIF (double immunofluorescence) microscopy. Interestingly, the infection of macrophages resulted in a significant reduction of intracellular pneumococcal survivors in the TIGR4 triple serine protease knockout mutants in compatison to the wild-type strain. In line with this finding, the enumeration of triple mutants internalized by macrophages (determined by immunofluorescence microscopy) was also markedly reduced compared to the wild-type. These results indicate that serine proteases may be of great importance for the resistance aginst to phagocytosis mediated by macrophages. A clear correlation exists in resistance to phagocytosis in both group A and group B streptococci, which show that expression of a peptidase of the

subtilisin family (S8) of serine proteases inhibits the clearance and phagocytosis of the bacteria[378]. Evidence suggesting that bacterial proteases are involved in the inhibition of complement immunity and phagocytosis has been previously published[379]. In particular, it was proposed that bacterial proteases in general can usually be used to disable the antibacterial functions of bacterial killing inside the macrophages. Therefore, it can be proposed that serine proteases could play a role to prevent bacterial killing inside phagocytes[379].

The real-time bioimaging with bioluminescent pneumococci is an excellent method to image in real-time multiplication and dissemination of pneumococci in living mice[380]. When using TIGR4*lux* in the acute pneumonia model, our data show that the deficiency of three serine proteases out of four did not impair the full virulence of TIGR4 in CD-1 mice. However, the quantification of bioluminescence, as well as the monitoring of mouse survival in **Figure 4.14**, suggests that the triple knockouts have a slightly reduced capacity to cause pneumonia and, in consequence, invasive disease. The triple mutant with a functional PrtA+ also showed a significant attenuation in the acute pneumonia model. This is an interesting finding because the PrtA positive mutant is deficient in HtrA. The protein HtrA was already shown to be a major virulence factor in pneumococcal pneumonia caused by *S. pneumoniae* D39 in C57BL/6 mice, while SFP and PrtA played no significant role[234]. It has to be mentioned that C57BL/6 are more resistant to pneumococcal infections compared to CD-1 mice[381] used in our study. Despite being able to cause severe pneumonia in mice, D39 and TIGR4 differ in their genomic content and show different regulatory processes[380, 382]. The importance of HtrA is furthermore evident in our pneumonia model because the HtrA expressing triple *serine protease* mutant showed, according to the bioluminescence data, a similar multiplication in the lung compared to the parental TIGR4 strain. Mutants lacking HtrA show significantly lower bioluminescence and therefore bacterial burden, compared to the isogenic parental strain TIGR4. During influenza-pneumococcal co-infections, HtrA induces an inflammatory response when highly expressed, thereby enhancing the bacterial load in a mouse pneumonia model[245]. In addition, the mouse survival rates confirm the importance of HtrA as a major virulence factor of *S. pneumoniae* TIGR4. This is in accordance with earlier observations, which showed that HtrA is an important virulence factor [234]and contributes to the transition from nasal colonization to infection. It has been shown that antibodies against HtrA are protective against invasive pneumococcal diseases[303]. These observations are promising for the development of new antimicrobial treatmant with HtrA antibodies against *S. pneumoniae*.

Another study showed that the deficiency of PrtA in *S. pneumoniae* D39 reduced the virulence in a sepsis mouse model after intraperitoneal infection[236]. Considering our colonization data with 19F_EF3030 and the pneumonia data with TIGR4, we hypothesize that PrtA contributes to colonization but not to lung infections. A finding that confirms earlier data proposing that some of the sortase anchored pneumococcal proteins including serine protease PrtA have adhesive functions[234, 383]. Other studies showed that the deletion of the gene encoding sortase A reduces the virulence in *in vivo* models of colonization, pneumonia and bacteremia[384]. However, to decipher the individual impact of PrtA or other serine proteases on adherence and colonization or even pneumonia, single knockout strains and *in trans*-complemented mutants should be use for further investigation. SFP has previously been shown to play only a minor role in pneumococcal virulence of strain D39[234]. CspA, a serine protease from *S. agalactiae* (group B streptococci), which is highly homologous to SFP, can inactivate chemokines[324]. Our data confirm that SFP is probably not crucial for *S. pneumoniae* virulence, while its role in colonization is still elusive. However, it should be noted that we used two different strains to study the role of SFP, the 19F_EF3030 where SFP is missing and the TIGR4 strain with a C-terminal truncated SFP. Therefore, the role of SFP for pneumococcal fitness, virulence, or immunomodulation needs further investigation in other pneumococcal serotypes and it will be interesting to identify SFP substrates as well.

Biofilm formation also increases the pathogenic potential of *S. pneumoniae*. Biofilms protect pneumococci from antibiotics and host defense clearance. However, biofilms can also disperse bacteria to otherwise sterile sites. Like many other pathogenic bacteria, pneumococci must break down the extracellular biofilm matrix that embeds them in the colonizing biofilm to spread and cause infections. Previous studies showed that the exogenous addition of serine and cysteine proteases (trypsin and proteinase K) onto biofilms leads to a release of pneumococci from the biofilm[199]. Therefore, pneumococcal serine proteases may play a significant role in both, biofilm formation and the release of pneumococci from biofilms. Although several groups have studied *S. pneumoniae* biofilm formation, most of these studies were carried out on abiotic surfaces rather than in the context of human tissues or epithelial cells, which mimicks more the *in vivo* situation. Therefore, in our study, the importance of serine proteases on biofilm formation was further evaluated *in vitro* using a biofilm model with living epithelial cells. With strain 19F_EF3030 it has been shown that this strain can form well-structured biofilms on epithelial cells[12, 355]. Furthermore, antibiotic resistance in biofilms was evaluated by using the bactericidal antibiotic gentamicin.

Morphologically, we found that 19F_EF3030 strains formed phenotypically and structurally different biofilms when grown on biological surfaces like the nasopharyngeal epithelial Detroit-562 cells, compared to abiotic surfaces (glass or plastic). Under the described experimental conditions, pneumococci adhered well to living epithelial cells and formed a multicellular three-dimensional structure (**Figure 4.20**). This observation suggests that mucosal cells play an essential role during biofilm development, which was previously been shown by others[12]. Our analysis by scanning electron microscopy showed that biofilms from the mutants lacking all serine proteases exhibited different biofilm structures compared to biofilms formed by their corresponding wild-type strain. The biofilm of 19F_EF3030 wild-type showed an extensive three-dimensional "honeycomb"-like structure where the bacteria were embedded in an amorphous matrix. This finding is consistent with that of Marks et al. (2012)[12]. In contrast, the biofilm formed by the mutant without serine proteases was less dense, despite they produced biofilms of approximately similar biomass (viable bacteria from biofilms). However, how these morphological and phenotypic differences between wild-type and mutant without serine proteases look like is an interesting question for further characterization for example by 3D analysis with CLSM.

Obviously, cell disruption in the mutant without serine protease started early at 24h and continued at 72h, where all the cells seemed to be detached (**Figure 4.16**). Therefore, bacteria cannot bind to the epithelial cells and thus less biofilm were formed at 72h. So far, two studies showed that pneumococci show a strong binding to epithelial cells after forming biofilms on abiotic surfaces than the original culture[79, 385]. The specific role of serine proteases in biofilm formation on living cells has not been addressed yet. Prior studies have noted the importance of HtrA from *S. mutans* as a chaperone/protease during biofilm formation and oxidative stress tolerance[386]. In another study using single serine protease deletion mutants lacking one serine protease, either PrtA or HtrA in *S. pneumoniae* D39 strain, no difference in biofilm structures could be observed in comparison to the wild-type[199]. The mentioned study showed that PrtA and HtrA are upregulated in heat-induced biofilms, which mimics fever conditions, but only HtrA modulated bacterial release. However, biofilm analysis in this study was performed using prefixed epithelial cells[199]. We speculate that the deletion of all serine proteases may result in the upregulation of other virulence factors in *S. pneumoniae* such as H_2O_2 production that may cause higher cell cytotoxicity.

Interestingly, we were able to show that the mutant with functional CbpG produced a high dense biofilm that enhanced resistance to gentamicin. This is maybe due to a high release of

extracellular DNA[205] and extracellular polymeric substances (EPS)[387], which builds the matrix of the biofilm and that protect the growing pneumococci from gentamicin.

Furthermore, an increased number of released pneumococci from the biofilm to the supernatant was detected after 48 and 72h of growth. Studies on the biofilm of *S. pyogenes*, showed that SpeB as cysteine protease plays an important role in dispersal of bacteria from biofilms and subsequent disease progression[388]. Choline-binding proteins like the LytB glucosaminidase, and LytC lysozyme are involved in pneumococcal biofilm formation, colonization and immune evasion[389, 390]. It has been shown that LytB is required for *in vitro* biofilm formation, because LytB negative mutants (either Δ*lytB* or producing the inactive LytB$_{E585A}$) are incapable of forming biofilms[389]. Taken together, this suggests a possible role for choline-binding protein (CbpG) in the nasopharynx, in not only adherence or colonization but also in biofilm formation and subsequent infection[100].

It should be mentioned that the structural phenotypes for the biofilms shown by scanning electron microscopy analysis are different compared to the counted viable pneumococci from the biofilm. The biofilm matrix in between the bacteria requires extracellular DNA, lysis and the death of bacteria. Therefore, in our results, the quantification of the biofilm by counting viable bacteria only reflects the living bacteria without these materials. The possible explanation for these differences when biofilms formed on epithelial surfaces may be associated with the expression of various virulence determinants combined with their ability to colonize or cause infection *in vivo*, as shown previously[12].

In conclusion, our biofilm results, on the one hand, are of major interest for future studies to understand the role of pneumococcal serine proteases in biofilm formation. Because biofilm may provide insight into strategies to prevent pneumococcal colonization of the human host. On the other hand, in the substratum model used in this study, so-far-unidentified factors are likely to be involved in both the biofilm maturity and disruption. The proteome analysis of pneumococci identified covalently anchored serine proteases at the surface and all serine proteases in the exoproteome suggesting immunomodulatory functions and a potential impact on biofilm formation[103]. Therefore, proteomic analysis and *in vivo* studies (chinchilla model)[391] of biofilms at different stages of development and biofilms formed by mutants or wild-type pneumococci should be performed to potentially identify novel factors of importance for biofilm formation in the mucosal niche.

5.3 Structural and functional characterization of the catalytic domain of pneumococcal serine proteases

The pneumococcal serine proteases were highly conserved in different pneumococcal serotypes and show sequence similarities to other microorganisms (**Table 5.1**). Most of these proteases were shown to have protease activity. Therefore, in order to gain insights into the substrate recognition by the serine protease NisP that encoded for a subtilisin-like serine protease from *Lactococcus lactis*[392] was selected to show structural and functional features in comparison to the pneumococcal subtilase family protein SFP (**8.4.2**). As mentioned, all serine proteases have in common the typical Ser/His/Asp triad where the histidine is polarized through hydrogen bonding by aspartate, resulting in a polarization of serine and increased nucleophilicity of the hydroxyl oxygen atom. The highly conserved arrangement and distance between these three amino acids are crucial for the forming of the catalytic center and the cleavage of peptide bonds. This modeling revealed a similar subtilisin-like catalytic domain of SFP and PrtA. Here the overall fold consists of 7-stranded parallel beta-sheets, with the catalytic Asp on the first strand (S1) and five alpha helices containing Ser and His. While the catalytic motif seems quite similar, a protease-associated domain is found within the amino acid sequence of the PrtA catalytic domain, which may mediate protein-protein interactions or substrate specificity. Due to low sequence identity, it was omitted from the homology modeling and should be further explored. Furthermore, HtrA of *E. coli* is well characterized and studied in detail for its functional role as chaperone and protease activity. Previously, the crystal structure of HtrA from *E. coli* and *Termotoga maritima* showed the protease activity at elevated temperature as a trimer by hydrophobic interaction of the subunits[289]. The *in silico* structural analysis revealed a quite similar structure of the catalytic domain between HtrA and CbpG with the common double β-barrel core motif adjacent to the catalytic triad. The Asp and Ser. residues are localized on flexible loop structures, whereas the His residue is localized on a small helical fold (**Figure 5.1**).

Since this is only a simplified view of the active proteolytic centers of these serine proteases, there are ongoing efforts to purify the recombinant serine proteases for X-ray crystallography.
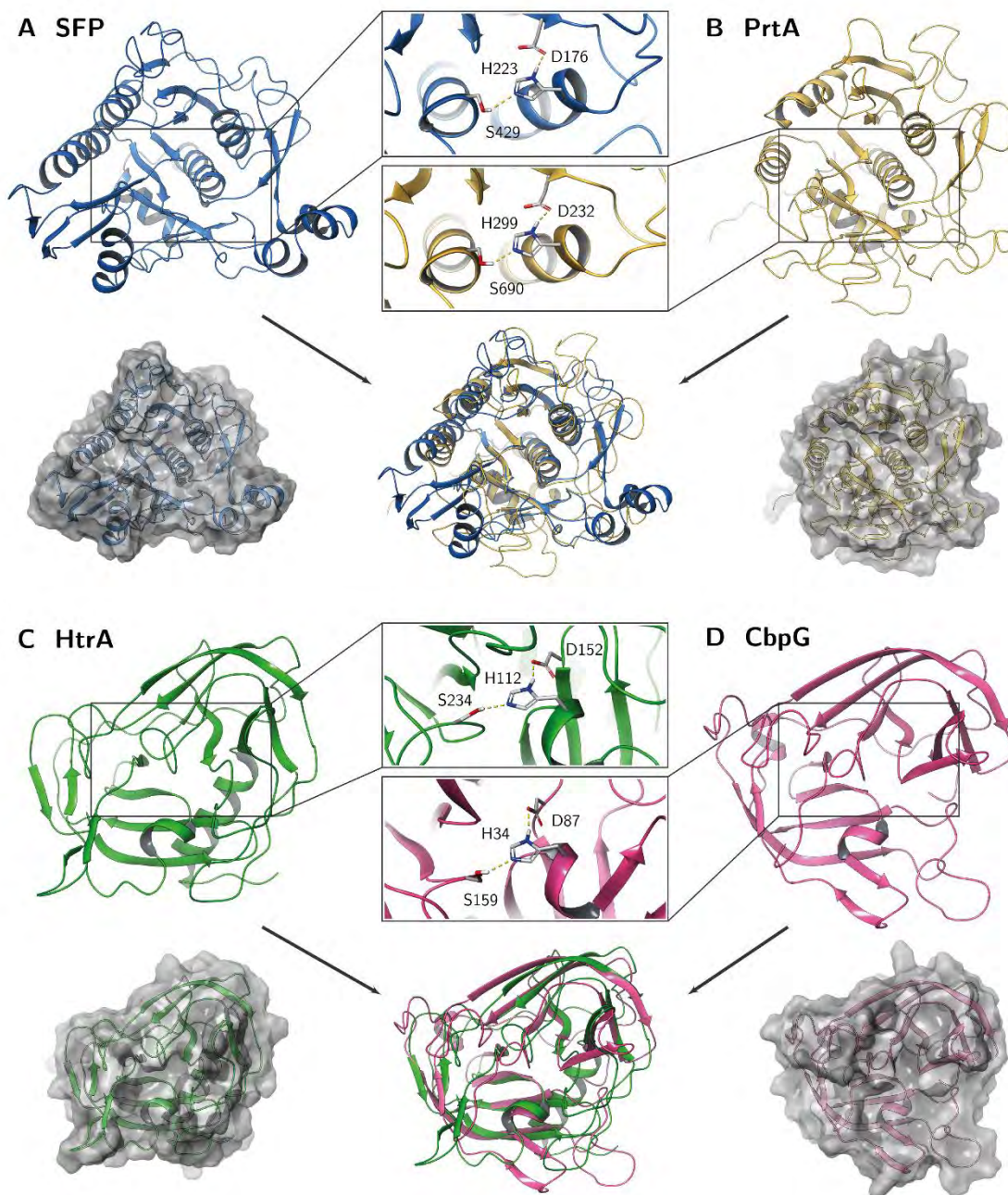
Figure 5.1: Comparison of predicted homology models of the pneumococcal serine proteases.

Catalytic residues (aspartate, histidine, and serine) are shown as sticks in detail. The calculations were performed within the Multiple Sequence Viewer/Editor application in Maestro 2020-4 (Schrödinger, LLC, New York, NY, 2020).[2]

---

[2] Computer based homology modelling was performed by Lukas Schulig Department of Pharmacy, University of Greifswald.

5.4 Serine protease degrade extracellular host proteins for invasion

Colonization is facilitated by the interaction between different bacterial adhesins with components of the ECM. *S. pneumoniae* secretes a variety of immune evasion molecules, including peptidases and proteases, which cleave components of the innate immune system or disrupt the integrity of the ECM[211]. In many pathogenic bacteria, it is believed that proteases and peptidases regulate ECM composition and facilitate cell migration by cleaving tissue barrier components such as collagen and non-collagenous glycoproteins[139, 393]. However, pneumococcal factors and the mechanism by which they breach host tissue barriers during the transition from colonization or biofilm to invasive infection are only partially understood. Pneumococcal serine proteases might be involved in the cleavage of adherence junctions or gap junction proteins to facilitate the pneumococcal paracellular route, which crosses the epithelial barrier, and leads to pneumococcal dissemination into the bloodstream.

It is assumed that pneumococci can avoid intracellular killing and invade human host tissues by taking the paracellular route[394]. Pneumococci thereby cleave adhesion junction (AJ) and tight junction (TJ) proteins such as epithelial cadherin (E-cadherin), occludins, and claudins to achieve this goal[395]. Interestingly, stimulation of toll-like receptors (TLRs) during pneumococcal infections down-regulate claudins, facilitating pneumococci movement across the epithelium[148]. Furthermore, in human lungs that are infected with pneumococci, a reduction of alveolar occludin, ZO-1, claudin-5, and E-cadherin, was observed[149]. There are only a few host proteins known to serve as biological substrates for PrtA and CbpG proteases which were recently identified. This includes collagen IV, plasminogen[225, 236] and fibronectin[106]. Therefore, to interpret the pathophysiological role of pneumococcal serine proteases and host protein targets further investigation is necessary. Former students of our working group previously performed the *in vitro* activity assay with recombinant purified serine proteases[36, 343, 356]. In their studies the recombinant purified serine proteases CbpG, HtrA and PrtA (functional domain) were incubated *in vitro* with different human host proteins such as collagen, fibronectin, fibrinogen, laminin, vitronectin, factor H, plasminogen, and thrombospondin-1. The initial protease activity assay of the recombinant serine protease with casein as substrate control under different conditions, such as buffer composition and incubation time, provided an overview of reaction conditions necessary for the catalytic activity of the proteases. However, these results still need to be confirmed with a different approach and more ECM proteins.

Fibronectin is a glycoprotein of the eukaryotic extracellular matrix and plasma which could be demonstrated to be a specific substrate for different pathogenic bacteria, including *S. pyogenes* and *S. aureus*[396, 397]. Our bioinformatic analysis shows that pneumococcal CbpG has homologies to other serine proteinases of different bacterial species such as *E. faecalis*, *S. aureus*, and *S. enterica*. These serine proteases have several functions, such as binding to the ECM fibronectin and promoting the invasion of human epithelial cells (**Table 5.1**). In *S. pneumoniae,* both PrtA and CbpG were shown to degrade fibronectin[106], which is consistent with our own results. This finding confirms the importance of CbpG in pneumococcal adherence and binding to epithelial cells[100, 106], which may also apply to the other serine proteases HtrA and PrtA as they cleave the same ECM proteins. Our finding, together with other studies, have indicated the importance of the serine protease CbpG as a factor modulating nasopharyngeal colonization and dissemination into the blood[100, 239]. Therefore, CbpG could play a role in the pneumococcal transition to the blood, which may be due to its fibronectin-cleaving potential[106]. The dual functions of CbpG, cleavage of host substrates and contributing to adherence to epithelial cells correlate with a substantial defect in the colonization of the nasopharynx by the *cbpG*-mutant. On the one hand, one can speculate that the proteolytic activity of CbpG on the bacterial cell surface can modify other pneumococcal surface proteins and enable them to interact with host cell receptors or soluble host proteins. On the other hand, CbpG modifies the ECM and eukaryotic cell surface, thereby facilitating adhesin-receptor interactions. These are still speculations and may also account for the other proteases.

Furthermore, many bacterial species besides *S. pneumoniae* possess a serine protease HtrA ortholog. The impact of HtrAs on bacterial pathogenesis was reviewed recently[287]. Most of the HtrAs can cleave adherence junctions, tight junctions, and ECM proteins such as fibronectin and proteoglycans, leading to a disruption of the epithelial barrier and, this mode of action is, therefore, critical for the host cell damage (**Table 5.1**). For example, the serine protease HtrA of *H. pylori* represents a crucial secreted virulence factor[398]. The disruption of the gastric epithelium leads to the transmigration of *H. pylori* across the epithelium and facilitates the oncogenic CagA protein injection into host cells. Consequently, HtrA can get into the extracellular space where it cleaves cell-to-cell junction factors, such as E-cadherin, leading to a disruption of the epithelial barrier, which then enables paracellular transmigration of the bacteria[399]. E-cadherin belongs to the cell adhesion molecule superfamily (CAM) and represents the target of several pathogenic bacteria, which invade the host[395, 400]. Interestingly, E-cadherin was described as an adherence receptor for the pneumococcal surface adhesin A

(PsaA), which also acts as a substrate-binding protein for manganese[401]. Pneumococcal HtrA may alternatively perform a function similar to those attributed to other members of the HtrA family. Our results did not show cleavage of E-cadherin by HtrA, which may be related to the temperature at which this experiment was performed (37°C). It has to be mentioned that HtrA protease activity is highest above 37°C (38-42°C)[286]. However, it seems that the involvement of HtrA in bacterial pathogenesis and the dual enzymatic activity of HtrAs have a common origin among (pathogenic) bacteria. Considering that, bacterial HtrAs show high similarities, particularly in their catalytic domain. Two strategies are possible and may explain the functionality of pneumococcal HtrA. First, the surface localization of HtrA can significantly influence adherence and colonization as this has been indicated by us and others published earlier[155, 239]. Second, HtrA undergoes the auto-cleavage process[372], and due to the secretion of HtrA into the environment, HtrA can degrade host components to facilitate invasion. These activities may explain why a deficiency of HtrA in *S. pneumoniae* leads to a reduced bacterial load in the blood, liver, and spleen[234, 242]. Therefore, whether pneumococcal HtrA functions in a similar role like in other bacteria with regard to degradation of E-cadherin *in vitro* remains a subject for further investigation.

The cleavage or degradation of ECM host proteins is not limited to the proteases activity of CbpG and HtrA. Also results from our group[344] show a cleavage of the tight junction protein E-cadherin by PrtA$^{-228-942aa}$. Many studies on pathogenic bacteria expressing serine proteases show that the bacterial proteases can cleave E-cadherin[402, 403]. This suggests that the pneumococcal serine protease PrtA can hydrolyze E-cadherin and open intercellular junctions enabling the pneumococci to take the paracellular path for pneumococcal invasion. Further studies also showed that the cell wall anchored PrtA is involved in the cleavage of host proteins like fibrinogen or collagen to penetrate tissues or escape from the immune system[383]. Besides, PrtA cleaves the host protein apolactoferrin to the even more bactericidal lactoferricin, thus facilitating the killing of pneumococci[238]. Therefore, it has been suggested that the presence of PrtA in the host may reduce pneumococcal load during systemic infection in mouse a model [234] due to the bactericidal effect of apolactoferrin[238].

Interestingly, our computer-assisted analysis and comparison revealed that catalytic center for CbpG and HtrA have almost identical three-dimensional structures besides the conserved catalytic triad aspartate, histidine, and serine. A similar subtilisin-like catalytic domain of SFP and PrtA was observed by this modeling.

Taken together, the presence of pneumococcal serine proteases seems to be required for colonization and dissemination and they seem to compensate for each other. This includes the serine proteases CbpG, HtrA and PrtA, which cleave fibronectin, vitronectin and thrombospondin-1. In addition, PrtA can also cleave the adherence junction E-cadherin and thus enable the pneumococci to disseminate to deeper tissues. Suggesting that pneumococcal serine proteases can interact with host proteins and facilitating host cell invasion and evasion of innate immunity. Therefore, serine proteases have the potential to facilitate pneumococcal colonization and binding to their host targets.
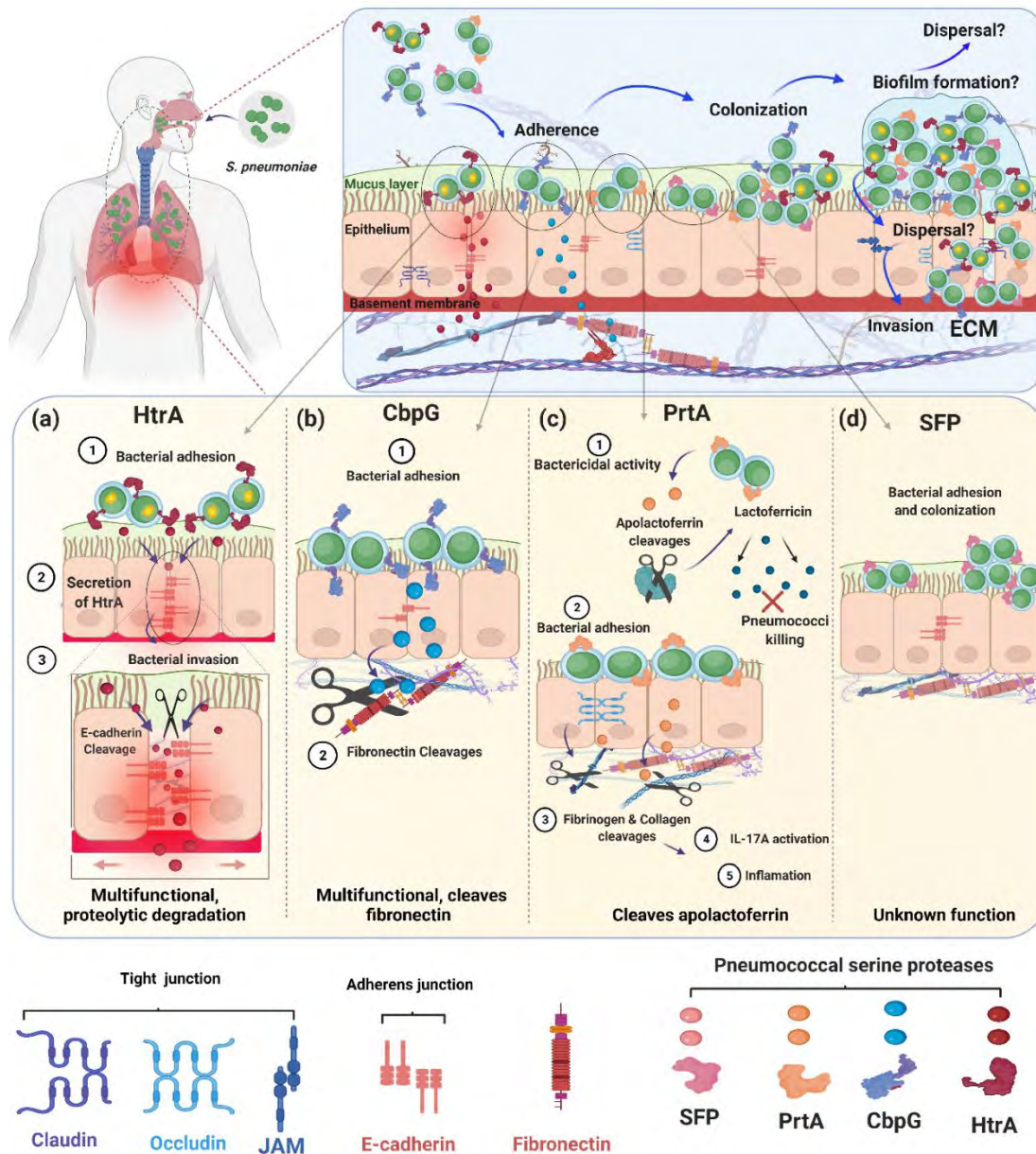
Figure 5.2: Role of pneumococcal serine proteases during adherence, colonization, and invasion of the human host.

The schematic models present different strategies of pneumococci during adherence, colonization and invasion and the role of extracellular pneumococcal serine proteases. The upper panel shows that all four serine proteases HtrA, CbpG, PrtA, and SFP are involved in attachment to epithelial cells during nasopharyngeal colonization. Bacterial colonization and persistence in different host niches are dependent on the pneumococcal adherence capacity to host cells and tissues. Attachment to host cells facilitates bacterial cell aggregation and the formation of sessile biofilm communities. During increased temperatures like fever, released bacteria can switch from asymptomatic colonization to severe infections. The lower panels **(A)** HtrA, **(B)** CbpG, **(C)** PrtA, and **(D)** SFP show the individual role of each serine proteases for pneumococcal adherence and their potential to cleave human ECM components. Figure created with **BioRender**.com.

Table 5.1. Comparison and distribution of pneumococcal serine proteases in other related bacterial species with amino acid sequence similarity and role in pathogenicity.

Protein (locus tag)	Bacterial species	Associated disease	Pathogenic function	Host-Targets	References
PrtA, cell wall-associated serine protease	*Streptococcus pneumoniae*	CAP1, sepsis, meningitis	Killing by apolactoferrin colonization adherence, pneumonia	cleaves human apolactoferrin, interact with collagen IV and plasminogen, cleavage of leader peptides from lantibiotics, possible adhesin	[13, 236, 239, 240]
PrtP (LP151), proteinase	*Lactobacillus paracasei*	dental caries, rheumatic vascular disease, septicemia, and infective endocarditis	degrades secreted, cell-associated, and tissue-distributed and other pro-inflammatory chemokines	degrade proinflammatory chemokines	[404, 405]
PrtP (SK11), Proteinase, PIII-type	*Lactococcus lactis subsp. cremoris*	Lactic acid bacteria (LAB), endocarditis chronic gastritis, central nervous infection	involved to adhesion and invasion, transit in the intestinal mucosa	adhesive properties, degrade alpha (S1)- and beta-caseins	[375, 406, 407]
PrtB, proteinase precursor	*Lactobacillus delbrueckii bulgaricus,*	LAB Non-pathogenic	antibacterial activity, probiotic function,	cleaves beta-casein	[408]
PrtH, cell envelope associated proteinase	*Lactobacillus helveticus*	LAB Nonpathogenic	antibacterial activity,	degrades alpha and beta-caseins	[409]
PrtS, cell envelope proteinase	*Streptococcus thermophilus*	LAB intestinal diseases	-	essential for growth	[410]
ScpA, C5a peptidase	*Streptococcus pyogenes*	necrotizing fasciitis, pharyngitis	facilitates the local infection	cleaves the human serum chemotaxis C5a	[411]
ScpB, C5a peptidase	*Streptococcus agalactiae*	bacteremia, pneumonia	virulence factor, promote Fn-independent GAS invasion of human epithelial cells	inactivates C5a	[337]

CbpG, choline-binding protein G	*Streptococcus pneumoniae*	CAP, sepsis, meningitis	adherence, colonization virulence factor,	cell-attached form promotes adherence, extracellular form degrades fibronectin, important formucosal and invasive disease	13, 100, 106, 239
GEJ60330 serine proteinase	*Enterococcus faecalis*	colonizing the gastrointestinal tract and oral cavity of animals and humans	endophthalmitis, peritonitis, endocarditis, and orthopaedic	-	412
serine protease	*Staphylococcus aureus*	CAP, bacteremia, endocarditis, osteomyelitis	involved in the evasion of host immunity	cleaves the ECM components	413
serine protease	*Salmonella enterica subsp.*	foodborne diseases (Salmonellosis)	epithelial cell invasion	cleavage of E-cadherin	414
Glu, endopeptidases	*Bacillus intermedius*	-	-	cleaves the peptide bond on the carboxyl end of glutamic acid	415
SFP, subtilisin-like serine protease	*Streptococcus pneumoniae*	CAP, sepsis, meningitis	facilitates bacterial growth, adherence, colonization	cleavage of leader peptides from lantibiotics	211, 234, 239
NisP, leader peptide-processing serine protease	*Lactococcus Lactis*	endocarditis infection	antibacterial lantibiotic	cleave leader peptides from lantibiotics	392, 416
CspA, cell surface serine endopeptidase	*Streptococcus agalactiae*	CAP, sepsis, meningitis	virulence factor, resistance to opsonophagocytosis	cleaves human fibronectin inactivates chemokines	324, 325
HtrA (DegP) serine protease/ chaperone	*Streptococcus pneumoniae*	CAP, sepsis, meningitis	chaperone, heat-shock protein, protease, virulence factor, competence pathways, growth advantage in influenza A virus co-infection, adherence, colonization	quality control of secreted proteins	155, 234, 239, 241-245
	Streptococcus pyogenes	purulent diseases of the pharynx and skin	processing of extracellular virulence factors and hemolytic activity	cleavage of complement factor C5a	417, 418

	S. agalactiae	bacteremia, pneumonia	-	-	
	Streptococcus mutans	dental carries	colonization	biofilm formation	[386]
	Campylobacter jejuni	Campylobacteriosis, Guillain Barré syndrome	bacterial adhesion, transmigration, and invasion	cleavage of E-cadherin, apoptosis, and immune responses	[419]
	Helicobacter pylori	gastritis, ulcers symptoms	bacterial transmigration, activation of type IV secretion	cleavage of occludin, claudin-8, E-cadherin, and fibronectin	[398, 420, 421]
	Mycobacterium tuberculosis	tuberculosis	cell wall hydrolases	degrades a putative cell wall muramidase (Ami3)	[422]

5.5 Conclusion

This study focused on the molecular biology of pneumococcal serine proteases and their pivotal role in pathogenesis, starting from adherence to colonization, biofilm formation and immune evasion. Pneumococcal serine proteases seem to have an important function during colonization and pneumonia. The ECM proteins serve as a molecular bridge, enabling pneumococci to interact with cellular receptors. However, the initial attachment of *S. pneumoniae* to epithelial cells is relatively weak. Therefore, the tight adherence of pneumococci in the respiratory tract is ensured by the expression of MSCRAMMs (microbial surface components recognizing adhesive matrix molecules) that recognize glycoproteins present in the ECM, thereby facilitating access to eukaryotic cell-surface receptors.

In this study, we confirmed the contribution of extracellular serine proteases in adherence to epithelial cells as demonstrated by *in vitro* adherence assays using the human pharyngeal epithelial cell line Detroit-562. This finding shed light on the importance of serine proteases during the process of colonization. The underlying molecular mechanisms behind this could be due to the proteolytic activity and adhesive function of these proteases. On the one hand, serine proteases likely contribute to host-pathogen interactions by degrading host proteins such as collagen, fibronectin and vitronectin. These host proteins represent the major components of the ECM. Thus, pneumococcal serine proteases can facilities the binding to host cells or even dissemination in the host. On the other hand, it cannot be excluded that serine proteases are also involved in the modification, cleavage, and release of other pneumococcal surface proteins, which might be a strategy to evade from the host immune system and facilitate adhesion to host cells by demasking important cell receptors. A direct adhesive activity has already been proposed for CbpG[106]. In addition, we show that CbpG enhances biofilm formation on living epithelial cells. This is likely due to its adhesive function. However, this adhesion function may also be possible for the other serine proteases but should be further investigated in future studies.

Our *in silico* analysis in combination with 3D hypothetical structural models revealed that the functional domains of the pneumococcal serine proteases CbpG, HtrA, and PrtA, are highly conserved. The exception is SFP, which seems to be produced only by a subset of strains. All serine proteases are secreted to the cell surface and, depending on the variant, are even released into the host environment. Secreted CbpG is able to degrade ECM-deposited Fn via its trypsin-like serine protease activity. Because pneumococcal HtrA is homologous to other bacterial

HtrAs and can also be secreted, it is plausible to hypothesize that it may cleave E-cadherin (**Figure 5.2**).

The 3D models show that the HtrA catalytic domain displays homologies to the CbpG catalytic domain, while SFP is quite similar to the catalytic domain of PrtA. Although all serine proteases have a typical catalytic triad, they might have different but also overlapping substrate specificities. The redundancy of serine proteases and probably their compensatory effect in the absence of one or more serine proteases makes it difficult to assess their individual contribution to pneumococcal fitness and virulence. This was clearly shown for PrtA, CbpG and HtrA, as they cleave fibronectin, vitronectin and thrombospondin-1, respectively. Thus, all studies are in parts limited in their conclusions because of the redundancy of these serine proteases. This, in turn, leaves gaps of knowledge such as e.g., substrate specificities and host compartment specificities that have to be deciphered in additional experiments *in vivo* and advanced *in vitro* models. The immunogenicity of functional domains of pneumococcal serine proteases in combination with their highly conserved protein sequences fulfills one of the requirements for a protein-based, serotype-independent (multi-) component vaccine. The individual potential as a vaccine candidate has, however, to be validated in future studies.

5.6   Outlook

Pneumococcal serine protease affects the recruitment of host factors and can also cleave and modify pneumococcal surface proteins resulting in the release of other protein fragments into the environment, which could modulate the host response. Therefore, it will be of interest to investigate the capacity of the selected serine proteases to cleave other pneumococcal proteins. The 3-D structures of pneumococcal serine proteases can be used to identify potential activators or inhibitors of proteases by virtual screening with different ligands. Thus, future aspects are the chemical synthesis and screening of compound libraries of promising candidates for *in vivo* studies and solving the structures in complex with its target protease. The purification of recombinant serine proteases heterologously expressed in *E. coli* was limited due to the formation of inclusion bodies. Therefore, the use of other bacterial systems such as *Bacillus subtilis* or *L. lactis* should be tested.

The contribution of pneumococcal serine proteases in host cell barrier disruption, as shown in *in vitro* activity assays of ECM degradation and biofilm formation, should be further investigated. The interaction of pneumococci with host cells or tissues induces host cell responses, thus leading to the activation of matrix metalloproteinases (MMPs) and the release

of pro-and anti-inflammatory cytokines. Hence, the role of pneumococcal serine proteases on MMP activation and induced inflammatory response could be characterized by measuring different cytokines and chemokines by flow cytometry and ELISAs

Chapter 6

MATERIALS

6.   Materials

6.1 Microorganisms and culture media

6.1.1 *Escherichia coli*

Table 6.1: *E. coli* wild type strains for storage of recombinant plasmids or protein expression

Strain	Genotype	Source
DH5α	F⁻ Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(rₖ⁻, mₖ⁺) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ⁻	Novagen
BL21 (DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	Stratagene
ArcticExpress (DE3)	*E. coli* B F⁻ *ompT hsdS* (rᵦ– mᵦ–) *dcm⁺* Tetʳ *gal* λ (DE3) *endA* Hte [*cpn*10 *cpn60* Gentʳ]	Agilent

Table 6.2: *E. coli* strains used in this study for protein expression

No.[1]	Strain	Recombinant plasmid	Properties	Resistance[2]	Reference
E1103	BL21 (DE3)	pTP1-His₆-*cbpG*	Protein purification of CbpG (*sp_ 0390*)	Kmʳ	BSc. Andreas Wüst, AG, Hammerschmidt, 2017
E1064	BL21 (DE3)	pTP1- His₆-*cbpG* ³²⁻¹⁸⁴ᵃᵃ	Protein purification of CbpG (*sp_ 0390*)	Kmʳ	BSc. Andreas Wüst, AG, Hammerschmidt, 2017
E1212	Arctic Express (DE3)	pTP1-His₆-*cbpG*	Protein purification of CbpG (*sp_ 0390*)	Kmʳ	This work
E1213	Arctic Express (DE3)	pTP1- His₆-*cbpG* ³²⁻¹⁸⁴ᵃᵃ	Protein purification of CbpG (*sp_ 0390*)	Kmʳ	This work
E951	BL21 (DE3)	pDB-His₆-MBP-*htrA*	Protein purification of HtrA (*sp_2239*)	Kmʳ	MSc. Nadine Henck, AG, Hammerschmidt, 2014
E1011	BL21 (DE3)	pTP1- His₆-*prtA* ²⁸⁻⁹⁴²ᵃᵃ	Protein purification of PrtA²⁸⁻⁹⁴²ᵃᵃ (*sp_0641*)	Kmʳ	MSc. Nadine Henck, AG, Hammerschmidt, 2014
E1060	BL21 (DE3)	pTP1-His₆-*sfp*	Protein purification of SFP (*sp_1954*)	Kmʳ	MSc. Robert Bolsmann, AG, Hammerschmidt, 2017
E1128	DH5α	pTP1-His₆-*sfp* ²⁹⁻³⁹⁷ᵃᵃ	Protein purification of SFP²²⁸⁻⁹⁴²ᵃᵃ (*sp_1954*)	Kmʳ	MSc. Robert Bolsmann, 2017
E1140	Arctic Express (DE3)	pTP1-His₆-*sfp* ²⁹⁻³⁹⁷ᵃᵃ	Protein purification of SFP²²⁸⁻⁹⁴²ᵃᵃ (*sp_1954*)	Kmʳ	This work

6.1.2 *Streptococcus pneumoniae* strains

Table 6.3: Wild type strains of *S. pneumoniae*

Strain no*	Serotype	NCTC Number	Source or Reference
SP408 (19F)	19F_EF3030	-	Andersson, et al., 1983 [341]
SP261 (TIGR4)	4		Tetteline et al., 2001 [54]
SP257 (D39)	2	NCTC 7466	
SP37 (35A)	35A	NCTC10319	Hammerschmidt et al., 1995

*Numbering is the stock list of the Department of Molecular Genetics and Infection Biology, Interfaculty Institute of Genetics and Functional Genomics, Center for Functional Genomics of Microbes University of Greifswald.

Table 6.4: Pneumococcal mutant strains

Strain no.*	Genotype (gene locus tag)	Resistance**	Phenotype	Reference
S. pneumoniae 19F_EF3030 mutants:				
PN762	19F_EF3030Δ*htrA* (*EF3030_11105*)	Cm^r	-	This work
PN763	19F_EF3030Δ*prtA* (*EF3030_03025*)	Erm^r	-	This work
PN769	19F_EF3030Δ*cbpG* (*EF3030_01920*)	Spec^r	-	This work
PN768	19F_EF3030Δ*cbpG* (*EF3030_01920*)	Erm^r	-	This work
PN765	19FΔ*htrA*Δ*cbpG* (*EF3030_11105, EF3030_01920*)	Cm^r, Spec^r	*prtA$^+$*	This work
PN770	19F_EF3030Δ*prtA*Δ*htrA* (*EF3030_03025, EF3030_11105*)	Erm^r, Cm^r	*cbpG$^+$*	This work
PN766	19F_EF3030Δ*prtA*Δ*cbpG* (*EF3030_03025, EF3030_01920*)	Erm^r, Spec^r	*htrA$^+$*	This work
PN767	19F_EF3030Δ*htrA*Δ*cbpG*Δ*prtA* (*EF3030_11105, EF3030_01920, EF3030_03025*)	Cm^r, Spec^r, Erm^r	*no serine proteases*	This work
S. pneumoniae TIGR4 mutants:				
PN259	TIGR4Δ*cps*	Km^r	capsule deficient	Schulz et al., 2014 [380]
PN494	TIGR4Δ*cps*Δ*prtA* (*sp_0641*)	Erm^r, Km^r	-	This work
PN488	TIGR4Δ*cps*Δ*htrA* (*sp_2239*)	Erm^r, Km^r	-	This work
PN663	TIGR4Δ*cps*Δ*cbpG* (*sp_0390*)	Km^r, Erm^r	-	This work
PN674	TIGR4Δ*cps*Δ*sfp* (*sp_1954*)	Km^r, Spec^r	-	This work
PN681	TIGR4Δ*cps*Δ*htrA*Δ*sfp*	Km^r, Erm^r, Cm^r	*prtA$^+$, cbpG$^+$*	This work
PN530	TIGR4Δ*cps*Δ*prtA*Δ*htrA*	Km^r, Erm^r, Cm^r	*sfp$^+$, cbpG$^+$*	This work
PN682	TIGR4Δ*cps*Δ*prtA*Δ*sfp*	Km^r, Erm^r, Cm^r	*htrA, cbpG$^+$*	This work

PN692	TIGR4Δ*cps*Δ*htrA*Δ*cbpG*Δ*sfp* (*sp_2239, sp_0390, sp_1954*)	Km^r, Erm^r, Cm^r, Spec^r	*prtA$^+$*	This work
PN685	TIGR4Δ*cps*Δ*prtA*Δ*htrA*Δ*sfp* (*spd_0558, sp_2239, sp_1954*)	Km^r, Erm^r, Cm^r, Spec^r	*cbpG$^+$*	This work
PN693	TIGR4Δ*cps*Δ*prtA*Δ*cbpG*Δ*sfp* (*spd_0558, sp_0390, sp_1954*)	Km^r, Erm^r, Cm^r, Spec^r	*htrA$^+$*	This work
PN695	TIGR4Δ*cps*Δ*htrA*Δ*prtA*Δ*cbpG* (*sp_2239, spd_0558, sp_0390*)	Km^r, Erm^r, Cm^r, Spec^r	*sfp$^+$*	This work
PN315	TIGR4*lux*	Km^r	-	Schulz et al., 2014 [380]
PN675	TIGR4*lux*Δ*sfp* (*sp_1954*)	Km^r, Spec^r	-	This work
PN489	TIGR4*lux*Δ*htr*A (*sp_2239*)	Erm^r, Km^r	-	This work
PN495	TIGR4*lux*Δ*prt*A (*sp_0641*)	Erm^r, Km^r	-	This work
PN750	TIGR4*lux*Δ*prtA*Δ*cbpG* (*sp_0641, sp_0390*)	Km^r, Erm^r, Spec^r	*htrA$^+$* *sfp$^+$*	This work
PN743	TIGR4*lux*Δ*sfp*Δ*htr*A (*sp_1954, sp_2239*)	Km^r, Spec^r, Cm^r	*cbpG$^+$, prtA+*	This work
PN531	TIGR4*lux*Δ*prtA*Δ*htr*A (*sp_0641, sp_2239*)	Km^r, Erm^r, Cm^r	*sfp$^+$* *cbpG$^+$*	This work
PN747	TIGR4*lux*Δ*htrA*Δ*cbpG*Δ*sfp* (*sp_2239, sp_0390, sp_1954*)	Km^r, Erm^r, Cm^r, Spec^r	*prtA$^+$*	This work
PN686	TIGR4*lux*Δ*htrA*Δ*prtA*Δ*sfp* (*sp_2239, sp_0641, sp_1954*)	Km^r, Erm^r, Cm^r, Spec^r	*cbpG$^+$*	This work
PN752	TIGR4*lux*Δ*prtA*Δ*cbpG*Δ*sfp* (*sp_0641, sp_0390, sp_1954*)	Km^r, Erm^r, Cm^r, Spec^r	*htrA$^+$*	This work
PN760	TIGR4*lux*Δ*htr*AΔ*prt*AΔ*cbp*G (*sp_2239, sp_0641, sp_0390*)	Km^r, Erm^r, Cm^r, Spec^r	*sfp+*	This work

S. pneumoniae 35A mutants:

PN677	35AΔ*sfp*	Spec^r	-	This work
PN666	35AΔ*cbpG*	Erm^r	-	This work
PN487	35AΔ*htrA*	Erm^r	-	This work
PN523	35AΔ*prtA*	Erm^r	-	This work
PN528	35AΔ*prtA*Δ*htrA*	Erm^r, Cm^r	*sfp$^+$,cbpG$^+$*	This work
PN744	35AΔ*htrA*Δ*cbpG*	Erm^r, Cm^r	*sfp$^+$, prtA$^+$*	This work
PN745	35AΔ*prtA*Δ*sfp*	Erm^r, Cm^r	*htrA$^+$,cbpG$^+$*	This work
PN688	35AΔ*htrA*Δ*prtA*Δ*sfp*	Erm^r, Cm^r, Spec^r	*cbpG$^+$*	This work
PN748	35AΔ*htrA*Δ*cbpG*Δ*sfp*	Erm^r, Spec^r, Cm^r	*prtA$^+$*	This work
PN741	35AΔ*prtA*Δ*htrA*Δ*cbpG*	Erm^r, Cm^r, Spec^r	*sfp$^+$*	This work
PN753	35AΔ*prtA*Δ*sfp*Δ*cbpG*	Erm^r, Cm^r, Spec^r	*htrA$^+$*	This work

S. pneumoniae D39 mutants:

PN742	D39Δ*lux*Δ*sfp*Δ*htrA*	Km^r, Spec^r, Cm^r,	*cbpG$^+$, prtA$^+$,*	This work
PN749	D39Δ*lux*Δ*prtA*Δ*cbpG*	Km^r, Erm^r, Cm^r,	*htrA$^+$,sfp$^+$*	This work
PN684	D39Δ*lux*Δ*htrA*Δ*prtA*Δ*sfp*	Km^r, Erm^r, Cm^r, Spec^r	*cbpG$^+$*	This work

PN746	D39ΔluxΔsfpΔhtrAΔcbpG	Km^r, Spec^r Cm^r, Erm^r	prtA$^+$	This work
PN740	D39ΔluxΔprtAΔhtrAΔcbpG	Km^r, Erm^r, Cm^r, Spec^r	sfp$^+$	This work
PN751	D39ΔluxΔprtAΔcbpGΔsfp	Km^r, Erm^r, Cm^r, Spec^r	htrA$^+$	This work

[*]Numbering is the stock list of the recombinant pneumococci at Department of Molecular Genetics and Infection Biology, Interfaculty Institute of Genetics and Functional Genomics, Center for Functional Genomics of Microbes University of Greifswald.

[**]Erm: Erythromycin, Km: Kanamycin, Cm: Chloramphenicol, Amp: Ampicillin, R: resistance.

6.1.3 Culture and media used for cultivation of microorganisms

Table 6.5: *S. pneumoniae* and *E. coli* culture media used for cultivation

Microorganisms	Culture media	Composition
S. pneumoniae	Blood agar plates	23 g/l Pepton, 5 g/l NaCl, 10 g/l Agar, 50ml/l sheep blood, pH 7.4 (Oxoid, Wesel, Deutschland)
	THY medium	36.4 g/l Todd Hewitt Broth, 0.5 % yeast extract
	RPMI$_{modi}$	See **Table 6.8** and **Table 6.9**
	CD medium	See **Table 6.6** and **Table 6.7**
E. coli	LB agar plates	1% Bacto-tryptone, 0.5 % yeast extract, 0,5% NaCl, 1,2% Agar
	LB medium	1% Bacto-Trypton, 0.5 % yeast extract, 0.5 % NaCl, pH 7.5

Table 6.6: Chemically defined medium (CDM)*

Solutions	Ingredients	Final conc. [g/L]	Vol [ml]	Remarks
1a	FeSO$_4$ x 7H$_2$O	0.504	50	Dissolve with 1N HCL at 55°C in water bath, store at 4°C
	(FeNO$_3$)$_2$ x9H$_2$O	0.106		
1b	K$_2$HPO$_4$	20	50	Readily dissolve, Store at 4°C
	KH$_2$PO$_4$	100		
1c	MgSO$_4$ x 7H$_2$O	70	50	Readily dissolve, Store at 4°C
	MnSO$_4$ x H$_2$O	558		
3a	p-aminobenzoic acid	0.02	500	1. Resolve all the components except of biotin and folic acid, and adjust the pH to 6.9
	Biotine	0.02		
	Folic acid	0.08		
	Nicotinamide	0.1		2. Dissolve the biotin in 2ml PBS and folic acid in 5ml 2.5 NaOH and mix all the three solutions, Store at 20°C
	Nicotinamide adenine dinucleotide	0.25		
	Pathothenate calcium salt	0.2		

	Pyridoxal	0.1			
	Pryidoxamine 2HCL	0.1			
	Thiamine hydrochloride	0.1			
	Vitamin B12	0.01			
3b	Riboflavin	0.04	400	Store at -20°C	
4	Glucose or galactose	100	500	Sterile filtration	
5	Adenine	4	500	Resolve first in 200 ml 2N	
	Uracil	4		HCl at 55°C	
6a	$CaCl_2$ x $2H_2O$	0.2	500	Sterile filtration, store at	
	Choline chloride	0.2		RT.	
6b	NaH_2PO_4 X H_2O	27.2	500	Sterile filtration, store at	
	$NaHPO_4$ x $2H_2O$	32		RT.	
	$NaC_2H_3O_2$ x H_2O	25			
6c	Cystine	50	20	Resolve in 20 ml 1N HCL, Sterile filtration, store at RT.	

*The CDM medium was prepared without the essential amino acid solution **I** and **II**, supplemented instead with 2% casein hydrolysate. **L-amino acid I:** alanine, asparagine acid, asparagine, glutamine acid, glycine, isoleucine, leucine, lysine, proline, hydroxyproline, serine, valine, glutamine, and threonine
L-amino acid II: arginine, histidine, methionine, phenylalanine, tryptophan and tyrosine.

Table 6.7: Composition of 1liter of CDM

Stock solution	Volum [ml]
dH_2O	608
1a	10
1b	10
1c	10
3a	10
3b	40
4	100
5	10
6a	100
6b	100
6c	2

Table 6.8: Supplements of the RPMI$_{modi}$ medium

Solutions	Components	Concentration [g/l]	Volume [ml]	Remarks
Adenine/Uracil-solution	Adenine	0.4	100	Dissolved in 40ml 1N HCL at 90°C. Storage at RT.
	Uracil	0.8		
Buffer solution	Glucose	74	1000	Sterile filtration, and storage at 4°C
	NaHCO$_3$	24.7		
	Glycine	1.11		
	Choline chloride	0.456		
	NaH$_2$PO$_4$ x H$_2$O	3.195		
	Na$_2$HPO$_4$ x H$_2$O	9.22		

Table 6.9: Preparation of RPMI$_{modi}$ medium

Buffers	Volume [ml]	Remarks
Buffer solution	40.54	Sterilize filtration
Adenine/Uracil	5	Re-heat at 90°C before use

The supplements solution was added to 500ml of RPMI-1640 medium with glutamine (HyClone™, Germany). The solution were mixed gently and were stored at 4°C

6.2 Antibiotics

Table 6.10: Final concentration of the antibiotics used for the selection of the microorganisms

Antibiotics	*E. coli* [µg/ml]	*S. pneumoniae* [µg/ml]	Source
Ampicillin	100	-	Applichem
Kanamycin	50	50	Roth
Erythromycin	250	5	Sigma-Aldrich
Spectinomycin	600	100	Applichem
Chloramphenicol	30	5	Merck
Penicillin G	-	100 Units/ml	Sigma-Aldrich
Gentamicin	-	100	Sigma-Aldrich

6.3 Plasmids and oligonucleotides

Table 6.11: Constructed plasmids

No[*]	Vector name	Properties	Resistance[**]	Reference
SP72		Cloning vector	Ap^r	Promega
pSP72D		Cloning vector for PCR products (derivative of pSP72)	Ap^r	G. Burchhardt
1019	pUC18Δ*prtA*::Erm^r	pUC18 vector with *sp_0641* (*prtA*) gene partial deleted replaced with Erm^r resistance gene cassette, (4861 bp.)	Ap^r, Erm^r	N. Henck
1101	pSP72DΔ*cbpG*::Spec^r	pSP72D vector with *sp_0390* (*cbpG*) gene partial deleted replaced with Spec^r resistance gene cassette, (4375 bp.)	Ap^r, Spec^r	A. Wüst
1100	pSP72DΔ*cbpG*::Erm^r	pSP72D vector with *sp_0390* (*cbpG*) gene partial deleted replaced with Erm^r resistance gene cassette, (4313 bp.)	Ap^r, Erm^r	A. Wüst
991	pSP72DΔ*htrA*::Cm^r	pSP72D vector with *sp_2239* (*htrA*) gene partial deleted replaced with Cm^r resistance gene cassette, (4923 bp.)	Ap^r, Cm^r	N. Henck
1131	pSP72DΔ*sfp*::Cm^r	pSP72D vector with *sp_1954* (*sfp*) gene partial deleted replaced with Cm^r resistance gene cassette, (4926 bp.)	Ap^r, Cm^r	R. Bolsmann
1132	pSP72DΔ*sfp*::Spec^r	pSP72D vector with *sp_1954* (*sfp*) gene partial deleted replaced with Spec^r resistance gene cassette, (5029 bp.)	Ap^r, Spec^r	R. Bolsmann
1119	pSP72DΔ*sfp*::Erm^r	pSP72D vector with *sp_1954* (*sfp*) gene partial deleted replaced with Spec^r resistance gene cassette, (4968 bp.)	Ap^r, Erm^r	R. Bolsmann

[*]Numbering is the stock list of the Department of Molecular Genetics and Infection Biology, Interfaculty Institute of Genetics and Functional Genomics, Center for Functional Genomics of Microbes University of Greifswald.
[**]Erm: Erythromycin, Km: Kanamycin, Cm: Chloramphenicol, Amp: Ampicillin, R: resistance.

Table 6.12: Oligonucleotides

Primer intended use	Primer No*	Primer name/ restriction enzyme	Sequence (5'-3')**
Primers used for the amplification of antibiotic resistance genes:			
Erythromycin (*ermB*)	99	MC_Erm^r_F (*ermB*) (*EcoRI*)	5'-CCCGGGGAAATTTTGATATCGATGGATCCGAATTCGACGGTTCGTGTTCGTGCTG-3'
	100	MC_Erm^r_R (*ermB*) (*EcoRI*)	5'-CCCGGGGAAATTTTGATATCGATAAGCTTGAATTCCCGTAGG CGCTAGGGACCTC-3'
Spectinomycin (*aad9*)	118	Spec^r_R (*aad9*) (*HindIII*)	5'-AAAAGCTTGCTAGCAATTAGAATGAATATTTCCC-3'
	117	Spec^r_F (*aad9*)	5'-GTACAGGATCCGAATTCATCGATTTTCGTTCGTGAATAC-3
Chloramphenicol (*cat*)	158	BM19 Cm^r_F (*cat*) (*EcoRI*)	5'-GCGCGAATTCGAAAATTTGTTTGATTTTTAATGG-3'
	159	BM18 Cm^r_R (*cat*) (*SacI*)	5'-ATATGAGCTCGGGTTCCGAGGCTCAACGTCAA-3'
Chloramphenicol (*cat*)	180	BM19 Cm^r (*cat*) (*BamHI*)	5'-GCGCGGATCCGAAAAT TTGTTTGATTTTTAATGG-3'
	181	BM19 Cm^r (*cat*) (*HindIII*)	5'-GCGCAAGCTTGGGTTCCGAGGCTCAACGTCAA-3'
Primers used for insertion-deletion mutagenesis			
Amplification of *cbpG* (*sp_0390*) 5′ and 3′ flanking region	1421	v*cbpG*_F (*SacI*)	5'-GCGCGAGCTCGAAGGTGG TAGATTTCTTGATTC-3'
	1422	h*cbpG*_R (*SacI*)	5'-GCGCGAGCTCGTAATACA CCATCTTGACC-3 '
Inverse PCR of *cbpG* (*sp_0390*) 5′ and 3′ flanking region (pSP72D vector)	1423	i*cbpG*_R (*BamHI*)	5'-GCGCGGATCCGTGAGCCG CTGTAATTAACAC-3'
	1424	i*cbpG*_F (*HindIII*)	5'-GCGCAAGCTTGGTAAGAT GCTTACAGATTG-3'
Mutations analyse of upstream region *cbpG* (*sp_0390*)	1467	vv*cbpG*_F	5'-GAATGGCTGAACTTAGTAT C-3'
Amplification of *sfp* (*sp_1954*) 5′ and 3′ flanking region	1284	v*sfp*_F (*SacI*)	5'-GCGCGAGCTCGGAGCAGTGTTACAAAATTC-3'
	1285	h*sfp*_R (*SacI*)	5'-GCGCGAGCTCGTTGTGGTAACCTGTTTGC-3'
Inverse PCR of *sfp* (*sp_1954*) 5′ and 3′ flanking region (pSP72D vector)	1286	i*sfp*_R (*BamHI*)	5'-GCGCGGATCCCGCTAGTCTGAGTGTGAG-3'
	1287	i*sfp*_F (*HindIII*)	5'-GCGCAAGCTTGATCAGCCCTATAATTATATG-3'
Mutations analyse of upstream region *sfp* (*sp_1954*)	1416	vv*sfp*_F	5'-GAACCTAATATTGGTTCAATAG-3'
	1061	*BVhtrA*_F (*BamHI*)	5'-CGCGCGGATCCAGTCAATTTTCTATTTATG-3'

Amplification of *htrA* (*sp_2239*) 5´ and 3´ flanking region	1062	PH*htrA*_R (*PstI*)	5'-CTCA<u>CTGCAG</u>AAGAGCTTCTAATTTCC-3'
Inverse PCR of *htrA* (*sp_2239*) 5´ and 3´ flanking region (pSP72D vector)	1063	IH*htrA*_R (*EcoRI*)	5'-CATGCG<u>GAATTC</u>GCTAATGACGATAACGAC-3'
	1064	IH*htrA*_F (*EcoRI*)	5'-GCGC<u>GAATTC</u>CTTAACAAGAGTTCAGGTG-3'
Mutations analyse of upstream and downstream region *htrA* (*sp_2239*)	1088	vv*htrA*_F	5'-CCAGCTTTGCTATTATATTG-3'
	1089	hh*htrA*_R	5'-ACAGCCTTATTTCAGGCTG-3'
Amplification of *prtA* (*sp_0641*) 5´ flanking region	1073	v*prtA*_F (*BamHI*)	5'-GCAC<u>GGATCC</u>TTAAGCCTTACTCTTAGCG-3'
	1074	v*prtA*_R (*SacI*)	GCGC<u>GAGCTC</u>ATAAACTTTAACTTTGCTAGC
Amplification of *prtA* (*sp_0641*) 3´ flanking region	1075	h*prtA*_F (*SacI*)	GCGA<u>GAGCTCG</u>TTTATGTACTGAGATTAGATAG-3'
	1076	h*prtA*_F (*SalI*)	5'-GCGA<u>GTCGAC</u>CACTTTCAGAATAAGGAGCCTG

*Numbering is the stock list of the Department of Molecular Genetics and Infection Biology, Interfaculty Institute of Genetics and Functional Genomics, Center for Functional Genomics of Microbes University of Greifswald.
**The primers were synthesized by Eurofins MWG Operon, Germany. The restriction sites used for cloning are underlined.

6.4 Eukaryotic cells and culture media

6.4.1 Eukaryotic cell lines

Table 6.13: Characteristics and origin of human and murine cells used in this study

Cells	Properties	Source	Reference
Detroit-562	Human nasopharyngeal epithelial	ATCC CCL-138	Peterson et al., 1968
J774A.1	Murine macrophages (BALB/c)	DSMZ, Braunschweig	Ralph et al., 1975

6.4.2 Culture media for eukaryotic cells

Table 6.14: Culture media for cell lines

Cells	Medium
Detroit-562	RPMI-1640 (HyClone™), 10 % (v/v) heat inactivated FCS, 2 mM glutamine, 1 mM sodium pyruvate, 1 mM sodium pyruvate, 1% HEPES, 0.1 mg/ml streptomycin, 100 units/ml penicillin
J774A.1	RPMI-1640 (HyClone™), 10 % (v/v) heat inactivated FCS, 2 mM glutamine, 1 mM sodium pyruvate, 1 mM sodium pyruvate, 0.1 mg/ml streptomycin, 100 units/ml penicillin

Table 6.15: Cell culture supplements for the cultivation of eukaryotic cells

Additive	Components	Source
FBS	Fetal Bovine Serum	Gibco
Penicillin/Streptomycin	0.1 mg/ml Streptomycin, 100 Units/ml Penicillin	PAA, Laboratories GmbH
Trypsin/EDTA	5 mg/ml Trypsin, 2.2 mg/ml EDTA	PAA, Laboratories GmbH
PBS/EDTA	2.2 mg/ml EDTA dissolved in PBS pH 7.4, sterilized by filtration	
Freeze medium	Cell-culture medium (HyClone™), 10% FBS, 10% DMSO	

6.5 Mouse strains, pharmacological and immunomodulatory substances

8-10-week old female CD-1® outbred mice from Charles River (Sulzfeld, Germany) were used. Mice were anesthetized intraperitoneally with ketamine (Ketanest S; Pfizer Pharma, Karlsruhe, Germany) and xylazine (Rompun®; Provet AG, Lyssach, Germany).

Table 6.16: Pharmacological substances

Substances	Source
Ketamine, KetanestS	Pfizer pharma, Karlsruhe,Germany
Xylazine, Rompun®	Provet AG, Lyssach, Germany
Isoflurane CP®	Baxter

6.6 Enzymes used in this study

Commercially available enzymes used with the respective buffers, as described by the manufacturer.

Table 6.17: Enzymes

Enzyme	Source
Dream *Taq*-DNA-Polymerase	Theromo Fisher
Pfu-DNA-Polymerase, recombinant	Department of Molecular Genetics and Infection Biology, *E. coli* clone E162*
Taq- DNA-Polymerase	Department of Molecular Genetics and Infection Biology, *E. coli* clone E175*
Phusion high fidelity polymerase	Thermo Fisher
T4 DNA ligase	NEB/Thermo Fisher
Lysozyme	Serva
Hyaluronidase Type IV-S	Sigma-Aldrich
Tobacco Etch virus (TEV), protease	Promega

*Numbering is the stock list of the Department of Molecular Genetics and Infection Biology, Interfaculty Institute of Genetics and Functional Genomics, Center for Functional Genomics of Microbes University of Greifswald.

Table 6.18: Restriction enzymes

Enzyme	Recognition site (5'-3')*	Source
*Eco*RI	G↓AATTC	NEB/Thermo Fisher
*Hin*dIII	A↓AGCTT	NEB/Thermo Fisher
*Sac*I	GAGCT↓C	NEB/Thermo Fisher
*Bam*HI	G↓GATCC	NEB/Thermo Fisher
*Pst*I	CTGCA↓G	NEB/Thermo Fisher
*Nde*I	CA↓TATG	NEB/Thermo Fisher
*Ecl*136II	GAG↓CTC	NEB/Thermo Fisher

Table 6.19: Purified recombinant proteases

Protein	*E. coli* clones	Expression conditions	Column for purification
His_6-*cbpG*	BL21 (DE3) Arctic Express	1 mM IPTG/ 12 °C/ 16 h	HisTrap^{TM} FF Crude, 1 ml (GE Healthcare)
His_6-*cbpG*$^{32-184aa}$	BL21 (DE3) Arctic Express	1 mM IPTG/ 12 °C/ 16 h	HisTrap^{TM} FF Crude, 1 ml (GE Healthcare)
His_6-MBP-*htrA*	BL21 (DE3)	1 mM IPTG/ 27 °C/ 2 h	HisTrap^{TM} FF Crude, and MBPTrap^{TM}1 ml (GE Healthcare)
His_6-*prtA*$^{28-942aa}$	BL21 (DE3)	1 mM IPTG/ 30 °C/ 3 h	HisTrap^{TM} FF Crude, 1 ml (GE Healthcare)
His_6-*sfp*	BL21 (DE3)	1 mM IPTG/ 27 °C/ 3 h	HisTrap^{TM} FF Crude, 1 ml (GE Healthcare)

6.7 Antibodies

Table 6.20: Antibodies

Antibodies	Dilution	Source
Alexa-Fluor® 488 goat-anti-mouse IgG	1:1000	Invitrogen
Alexa-Fluor® 568 goat-anti-mouse IgG	1:1000	Invitrogen
Mouse anti-pneumococci IgG	1:1000	Davids Biotechnologie
Mouse-monoclonal Penta-His™	1:600/1:2000	QIAGEN
Mouse anti-MBP IgG	1:1000	AG Hammerschmidt

6.8 Commercial kits, columns and markers

Table 6.21: Commercial kits and columns

Commercial kits*	Source
Molecular Biology	
Wizard® Plus SV Minipreps DNA Purification System	Promega
Wizard® SV Gel and PCR Clean-Up System	Promega
Protein purification and cell biology	
HisTrap™ Ni-NTA column, FF Crude, 1 ml	GE Healthcare
Protino® Ni-TED 1000 column	Macherey-Nagel
PD-10 column	GE Healthcare
Vivaspin 20 and 6	Sartorius
Phalloidin-iFlour®-594 conjugate 1:1000	Abcam

Kits were used according to respective manufacturer's instructions

Table 6.22: DNA and protein ladders used for electrophoresis

Markers	Size (bp /kDa)	Source
Molecular biology:		
GeneRuler™ 1 kb DNA ladder	250, 500, 750, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 5000, 6000, 8000, 10,000 [bp]	Fermentas
Protein analysis:		
PageRuler™ prestained protein ladder	10, 15, 25, 35, 40, 55, 70, 100, 130, 170 [kDa]	Fermentas
PageRular ™ unstained Protein Ladder	10, 15, 25, 35, 40, 55, 70, 100, 130, 250 [kDa]	Fermentas

6.9 Solutions, and buffers

Table 6.23: Buffers and solutions used for molecular biology experiments

Buffer/solutions	Compositions/application	Source
TE-buffer	20 mM Tris-HCl, 1 mM EDTA, pH 7.0	Department of
TES-buffer	50mM Tris-HCl, 5 mM disodium EDT, 10mM NaCl, pH 8.0	Molecular Genetics and Infection Biology
TAE buffer (20x)	800 mM Tris-HCl, 200 mM EDTA, 2.28% glacial acetic acid, distilled water, pH 8.0	
Agarose (0.8%)	Agarose 2.4g in 300ml 1xTAE buffer	Genaxxon bioscience
HD-green plus DNA stain	Dissolve 18μl of HD green in 300ml agarose solution	INTAS
DNA loading buffer	Bromophenol blue 0.25%, sucrose 40% dissolved in distilled water	fermenters
Gel loading dye purple	2.5% Ficoll-400, 10mM EDTA, 3.3 mM Tris-HCl (pH 8.0 at 20°C), 0.02%	NEB
Lysozyme, stock solution	10 mg/ml, dissolved in TES-buffer	
Phenol		Roth
RNase A- stock solution	5 mg/ml, dissolved in TES-buffer	
Pronase E	5 mg/ml, dissolved in TES-buffer	
N-Lauroylsarcosine	10% in 250 mM disodium-EDTA	
Phenol-chloroform-isoamyl,	25:25:1	Roth

Table 6.24: Buffers and chemicals used for protein separation

Buffer/solutions	Compositions/ Source
Sodium Dodecyl Sulfate Polyacrylamide Gel (SDS-PAGE)	
Solution for SDS-Gel	Acrylamide (30%), Tris (pH 8.8) 1.5M, SDS 10%, APS 10%
	Tris-(pH 6.8) 0.5M, TEMED
Resolving gel buffer	Tris-1.5M pH 8.8
Stacking gel buffer	Tris-0.5 pH 6.8
10% SDS-solution	10% SDS in 100ml distilled water
10% Ammonium per sulfate (10% APS)	10g APS in 100ml distilled water, store at -20°C. Polymerization of bisacrylamide inducer
SDS gel-loading buffer	
2x Protein sample buffer	dH$_2$O-6ml, stacking gel buffer pH 6.8-4ml, 10% SDS-4ml, Glycerol-4ml, B-mercaptoethanol-2ml, and a pinch of bromophenol blue. Loading-dye for protein electrophoresis (SDS-PAGE)
5x protein sample buffer	dH$_2$O-6ml, stacking gel buffer (pH6.8) 10ml, 10% SDS-10ml, Glycerol-5ml, β-mercaptoethanol-2ml, and a pinch of bromophenol blue.
10x running buffer	Tris-230mM, Glycine-1.92M, SDS 1%.

Table 6.25: Ingredients used for preparation of polyacrylamide gels

Ingredient	Resolving gel (10%, 10ml)	Resolving gel (12%, 10ml)	Stacking gel (4%, 5ml)
A. dest	4 ml	3.3 ml	3.4 ml
Acrylamide	3.3 ml	4 ml	830 μl
Resolving-stacking gel buffer	2.5 ml	2.5 ml	630 μl
SDS (10% (w/v))	100 μl	100 μl	50 μl
APS (10% (w/v))	100 μl	100 μl	50 μl
TEMED	4 μl	4μl	5 μl

Table 6.26: Buffers and chemicals used for protein biochemical work

Buffer/chemicals	Compositions/Applications
EDTA	5mM dissolved in distilled water
IPTG stock	Dissolve 2.38g IPTG (1M) in 10ml distilled water, store at -20°C
Lysozyme solution	1mg/ml lysozyme in distilled water. Bacterial lysis
BSA (100μg/ml)	dissolve BSA in 1xPBS pH 7.4. Protein standard
DTT-stock	100mM in dH2O, Store at -20°C. Disulfide reducing agent
DNase-stock solution	1mg/ml DNase in 10Mm MgSO$_4$.
Bradford reagent	Sigma

Table 6.27: Buffer and solutions for protein staining

Buffer/solutions	Compositions/Applications
Coomassie brilliant blue staining Protein staining (SDS-PAGE)	Ethanol 25%, acetic acid 10%, R250 Coomassie brilliant blue^{TM} 0.05%
Commassie brilliant blue staining Destaining solution	Ethanol 25%, acetic acid 10%
Silver staining solution 1. **Fixation**	Ethanol 50%, acetic acid 12%, Formaldehyde (37%) 500µl/L
2. **Washing**	Ethanol 50%
3. **Sensitization**	$Na_2S_2O_3$ x $5H_2O$- 0.2g/L
4. **Silver staining**	AgNO3- 2g/L, Formaldehyde (37%), distilled water
5. **Development**	Na_2CO_3-30g/L, $Na_2S_2O_3$ x $5H_2O$- 4mg/L, Formaldehyde (37%)-500µl/L, and distilled water
6. **Stope reagent**	Glycine – 1%, and distilled water

Table 6.28: Buffer used for protein purification

Systems/Buffers	Compositions
Buffers for affinity chromatography by HisTrap^{TM} FF column 1ml	
Tris-binding buffer	Tris HCl-100mM, NaCl-500mM, Imidazol-20mM, pH 7.4
Tris-elution buffer	Tris HCl-100mM, NaCl-500mM, Imidazol-500mM, pH 7.4
Tris buffer for rCbpG by HisTrap^{TM} FF column 1ml	
Tris-binding buffer	Tris HCl-100mM, NaCl-500mM, Imidazol-5mM, Choline chloride-250mM, Glycerol-10%, pH 7.4
Tris-elution buffer	Tris HCl-100mM, NaCl-500mM, Imidazol-500mM, Choline chloride-200mM, Glycerol-10%, pH 7.4
Buffer for affinity chromatography MBPTrap^{TM} column 1ml	
Binding buffer	Tris 20mM, NaCl 500mM, EDTA 1mM, DTT 1mM, pH 7.4
Elution buffer	Binding buffer + 10Mm Maltose
Buffer used for TEV-protease purification:	
Lysis buffer	300mM NaCl, 50mM NaH_2PO_4, 20 mM Imidazole, pH 7.4
Elution buffer	300mM NaCl, 50mM NaH_2PO_4, 500mM Imidazole, pH7.4
Buffer for His-trap FF crude column regeneration	
Stripping buffer	$NaPO_4$ 20mM, NaCl 500mM, EDTA 50mM, pH 7.4
$NiSO_4$	100mM $NiSO_4$ in 100ml dH_2O, filter and sonication

Table 6.29: Buffers and solutions used in cell biological work

Buffer/solutions	Compositions/Applications
1 x PBS	137 mM NaCl, 2,7 mM KCl, 80 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4
PFA-stock solution	37% Paraformaldehyde (v/v) in dest. H_2O
Citrate buffer	100 mM Na-Citrate, dissolved in dest. H_2O
Sodium carbonate buffer	100 mM Na_2CO_3, 100 mM $NaHCO_3$, pH 9.2
Mowiol	20g Mowiol 4-80 in 80 ml 1 x PBS
Permeabilization solution	1% Saponin (w/v), dissolved in RPMI-1640
DIF-Permeabilization	1% Triton X-100 in 1 x PBS
DIF-Fixation solution	2% PFA in 1 x PBS
DIF-Blocking solution	10% FBS in 1 x PBS

Chapter **7**

METHODS

7. Methods

7.1 Microbiological Methods

7.1.1 Cultivation of *E. coli*

The recombinant *E. coli* used are listed in **Table 6.1** and **Table 6.2**. *E. coli* were activated from cryopreservation stocks (-80°C) on LB agar plates and were incubated overnight at 37°C. Afterwards, bacteria were grown in liquid LB medium supplemented with corresponding antibiotic at 37 °C under shaking conditions (110-120 rpm). The bacterial cultures were incubated at 110-120 rpm and 37°C on a shaking incubator.

7.1.2 Growth and cultivation of *S. pneumoniae*

All *S. pneumoniae* strains and mutants used are listed in **Table 6.3** and **Table 6.4**. The cultivation of *S. pneumoniae* was conducted by plating bacteria from glycerol-stocks on blood agar plates (Oxoid, Germany) without antibiotics. The plates were incubated at 37°C and 5% CO_2 for 6-8h. Afterward, bacteria were transferred on blood agar plates supplemented with the appropriate antibiotics and incubated at 37°C and 5% CO_2 for up to 10h. For the cultivation of the main culture, strains were grown in liquid media such as complex Todd-Hewitt medium supplemented with 0.5% yeast extract (THY) (Roth, Germany), minimal medium $RPMI_{modi}$ (**Table 6.8**, and **Table 6.9**) or in CDM (composition in **Table 6.6** and **Table 6.7**) in a water bath at 37°C. The bacteria were incubated for most of the experiments until the mid-log phase (OD_{600nm} = 0.35-045). Bacterial growth was monitored by measuring the optical density at 600nm (OD_{600nm}) for up to 10h.

7.1.3 Cryopreservation

For long-term use and storage of the bacterial strains, pneumococcal strains freshly grown on blood agar plates or overnight cultures of *E. coli* were resuspended in the appropriate culture medium, supplemented with 20% glycerol (v/v) and frozen for cryopreservation at -80°C.

7.1.4 Transformation of *E. coli*

Transformation of *E. coli* with recombinant plasmids was performed using $CaCl_2$-treated competent *E. coli* cells. Aliquots of competent *E. coli* were gently thawed on ice. Then, bacteria were mixed with DNA or ligation reaction and incubated for 10 minutes on ice. A heat shock

followed this for 5 minutes at 37°C and incubation on ice for two minutes. Afterward, 600µl pre-warmed LB medium was added, and the transformation mixture was incubated for 45-60 minutes in a shaking incubator at 30°C. Positive colonies were selected by plating aliquots on a selective LB medium with appropriate antibiotics and were incubated at 30°C overnight.

7.1.5 Transformation of *S. pneumoniae*

S. pneumoniae is a naturally competent bacterium that can uptake extracellular DNA from the surrounding environment. The DNA uptake can be facilitated using competence-stimulating peptides (CSP)[423]. The most effective competence was obtained from OD_{600nm} 0.1-0.15 in the early exponential growth phase. *S. pneumoniae* transformations were performed as described [424]. Pneumococci strains were cultured in 10 ml THY medium with an initial OD_{600nm} 0.05-0.06 at 37°C in a water bath. After reaching the desired optical density of 0.1-0.15, 200µl of the culture were transferred into a new 1.5ml tube. After adding the strain-specific CSP (1µg, CSP1 for D39) and the DNA (1µg), pneumococci were incubated on ice for 10 min. Afterward, bacteria were incubated for 30 minutes at 30°C and then 2h at 37°C. Finally, positive colonies were selected from blood agar plates containing appropriate antibiotics (after overnight incubation at 37 °C and 5% CO_2). Grown colonies were analyzed by colony PCR. The pneumococcal isolate EF3030 serotype 19F (kindly provided by Anders P. Håkansson, Lund University, Sweden) was transformed using another protocol. Briefly, pneumococci were grown in 1 ml THY competence medium (THY supplemented with 0.2% bovine serum albumin (BSA), 0.2% glucose, and 0.02% $CaCl_2$), and incubated at 37°C in a water bath. The bacterial culture was diluted 1:25 ratio into 200µl fresh competence medium after reaching OD_{600nm} 0.4. Subsequently, 200ng of CSP1 was added, and the bacteria incubated for 15 minutes at 34°C. DNA (10ng-1000ng) was added to the pneumococci and incubated for 30 minutes at 34°C followed by 2.5-3h at 37°C. Finally, bacteria were plated on blood agar plates with appropriate antibiotics and incubated at 37°C and 5% CO_2 for 24 or 48h.

7.1.6 Growth studies of *S. pneumoniae*

The growth behavior of *S. pneumoniae* wild-type and isogenic mutants were studies in different growth media. Pneumococcal strains were plated from glycerol stocks (-80°C) on blood agar plates without antibiotics and incubated for 8h at 37°C and 5% CO_2. Afterwards, grown bacteria were transferred to blood agar plates supplemented with antibiotics and incubated for a maximum of 8h overnight. The following day, pneumococci were collected using a sterile cotton swab and resuspended in a 35-40ml liquid growth medium (THY, $RPMI_{modi}$, or CDM) and incubated at 37°C in a water bath without agitation. The initial optical density (OD_{600nm}) was adjusted to 0.08-0.1. The optical density was measured every hour for up to 10h. Afterwards, the generation time ("g") of the strains was calculated using the formula below:

$$\mu \ [\text{min}^{-1}] = (\ln X \ (t_2) - \ln X \ (t_1)) \setminus (t_2 - t_1)$$

$$g \ [\text{min}] = \ln 2\mu \setminus \mu$$

μ = Growth rates
$X \ (t1)$ = OD_{600nm} at time t1
$X \ (t2)$ = OD_{600nm} at time t2
g = Generation time

7.1.7 Determination of colony-forming unit (CFU)

In order to perform the *in vitro* and *in vivo* infection experiments with an accurate number of bacteria, the colony-forming units (CFU) per ml were determined. Therefore, pneumococci were grown in 35-40 ml THY medium to mid-log phase (OD_{600nm} = 0.35-04). After centrifugation (3000 xg for 6 minutes), the bacterial pellet was resuspended in 1ml PBS buffer. The bacterial suspension was then adjusted to an OD_{600} of 1.0, corresponding to $\sim 1 \times 10^9$ CFU/ml. The pneumococcal suspension was serially diluted up to 10^7 and 0.1 ml plated on blood agar plates, followed by incubation at 37°C and 5% CO_2. Colonies were counted, and CFU per ml as determined by integrating the corresponding dilution factor in the formula:

$$CFU \ per \ ml = \frac{Average \ number \ of \ colonies \ \times dilution \ factor}{0.1}$$

7.2 Molecular biological methods

7.2.1 Isolation of plasmid DNA from *E. coli*

Plasmid DNA was isolated from 15ml overnight culture, grown at 37°C and 120 rpm on a shaking incubator (as described under 7.1.1). According to the manufacturer's instructions, the overnight culture was harvested by centrifugation and DNA was isolated using the Wizard® Plus SV Minipreps DNA purification kit (Promega).

7.2.2 Isolation of the plasmid DNA from *E. coli* by alkaline lysis

Isolation of the plasmids from the recombinant *E. coli* was performed by alkaline lysis procedure (Birnboim)[425]. A full-loop of bacterial colonies or 15 ml cells pelleted bacteria from an overnight culture were resuspended in 200µl lysis buffer-1 (Tris-HCl 50mM, EDTA 10mM, RNase 0.1mg/ml and pH 7.5). Afterwards, 400µl of buffer-2 containing 0.2 M NaOH, 10% SDS was added to the bacterial suspension and mixed gently by inverting. After incubation at room temperature for 4 minutes, 300µl of buffer-3 (3M potassium acetate, pH 4.8) was added and followed by incubation on ice for 5 minutes. The mixture was centrifuged at 12,000xg for 10 minutes at 4°C and supernatant was transferred to a new tube containing 600µl of isopropanol, followed by incubation for 5 min at -20°C. Afterwards, the centrifugation was repeated and the received DNA pellet was washed with 70% ethanol. The suspension plasmid solution was then centrifuged for 4 minutes at room temperature and the supernatant was taken away. The plasmid DNA was later dried for 10 minutes at 45°C in a vacuum centrifuge. The plasmid DNA was resuspended into 30µl of sterile dH_2O.

7.2.3 Isolation of genomic DNA from *S. pneumoniae*

The genomic DNA of pneumococci was prepared by using one of the following methods:

a) Pneumococci were cultured in 30 ml THY medium, until $OD_{600nm}= 0.6$-0.8. Cultures were centrifuged for 10 minutes at 2956xg. The cell pellet was resuspended in TES buffer and 1ml of lysozyme (5mg/ml) was added, followed by incubation for 30 minutes at 37°C. The bacterial lysate was again incubated for another 15 minutes at 37°C after adding 50µl of Pronase E (5mg/ml) and RNase A (5mg/ml). Subsequently, 500µl (10%) of *N*-Lauroylsarcosine was applied and incubated for 30 minutes at 37°C in water bath. After the incubation time, the proteins were precipitated by adding 3 ml of the phenol and phenol/chloroform/isoamyl solution, and followed by centrifugation at 12,000xg for 15 min. DNA was extracted by adding 1ml of 3M sodium acetate and isopropanol treatment,

followed by seconds centrifugation at 13,000xg for 15 min. The supernatant was discarded, and isolated genomic DNA was washed with 70% cold ethanol and the sedimented DNA dissolved into 100µl sterile dH2O. The concentration of the genomic DNA was determined by measuring the absorbance at 260nm using NanoDropTM-1000 (PeqLab). In addition, sample of the isolated DNA was separated by agarose gel electrophorese to determine the integrity of the DNA.

b) *S. pneumoniae* whole genomic DNA was extracted by growing the pneumococci in 10ml THY medium to mid logarithmic phase (OD_{600nm} 0.3-0.4). Afterwards, the culture was centrifuged at 12,000xg for 3 minutes. The bacterial pellet was resuspended in a 100µl PBS buffer followed by centrifugation at 16,000xg for 4 minutes. The bacterial pellet was resuspended in 50µl dH_2O and heated at 96°C for 8 minutes to extract the DNA. Afterwards, the bacterial lysate was cooled down to room temperature and another centrifuge step was conducted at 16,000 xg for 2 minutes. 2µl of the supernatant was used as a template for the colony PCR.

7.2.4 Digestion and ligation of DNA fragment

Restriction endonucleases (**Table 6.18**) were used to digest DNA fragments and plasmids. According to the manufacturer instructions, the DNA digestion was conducted for 60 minutes at 37°C. T4 DNA ligase (New England, Biolabs) was used to ligate the target DNA fragment with a cleaved vector overnight at 16°C in a ligation chamber (Grant-Boekal). In a reaction volume of 10-30l, the quantity ratio of DNA fragment to vector was 3:1.

7.2.5 Polymerase chain reaction (PCR)

The amplification of DNA fragments was conducted using specific oligonucleotides (**Table 6.12**) in a polymerases chain reaction (PCR)[426]. Different thermostable DNA polymerase and reaction conditions were selected as recommended by the supplier. *Taq*-DNA-polymerase or *Pfu*-DNA-polymerase was used for DNA amplification.

Table 7.1: Standard PCR reaction mixture [50 µl]

Components	Stocks	Volume
Reaction buffer	10x	5µl
dNTPS	10mM	1µl
MgCl$_2$	25mM	1µl
Oligonucleotide, forward	20 pmol	1µl
Oligonucleotide, reverse	20 poml	1µl
DNA	~ 100ng	1µl
DNA Polymerase		1µl
A.dest		39µl

Table 7.2: Reaction condition for amplification using *Taq*-DNA-polymerase

Parameters	Temperature	Time	
Denaturation	95°C	2 min	
Denaturation	95°C	30 sec	
Annealing	55-58°C	60 sec	30 cycles
Extension	72°C	1kb/min	
Final extension	72°C	5 min	
Cooling	7°C		

Table 7.3: Reaction condition for amplification with *Pfu*-DNA-polymerase

Parameters	Temperature	Time	
Start	95°C	2 min	
Denaturation	95°C	30 sec	
Annealing	55-58°C	60 sec	30 cycles
Extension	72°C	1kb/**30sec**	
Final extension	72°C	5 min	
Cooling	7°C		

7.2.6   Agarose gel electrophoresis for DNA separation

To separate DNA based on its size, agarose gel electrophoresis was used. The agarose gel was prepared by dissolving 2.4 grams of agarose (Genaxxon bioscience) in 300ml of 1x TAE buffer. After boiling of the agarose solution and cooling to 50°C, 10µl of fluorescent dye HDGreen™ was added. Afterwards, 2µl of loading dye was added to 5µl of each reaction and loaded on the poured 0.8% agarose gel. In addition, 5µl of 100 bp gene ruler (NEB) was used as reference to

determine band sizes. Gels were allowed to run at 100volts for 1h and 30 minutes. The agarose gels were visualized in a gel documentation system (Intas).

7.2.7   Construction of pneumococcal serine proteases mutants

Pneumococcal mutants deficient for the serine proteases genes *cbpG*, *sfp*, *htrA*, and *prtA* were generated by insertion-deletion mutagenesis, which includes the allelic substitution of a target gene by an antibiotic resistance gene. The genes, including the upstream and downstream oligonucleotide primer combination and plasmids, are listed in **Table 6.11** and **Table 6.12**. Single, double and triple mutants were generated in different serotypes such as 19F_EF3030 in TIGR4 strains (non-encapsulated Δ*cps*, and bioluminescent *lux* strains) D39 and *S.p.*35A strains. The plasmids constructions for the serine proteases mutants were produced in line with former MSc. thesis Nadine Henck, 2014, MSc., thesis Robert Bolsmann, 2017 and BSc., thesis Andreas Wüst, 2017[424]. Briefly, *cbpG* mutant, the up-and downstream 500 bp of the corresponding *cbpG* was amplified with primers pair combination pair P1421 and P1422 using the TIGR4 chromosomal DNA as a template. The PCR amplified product was cloned into the *Eco*RV-cleaved plasmid pSP72D, which has a modified poly-linker region by deleting the *Xho*I-*Sac*I site of pSP72D. (Promega, Germany). The plasmid containing the *cbpG* gene region was used as a template for deleting 642 bp using the primer pair P1423 and P1424 in an inverse PCR. The primers included a *Hin*dIII and *Bam*HI site for cloning the *ermB* or *aad9* gene cassettes, which were amplified with primers P99 and P100 for *ermB* and p177/118 for *aad9* to generate the plasmids pAW1100 (pSP72D*cbpG*::Erm[r]) and pAW1101 (pSP72D*cbpG*::Spec[r]). While for *sfp* mutant generation, primer pairs P1284 and P1285 were used to amplify the *sfp* gene region, including up-and downstream sequences from TIGR4. Afterwards, the PCR product was cloned to the available cloning vector pSP72D into the *Eco*RV restriction site. To delete the entire *sfp* gene sequence from the recombinant plasmid, the plasmid was used as a template for an inverse PCR with primer pairs P1286 and P1287, which contain *Bam*HI and *Hin*dIII restriction sites. The antibiotic resistance gene cassettes *ermB*, *aad9* or *cat* were ligated separately after digestion with *Bam*HI and *Hin*dIII to generate the plasmids pRB1119 (pSP72D*sfp*::Erm[r]), pRB1132 (pSP72D*sfp*::Spec[r]) and pRB1131 (pSP72D*sfp*::Cm[r]), which were used to remove the entire *sfp* gene in TIGR4. Notably, the *sfp* gene and six upstream and downstream genes of TIGR4 strains were missing in the 19F_EF3030 strain.

To generate the *htrA* mutant, a deletion plasmid was generated using similar strategey. The *htrA* gene was cloned into the *Eco*RV region of cloning vector pSP72D after PCR amplification with

an upstream and downstream region with the primer pairs P1061 and P1062. Afterward, the primer pairs P1063 and P1064 containing the restriction site *Eco*RI were used to delete portions of *htrA* region 1104 bp from the plasmid for the inverse PCR. The antibiotic *cat* gene cassette was amplified with the primer pairs P158 and P159, *Eco*RI digested and cloned into the digested plasmid to construct the *htrA* deletion plasmid pNM991 (pSP72DΔ*htrA*::Cmr).

The construction of *prtA* deficient mutant was conducted via amplification of 500 bp of upstream and downstream of *prtA* gene sequences due to the large size of the gene. The primer combinations P1073 and P1074 (containing *Bam*HI and *Sac*I restriction sites) for the 5'- region, and P1075 and 1076 (containing *Sac*I and *Sal*I restriction sites) for the 3'- region wre used. The *Bam*HI/*Sal*I digested vector pUC18 was used to clone the PCR amplified fragment. After that, the recombinant plasmid was cleaved with *Ecl*136II and replaced with the erythromycin resistance gene cassette *erm*B resulting in plasmid pGB1019 (pUC18Δ*prtA*::Erm^r).

The deletion of the serine protease encoding genes was performed with the generated mutant plasmids, where an antibiotic cassette replaced the gene. After transformation into pneumococci and homologous recombination (*see* Subheading **7.1.5.**), the serine protease encoding genes were inactivated and mutated clones were selected on blood agar with the corresponding antibiotics. The pneumococcal mutants were further characterized by colony PCR (*see* Subheading **7.2.5**). Finally, pneumococci were transformed with the recombinant plasmids pAW1100, pAW1101, pRB1131, pRB1132, pRB1119, pNM991 and pGB1019 to produce double and triple serine protease deficient mutants.

7.3 Biochemical and protein analytical methods

7.3.1 Heterologous expression of pneumococcal serine proteases in *E. coli*

For heterologous expression primers were selected to amplify the corresponding genes without the signal sequence encoding part by PCR. Specific expression vectors (pTP1 or pDB) were used to express the recombinant enzymes. These vectors allow the generation of N-terminal His-tagged fusion proteins. The pTP1 contained a single *Nhe*I site to generate an N-terminal His-tagged fusion to the serine proteases PrtA, SFP and CbpG. In contrast, HtrA has been fused to a maltose-binding protein by cloning the sequence of the *htrA* gene into pDBHisMBP. The heterologous expression of recombinant serine proteases was carried out in *E. coli* BL21 (DE3) or *E. coli* ArcticExpress (DE3). The recombinant clone (**Table 6.2**) was briefly cultivated overnight in 10-15 ml in LB medium supplemented with kanamycin (50μg/ml) on a shaking

incubator. On the following day, the overnight culture was used to inoculate the 500 ml main culture with an initial OD_{600nm} of 0.15. The culture was incubated at 27-30°C and 110-120 rpm until an OD_{600nm} of 0.7-0.8. Protein expression was induced by adding IPTG (isopropyl-β-D-1-thiogalacto-pyranoside) (AppliChem) to a final concentration of 1mM. After the incubation, cultures were incubated in different conditions for 2-3h. Samples were collected prior and after the induction with IPTG to determine the level of the protein expression via SDS-PAGE analysis. (*see* Subheading **7.3.7**) and Coomassie staining of the gel (*see* Subheading **7.3.8**). Due to the inclusion bodies formation in *E. coli* BL21 (DE3) for CbpG protein expression, the *E. coli* ArcticExpress was used to reduce the incubation temperature to 12°C overnight after the induction with IPTG.

7.3.2 Protein purification of heterologous expressed serine proteases by affinity chromatography

The purification of the recombinant serine proteases rCbpG, rPrtA[28-942aa], and rSFP were carried out using HisTrap™ columns. The bacterial pellet of a 500ml culture was resuspended in a 5ml binding buffer supplemented with lysozyme (100μg/ml) and DNase I (Roche) and incubated on ice for 30 minutes. Bacterial cells were lysed by ultrasonic treatment (Branson Sonifier 250, 3x with 30 secs, 30-sec rest, 70% power). After centrifugation at 14000 xg for 30 minutes at 4°C, the supernatant was used for protein purification via ion affinity chromatography (IMAC). The purification of the His_6-tagged serine proteases was performed using the HisTrap™ HP Ni-NTA column with the "ÄKTApurifier liquid chromatography system" (GE Healthcare GmbH) according to the manufacturer instructions. Tris buffer (**Table 6.28**) was used to purify the protein under native conditions. The proteins were eluted from the column by increasing the concentration of imidazole (0-500 mM) in a linear gradient. The rPrtA[28-942aa] protein needed a two-step HisTrap™ purification. For the purification of rHtrA, an MBPTrap purification (*see* subheading **7.3.3**) was used, followed by the second purification via HisTrap. The elution fractions containing the purified protein were loaded on SDS-PAGE gel (*see* subheading **7.3.7**), followed by either Coomassie or silver staining (*see* subheadings **7.3.8** and **7.3.9**). The purified proteins were used to determine the protease activity with casein as a substrate (*see* subheading **7.3.10**).

7.3.3   Purification of rHtrA via MBPTrap^{TM} column

To purify rHtrA, the protein was fused to maltose-binding protein by cloning the gene sequence into the expression vector pDB-His$_6$-MBP. This fusion protein was purified by affinity chromatography using an MBPTrap^{TM} column (1 ml, GE Healthcare). The cultivation and induction of HtrA was performed as previously described (*see* subheading **7.3.1**). Briefly, the heterologous expression of His$_6$-MBP-HtrA in *E. coli* BL21 (DE3) was performed for 2h after induction at 27°C with a final concentration of 1mM IPTG. Bacterial cells were lysed by ultrasonic treatment (Branson Sonifier 250, 3x with 30 secs, 30-sec rest, 70% power). After centrifugation at 14000 xg for 30 minutes at 4°C. The supernatant was loaded on MBPTrap^{TM} column for protein purification. The purification was carried out using an "ÄKTApurifier liquid chromatography system" according to the manufacturer instructions. The fusion proteins were eluted from the column using MBP elution buffer with increasing concentrations of maltose. Due to the autocatalytic cleavage of the fusion protein, the elution fraction from the first purification containing His$_6$MBP and rHtrA was loaded onto a HisTrap™ HP Ni- NTA column for the second purification step to separate rHtrA from the MBP. The purified protein rHtrA was used for protease activity assays with casein (*see* subheading **7.3.10**).

7.3.4   Dialysis of purified serine proteases

Protein dialysis is commonly used for buffer exchange and to remove imidazole or high salt concentrations. The purified protein from the elution fractions was collected in a dialysis tube (16mm, MWCO 12-30,000 Dalton) and incubated for 24h at 4°C in 1l dialysis buffer. The buffer was exchanged two times to ensure proper dialysis.

7.3.5   Buffer exchange of purified protein using a PD-10 column

The PD-10 desalting column (GE Healthcare) was used to change the buffer of purified protein according to the manufacturer's instructions. Briefly, the column was adjusted with 25ml of elution buffer. Protein samples were applied to the column, and the flow-through was discarded. Protein was eluted with a 3.5 ml buffer and the flow-through was collected in 0.5ml fractions on ice. The protein concentration of the fractions was determined by using the Bradford assay.

7.3.6   Determination of protein concentrations

Bradford assay was used to assess the concentration of the purified protein. To measure the protein concentration, 1:200 dilutions were made by mixing 5µl of protein with 495µl PBS

buffer. After adding 500µl Bradford reagent (Sigma-Aldrich), reactions were incubated for 15 min at room temperature in the dark. The absorbance at 595nm was determined photometrically (Denovix). Protein concentration was calculated by using a standard curve (0.5-10µg/ml) of bovine serum albumin (BSA) (Roth).

7.3.7 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE gel electrophoresis [427] was performed to separate proteins and to determine the molecular protein weight. All proteins were separated under denaturation conditions. The gel was prepared by following the standard protocol for protein analysis (**Table 6.25**). Acrylamide concentration was adjusted depending on the molecular weight of the protein (10-12%) and the stacking gel concentration was 4%. Protein samples were prepared for gel electrophoresis by adding a 2x or a 5x (v/v) sample buffer and incubated for 10 minutes at 95°C for denaturation. Gel electrophoresis was performed at a voltage of 100V until the protein band reached the separation gel then voltage was increased to 120-150 volts. The proteins were visualized using either Coomassie brilliant blue or silver staining. Protein molecular weight was determined by comparison of the loaded prestained protein ladder (PagerRuler[TM] Prestained Protein Ladder, (*see* **Table 6.22**).

7.3.8 Coomassie Brilliant Blue staining (CBB)

To visualize the separated proteins in SDS-PAGE gels, Coomassie Brilliant Blue staining (G250) was carried out. The protein gels were incubated with Coomassie dye solution for 3h or overnight at room temperature on an environmental shaker. Afterwards, the gels were distained with Coomassie distaining solution (**Table 6.27**) for 10h until the protein bands became visible. Gels were documented using a computer scanner.

7.3.9 Silver staining

Silver staining is a high-sensitive method to visualize even low amounts of separated proteins in polyacrylamide gels. The method is subdivided into several steps, including protein fixation, sensitization, silver staining and image development (**Table 6.27**). After protein separation, the polyacrylamide gel was incubated in a 50 ml fixation solution for 1-3h before being washed twice in 50% ethanol for 20 minutes each. Subsequently, the sensitizing solution (0.02% sodium thiosulfate) was applied to the gel for 1 minute and the gel was afterward washed three times with distilled water. Afterwards, the silver stain solution (0.2% silver nitrate) was added to the gel and incubated for 20 minutes, followed by three-time washing with distilled water. Finally,

the protein was visualized by incubating the gel with the development solution. The reaction was stopped immediately after bands appeared by adding 1% glycine solution. Finally, the gel was washed with distilled water several times and stored until documentation.

7.3.10 Enzymatic assay with recombinant serine proteases

The purified serine proteases were prepared by dialyzing the isolated proteases with 50 mM-Tris-HCL, 200 mM-NaCl 5 mM-CaCl$_2$, at pH 7.4. The dialysis was performed with three buffer changes at 4°C for 24h. To check enzymatic activity, serine proteases were tested first by using β-casein (Merk) as an unspecific substrate control. Subsequently, the enzymatic assay was conducted at 37°C by incubating the purified recombinant serine proteases with 2µg of different extracellular matrix proteins. The concentration of purified recombinant serine proteases (1-5µg) were used in these protease assays depending on the substrate. Briefly, 5 µg of the His$_6$-tagged serine protease rCbpG and rCbpG^{184aa}, 1.5µg of the rPrtA$^{28-942aa}$ catalytic domain, 5µg of rSFP, and rHtrA were used to determine the catalytic activity in a Tris buffer system containing 50mM-Tris-HCL, 200mM-NaCl and 5mM-CaCl$_2$ in pH 7.4. The proteases were incubated with 20µg β-casein and 2µg of fibronectin (Fn), fibrinogen (Fbg), vitronectin (Vn), factor H (FH), thrombospondin-1 and laminin at 37°C for different time points. To verify that the degradation of casein or the substrate occurred only in the presence of recombinant serine protease, a negative control containing only a buffer was added. Substrate degradation by these recombinant serine proteases was detected by separation of the assay by SDS-PAGE gel (*see* subheading **7.3.7**) and visualization by either Coomassie or silver staining (*see* subheadings **7.3.8, and 7.3.9**).

7.4 Cell biological methods

7.4.1 Sub-cultivation (Splitting) of the eukaryotic cells

All eukaryotic cell lines were handled under sterile conditions. The supplemented cell line medium RPMI-1640 (HyClone™, Germany) was used (*see* **Table 6.15**) for culturing, seeding, and sub-cultivation in sterile cell culture flasks or plates at 37°C and 5% CO$_2$. The cell lines were cultured, sub-cultivated and seeded without using antibiotics. Cryopreserved cell stocks (1 ml aliquots, liquid N$_2$ storage) were thawed at 37°C. Afterwards, cells were diluted with medium and centrifuged for 700xg for 7 minutes. Each aliquot contains approximately 10^5 to 10^6 cells. The cells were passaged 2 to 3 times before seeding in wells for infection experiments, and the medium was exchanged after two days at 37°C and 5% CO$_2$, depending on the density

of the grown cells. For splitting, adherent cells were harvested at 80-90% confluency and adherent cells were detached by incubation with trypsin/EDTA (5ml in T75 cell culture flask) for approximately 5-10 minutes at 37°C incubation. After detachment of the cellular layer from the flask's surface, a cell culture medium dilutes the trypsin/EDTA solution and uniformly suspends the cells. The cell suspension was transferred to a 50 ml tubes and centrifuged at 700xg for 7 minutes at room temperature. The supernatant was discarded, and the cell pellet was resuspended in the respective cell culture medium. The cells (10^6 in 1ml) were then separated into different sterile cell culture flasks and cultured under the appropriate conditions. For the infection experiments, cells were counted in a Neubauer chamber and seeded into 24-well cell culture plates.

7.4.2   Cryopreservation of the cells lines

The cell lines were carefully maintained and stored at liquid nitrogen (N$_2$) to ensure a constant supply of the cell lines. After reaching 80-90% cell confluency, the medium was discarded and confluent cells were washed with 10 ml PBS. Afterwards, cells were trypsinized and harvested as described above. The cell pellet was resuspended in a fresh freeze medium (commonly 90% FBS, 10% dimethyl sulphoxide (DMSO) such that the cell count was between 5×10^5 to 10^6 cells/ml. Aliquots of 1ml were transferred into cryoprotective vials labeled with the cell line name, passage number, lot number, cell concentration, date and stocked in liquid N$_2$ for further usage.

7.4.3   Cells counting for *in-vitro* infection assay

The number of cells were determined using a Neubauer chamber to calculate the multiplicity of infection (MOI) for *in vitro* cell culture infections. Confluent monolayer cells were trypsinized and resuspended in 15ml fresh media and a 20µl sample of cell suspension was mixed with 20µl of Trypan blue staining to determine cell count. Afterwards, 10µl was used for counting the total number of cells using the chamber under the microscope. The central largest four squares were used for counting the cells. The total cell number was then calculated as the following formula:

$$Total\ cell\ count\ in\ 15\ ml\ media = \frac{Cells\ counted\ in\ 4\ squares}{4} \times 2 \times 10{,}000 \times 15$$

7.4.4 Cell-culture infection experiments with pneumococci

7.4.4.1 Cell adhesion assays

Pneumococcal adherence to human nasopharyngeal epithelial Detroit-562 cells (ATCC CCL-138) was conducted as described [140]. Epithelial cells were seeded (2×10^5 cells per well) in 24-well tissue culture plates in RPMI-1640 (HyClone™, Germany) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS) (Gibco, Germany), 2mM glutamine, 1mM sodium pyruvate, and 1% HEPES (Sigma, Germany) and incubated for 24 h at 37°C and 5% CO_2. The confluent monolayer (80-90% confluency) was washed three times with an infection medium (cell culture medium containing 1% heat-inactivated FBS) and infected with 19F EF3030 or TIGR4Δ*cps* wild-type or mutant pneumococci using a multiplicity of infection (MOI) of 50 pneumococci per epithelial cell (**Figure 7.1**). Prior to infection, pneumococci were grown in THY to mid-log phase (OD_{600} of 0.35-0.4) and after centrifugation, resuspended in phosphate-buffered saline (PBS, pH 7.4) and infection medium (RPMI-1640, 1% heat-inactivated FBS). The infection was carried out for the indicated time points at 37°C and 5% CO_2. After infection, non-adherent pneumococci were removed in three washing steps with RPMI-1640. Pneumococcal adherence was quantified by plating the attached and internalized (less than 0.1% of host cell-associated bacteria) pneumococci [140] on blood agar plates. The bacteria were counted using a colony counter (Bern University of Applied Sciences).

7.4.4.2 Phagocytosis study by antibiotic protection assay

Phagocytosis assays have been performed using the murine macrophage cell line J774A.1 as described [118]. Cells were seeded in 24-well cell culture plates (Greiner Bio-One, Germany) with 6×10^4 cells per well in RPMI-1640 (HyClone™, Germany) supplemented with 10% FBS incubated for 48h at 37°C, and 5% CO_2. Macrophages were infected with the invasive pneumococcal strain TIGR4Δ*cps* or isogenic *serine protease* mutants with an MOI of 50 pneumococci per cell for 30 minutes at 37 °C and 5% CO_2. The antibiotic protection assay evaluated the number of internalized and recovered pneumococci. Therefore, phagocytes were washed with an infection medium (RPMI-1640/1% FBS) to remove unbound pneumococci and incubated in RPMI-1640 medium containing gentamicin (200µg/ml) and penicillin G (100 units/ml) to kill non-internalized pneumococci. Afterwards, intracellular pneumococci were recovered by saponin-mediated lysis (1%) of macrophages. The number of recovered and survival pneumococci was determined by plating of the released intracellular pneumococci on blood agar plates.

7.4.4.3 Double immunofluorescence staining (DIF staining)

Double immunofluorescence was used to visualize the extra and intracellular pneumococci in the phagocytosis experiment. While, in the adherence assay, immunofluorescence microscopy was performed to visualize pneumococcal adherence to host epithelial cells. Staining of host cell-attached pneumococci was carried out as described[82, 118]. Cells were seeded on glass coverslips (diameter 12 mm, 24-well plates) and infected with an MOI of 50 pneumococci per cell without centrifugation. After the incubation time, infected cells were washed with PBS to eliminate unbound bacteria and fixed with 4% paraformaldehyde in PBS overnight at 4°C. Infected host cells were washed three times with PBS and blocked with 10% FBS/PBS for 3h at room temperature. Adherent pneumococci were stained using a polyclonal rabbit anti-pneumococci IgG (1:1000) followed by Alexa-Fluor® 488-labeled secondary goat anti-rabbit IgG (1:1000, green) (Abcam, Germany). Glass coverslips were washed with PBS and treated with 0.1% Triton X-100 in PBS (10 min, room temperature) to permeabilize the cells. Internalized pneumococci were stained using a polyclonal rabbit anti-pneumococci IgG (1:1000) and Alexa-Fluor 568-labeled secondary goat anti-rabbit IgG (1:1000, red). The number of intracellular pneumococci was analyzed by counting at least 50 phagocytes were counted per coverslips. The actin cytoskeleton was stained with Phalloidin-iFluor®-594 conjugate (red) (Abcam, Germany). Image acquisition was performed with a fluorescence microscope (Zeiss Axio-Observer. Z) using imaging software (Zen 2.6, Zeiss, Germany).
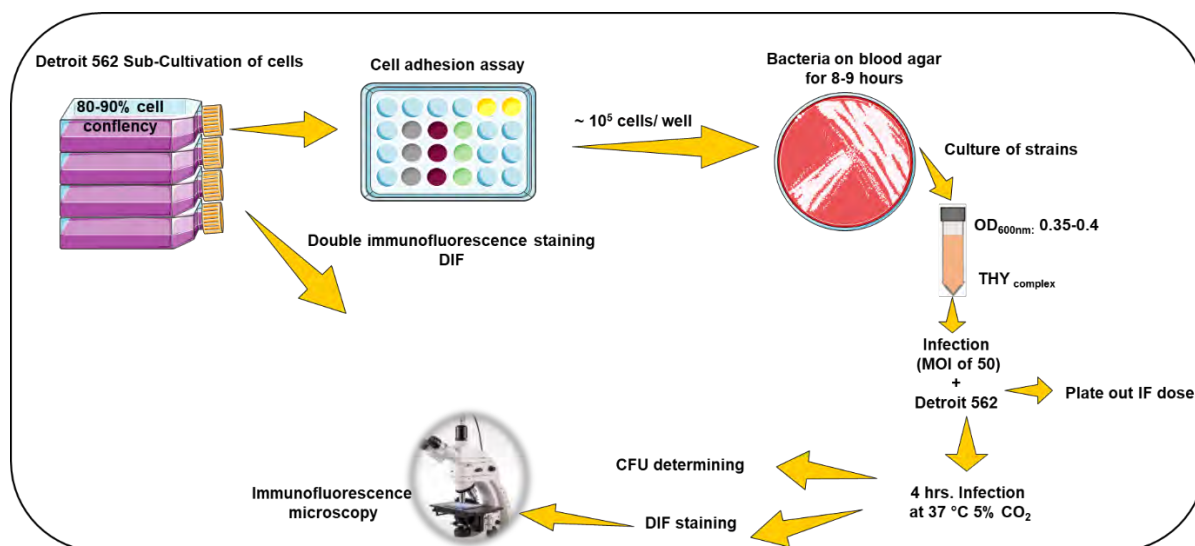
Figure 7.1: Schematic cartoon summarized the general workflow of the adherence assay.
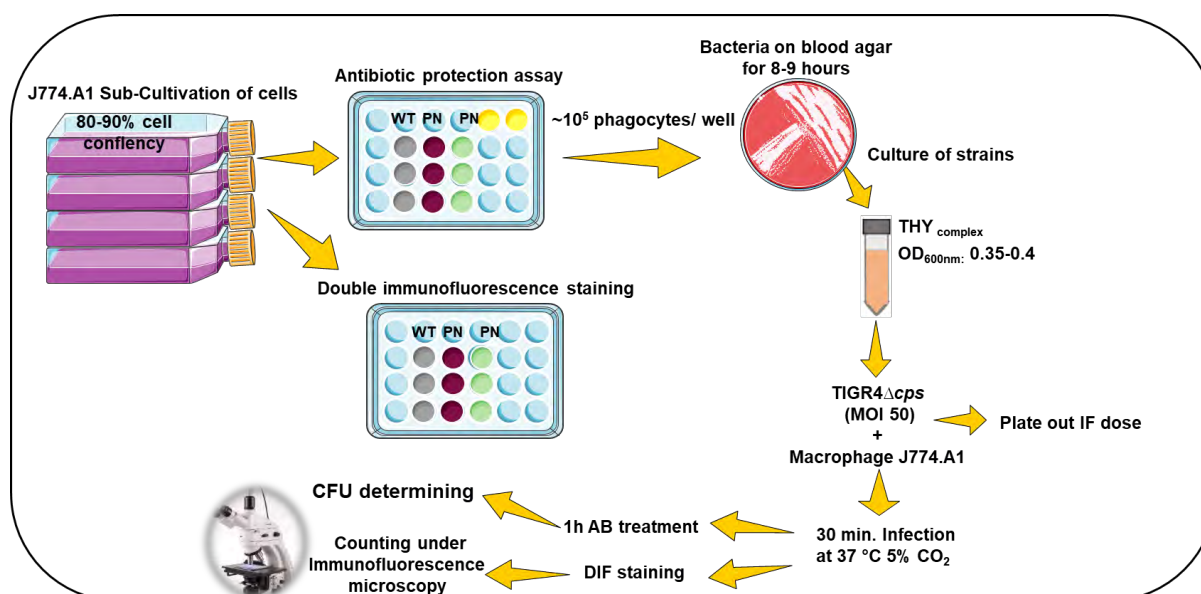
Figure 7.2: Schematic representation of the general workflow of the phagocytosis assay.

7.4.5 Biofilm formation models *in vitro*

7.4.5.1 General overview

To investigate the impact of extracellular serine proteases contributing to pneumococcal biofilm formation. We have established two models for biofilm formation. Substratum biofilm model, the pneumococcal biofilm were grown on live epithelial cells using cell line Detroit-562 as described by[198]. In addition, in static biofilm model, biofilm is grown on sterile round coverslips as described by Chao *et al*., 2014[355]. In the biofilm studies, the serotype 19F strains EF3030 and isogenic mutants were used (listed in **Table 6.11**). The living epithelial substratum appears to be essential for pneumococcal biofilm formation, as cells from the respiratory tract provide a more natural environment for the development of biofilms. The releases of pneumococcal hydrogen peroxide and pneumolysin in the medium during pneumococcal biofilm growth can cause cell toxicity toward the epithelial cells. Therefore, the biofilm was grown on pre-fixed epithelial cells to downregulate pneumococcal virulence factors and then transplanted to live epithelial cells [355, 428]. Each subheading provides a brief description of the protocol followed at each step of the study to illustrate the methods presented in this section (*see* **Figure 7.3**).

7.4.5.2 Preparation of frozen bacterial stocks

The pneumococcal biofilm inocula were prepared in advance by cultivating the pneumococcal strains in THY until the mid-log growth and were stored as glycerol-supplemented aliquots at -80°C. In brief, pneumococcal strains were grown on blood agar plates as described previously (*see* subheading **7.1.2**). The following day, pneumococci were cultured in 10 ml THY medium with an initial OD_{600nm} 0.08-0.1. Bacterial culture was incubated at 37°C in a water bath to mid-logarithmic phase (OD_{600nm} 0.6, approximately 10^8 CFU/ml). Afterwards, 2 ml of 80% glycerol was added into a 10 ml bacterial culture mixed with pipetting. The bacterial glycerol solution was aliquot in 1 ml cryotubes and stored at a -80°C freezer.

7.4.5.3 Preparation of the epithelial substratum cells

Detroit-562 epithelial cells were seeded into 24-well plates (1 ml, 2.5×10^5 cells per well) in RPMI-1640 (HyClone™, Germany) supplemented with 10% (v/v) heat-inactivated FBS, 2mM glutamine, 1mM sodium pyruvate, 1%HEPES (Sigma, Germany) and incubated for 24 h at 37°C and 5% CO_2 until a confluent monolayer was achieved. Further, cells were washed with PBS and fixed with 500µl of 4% buffered paraformaldehyde in PBS for 1h at room temperature. The pre-fixed epithelial cells were used immediately or stored at 4°C for a maximum of one month.

7.4.5.4 Biofilm formation on pre-fixed epithelial cells

The biofilm formation on pre-fixed epithelial cells for 48h at 34°C was performed to transfer the grown bacterial biofilms onto living epithelial cells. The temperature of 34°C simulates thereby the natural nasopharyngeal temperature[429]. Appropriate amounts of aliquots were thawed on the day of bacterial seeding, followed by adjusting the dose to 1×10^7 CFU/ml. Briefly, frozen stocks of pneumococci wild-type and their isogenic mutants were thawed from glycerol-stocks and sub-cultivated in RPMI$_{modi}$ medium (For the full recipe, *see* **Table 6.8** and **Table 6.9**) and grown at 37°C up to OD$_{600nm}$ = 0.2. The pneumococcal culture was diluted 1:2 in fresh medium to approximately 10^7 CFU/ml, verified by plating on blood agar plates. The fixed epithelial cells in a 24-well plate were washed three times with PBS (10 min between washes) to eliminate the rest of PFA. Afterwards, the bacterial suspension (0.5 ml) was seeded onto PFA-fixed epithelial cells in 24-well plates and grown for 48h at 34°C, 5% and CO_2 with fresh medium replaced every 12h (*see* **Figure 7.3**).

7.4.5.5 Biofilm formation on living epithelial cells

Mature biofilm grown for 48h on prefixed epithelial cells were suspended and diluted 1:3 in RPMI-1640 (HyClone™, Germany) supplemented with 2% (v/v) heat-inactivated FBS and 5mM sodium pyruvate (Sigma, Germany). The biofilm suspension (0.5 ml) was transferred to sterile, round glass coverslips placed in sterile polystyrene 24-well plates with or without a substratum of living epithelial cells. Biofilms were growing at 34°C in 5% CO_2 for the indicated times points 24, 48h and 72h with a change of pre-warmed medium every 4-6h for the living epithelial cells (biological surface) and every 12h for the biofilm on round glass coverslips (abiotic surface). Changing the medium frequently is very important for the viability of the epithelial cells and using a pre-warmed medium to keep the temperature constant. These biofilms can be used to assess the thickness, biomass, antibiotic resistance, dispersed bacteria, and 3-dimensional biofilm structures using confocal laser scanning microscopy (CLSM) or scanning electron microscopy (SEM).

7.4.5.6 Biofilm phenotype assessment

Mature biofilm were analyzed by plating the viable bacteria (biofilm biomass), antibiotic resistance (gentamicin sensitivity) and biofilm dispersal (released bacteria) after being incubated for 48 or 72h, respectively. Biofilms were exposed to 500µg/ml gentamicin (gentamicin sensitivity) or in the absence of antibiotics (total biofilm) with PBS alone for 3h and incubated at 34°C. To disrupt biofilm structures and dissociate bacteria, biofilms were scraped and resuspended by pipetting up and down. The bacterial suspension was serially

diluted in PBS and plated on blood agar plates. After incubation for 24-48h at 37°C and 5% CO_2, colonies were counted to determine the CFU/ml after antibiotic treatment. The dispersed bacteria released during the biofilm formation were assessed by collecting the supernatant from the biofilms after indicated times points and plating on blood agar plates.
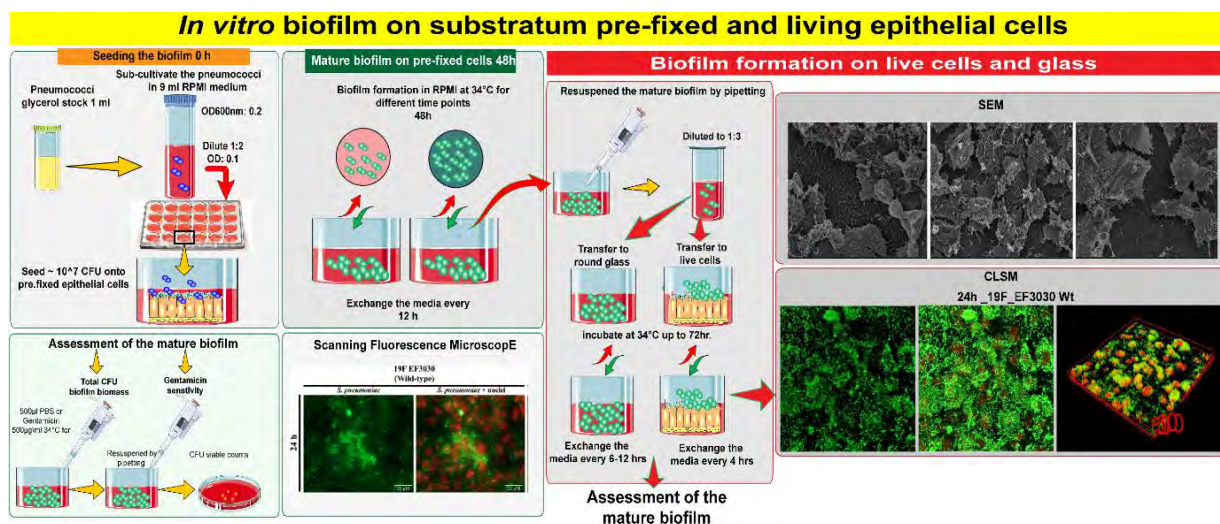
Figure 7.3: The methodology presents the biofilm formation models *in vitro*.

Pneumococcal strains were sub-cultivated in 9 ml RPMI$_{modi}$ medium, grown for OD$_{600}$ 0.2 and diluted 1:2 into fresh medium. Approximately 1×10^7 pneumococci were seeded on pre-fixed cells and incubated for 48h. The media was changed every 12h, as shown in the upper left corner. The mature biofilm was then resuspended and diluted 1:3 and transferred onto living epithelial cells. The medium was carefully replaced every 4-6h, as shown in the right section of the figure. The biofilm samples were further studies by either the imaging by SEM or quantification of CFU on blood agar plates.

7.5 Scanning electron microscopy (SEM) of pneumococcal biofilms

Scanning electron microscopy (SEM) with lysine-acetate-based formaldehyde-glutaraldehyde ruthenium red-osmium fixed samples (LRR fixation) were used to visualize and study the development of the biofilm over time. Thus, biofilms were prepared for SEM as described previously [77]. This approach has been shown to preserve carbohydrate structures and enhance biofilm morphology preservation [77]. Briefly, biofilms were prepared using the conditions described above. Biofilm samples grown on glass coverslips in 24-well plates were fixed with 300µl of precooled fixation solution-1 containing 2% paraformaldehyde, 2.5% glutaraldehyde and 0.075 M lysine-acetate in cacodylate buffer (0.15% ruthenium red, 0.2 M cacodylate, 0.01

M CaCl$_2$, 0.01 M MgCl$_2$, 0.09 M sucrose; pH 6.9) for 20 min on ice. Subsequently, biofilm samples were gently washed with 300µl cacodylate buffer (two times) and fixed for the second time with 300µl of precooled fixation solution-2 containing cacodylate buffer with 0.15% ruthenium red, 2% formaldehyde and 2.5% glutaraldehyde for 2h. After washing with cacodylate buffer biofilm samples were fixed with **1% osmium in ruthenium** red containing cacodylate buffer for 1 h at room temperature. After that, samples were washed several times with a **ruthenium red cacodylate buffer**. Fixed biofilms in 24-well plates were filled up with a cacodylate buffer and stored at 4°C. The electron microscopy has been carried out in cooperation with Prof. Manfred Rohde and Dr. Mathias Müsken (Department of Medical Microbiology, HZI, Braunschweig, Germany). Contrast and brightness were adjusted with Adobe Photoshop CS5.

Table 7.4: Fixing and washing solutions of LRR fixation for SEM

Buffer/solutions	Compositions
Fixation solution-1	
Cacodylat buffer & 0.15% ruthenium red	0.5 ml
Paraformaldehyde solution, 25%	80 µl
Glutaraldehyde solution, 25%	100 µl
L-lysine acetate salt (75 mM)	0.0155 g
A.dest	1 ml
Fixation solution-2 without lysine	
Cacodylat buffer & 0.15% ruthenium red	0.5 ml
Paraformaldehyde solution, 25%	80 µl
Glutaraldehyde solution, 25%	100 µl
A.dest	1ml
Washing solution	
Cacodylat buffer & 0.15% ruthenium red	0.5 ml
A.dest	1 ml
Osmium solution	
Washing solution	800 µl
Osmium tetroxide solution, 5%	200 µl

Table 7.5: Individual buffers for scanning electron microscopy of LRR-fixation

Buffer/solutions	Compositions	Amount
Cacodylat-buffer	Cocodylate (Natriumslatz Trihydrat), 0.2 M	21.41 g
(pH 6.9)	Calcium chloride dehydrate, 0.02 M	1.47 g
	Magnesium chloride hexahydrate, 0.02 M	2,04 g
	Fill up to 450 ml with A.dest, and adjust the to pH 6.9 with 1 M HCl. Fill up with 500 ml of A, dest	
Sucrose-Cacodylat	Cacodylat-buffer, pH 6.9	100 ml
	Sucrose	6.16 g
Cacodylat buffer and Ruthenium red (0.15%)	Cacodylat buffer	100 ml
	Ruthenium red	0.15 g
Paraformaldehyde solution, 25%	Paraformaldehyde	25 g
	Dissolve paraformaldydae by add 90 ml of A.dest, heat up to 60 °C, and add 10 M NaOH under stirring until it dissolves. Fill up with 100 ml of A, dest. Paraformaldehyde solution kept for a few days at room temperature, then filter the precipitation (sterile filter 0.2 μm) or remove it by centrifugation (13,000 xg, for 2 min).	
Glutaraldehyde solution, 25%	Ready-to-use solution (Roth)	

7.6 *In vivo* mouse infection model of *S. pneumoniae*

The influence of serine proteases on nasopharyngeal colonization and pneumonia was analyzed *in vivo* by applying two different mouse infection models, namely the nasopharyngeal colonization model and the acute pneumonia model.

All animal experiments were conducted in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals (National Research Council, USA), the guidelines of the ethics committee at The University of Greifswald, and the German regulations of the Society for Laboratory Animal Science (GVSOLAS) and the European Health Law of the Federation of Laboratory Animal Science Associations (FELASA). All experiments were approved by the Landesamt für Landwirtschaft, Lebensmittelsicherheit und Fischerei Mecklenburg-Vorpommern (LALLFV M-V, Rostock, Germany) and the LALLFV M-V ethical board (LALLF M-V permit no. 7221.3-1-056/16). All efforts were taken to minimize suffering, ensure the highest ethical standard and adhere to the 3R principle (reduction, refinement, and

replacement). All *in vivo* infection experiments were conducted with female CD-1 outbred mice (age of 8-10 weeks) purchased from Charles River, Sulzfeld, Germany. Before the experiments, the mice were adapted for 1 to 2 weeks to the new environment. For the experiments, 8-10-week old female CD-1 outbred mice were anesthetized intraperitoneally with ketamine and xylazine.

7.6.1 Mouse acute pneumoniae model and visualization via the IVIS® spectrum imaging system

TIGR4 is a clinical isolate and causes severe pneumonia and invasive diseases in mice (van Ginkel et al., 2003). Pneumococcal strains were grown in THY medium supplemented with 10% heat-inactivated FCS, up to OD600nm 0.35-0.4. After anesthesia, mice were intranasally infected with 20 µl PBS/1%FCS containing 9×10^7 bioluminescent *S. pneumoniae* TIGR4*lux* or isogenic triple *serine protease* mutants. Infected mice were imaged and monitored every 8h using the IVIS® Spectrum Imaging System (Caliper Life Sciences) to determine the dissemination of the bioluminescent pneumococci in the mice as described [380]. Besides, the bioluminescent intensity was quantified as the total photon emission using LivingImage® 4.1 software package (Caliper Life Sciences). The CFU of the bacterial inoculum was confirmed by plating serial dilutions on blood agar plates.

Figure 7.4: The summarized workflow for the murine acute pneumoniae model. Pneumococcal strains were cultured in THY medium to mid logarithmic phase followed by adjustment of the infection dose to 9×10^7 CFU in 10µl PBS per mouse for the intranasal infection (IN). Infected mice were imaged using an IVIS® Spectrum Imaging System.

7.6.2 Murine model of pneumococcal colonization

The murine colonization model was used to study the role of pneumococcal serine proteases in pneumococcal colonization. For that, strain 19F_EF3030 was used, which colonizes the nasopharynx of mice while being mostly noninvasive [342, 430]. The pneumococcal inoculum was prepared by cultivating the pneumococcal strain in THY supplemented with 10% heat-inactivated FCS until mid-log growth phase 0.35-0.4. Afterwards, mice were intranasally infected with 20µl PBS/1%FCS per mouse, containing 1×10^7 bacteria of 19F_EF3030 (wild-

type) or isogenic *serine protease* mutants. The infection dose was verified by plating on blood agar plates. An intraperitoneal injection of ketamine/xylazine (50mg ketamine and 5mg xylazine/kg) was used to anesthetize mice for the intranasal application. Nasopharyngeal washes (NP) and bronchoalveolar lavages (BAL) were collected at time points 2, 3, 7 and 14 days post-infection and serially diluted to determine the bacterial recovery (log CFU/ml) by plating on blood agar plates (Oxoid Microbiology Products).

7.7 Graphical presentation and Statistics analysis

Statistical significance between different groups was calculated using a one-way ANOVA (Kruskal-Wallis test) followed by Dunnett's post-test for the mouse colonization and bioluminescence measurements in the acute pneumonia model. Unpaired two-tailed Student's t-test (Mann-Whitney test) was performed to analyze the difference between the two groups. Two-way ANOVA analysis was used with the optical density for growth behavior. Kaplan-Meier survival curves of mice were compared by the log-rank (Mantel-Cox) test. A p-value of <0.05 was considered statically significant. All statistical analyses were performed using GraphPad Prism version 5.0 (GraphPad, Software, La Jolla, CA, USA).

7.8 Bioinformatics analysis of the genomes of *S. pneumoniae*

Gene and amino acid sequences of pneumococcal serine protease were extracted from the National Center for Biotechnology Information NCBI online tool database (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi), for the homolog analysis search by Clustal Omega. Furthermore, different database tools were used to analyze and predict the pneumococcal serine protease localization, including PSORT db 3.0[431] and the Surface Localization Extracellular Proteins pipeline SLEP (http://bl210.caspur.it/slep/). All analyzed serine protease gene sequences (*prtA* (*sp_0641*), *htrA* (*sp_2239*), SFP (*sp_1954*) and *cbpG* (*sp_0390*) of *S. pneumoniae* strain TIGR4 were retrieved from the KEGG (https://www.genome.jp/kegg/)[332]. Signal sequences were predicted using the software tool SignalP 4.0 (http://www.cbs.dtu.dk/services/SignalP-4.0/)[432, 433]. Choline binding proteins were identified via typical choline-binding domains consisting of characteristic repeat motifs[99]. Moreover, for transmembrane helices' prediction, the TMHMM Server 2.0 algorithm (Hidden Markov Model for transmembrane protein topology prediction) was applied (http://www.cbs.dtu.dk/services/TMHMM/)[434]. Functional domains were predicted using Pfam . Blast analysis and genome annotation of pneumococcal serine proteases revealed only four different serine proteases[92]. The genomes of various clinically relevant *S. pneumoniae* strains

were analyzed on DNA and protein levels with BlastN and BlastP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=&LINK_LOC=blasttab&LAST_PAGE=blastp) respectively, for the homology analysis of pneumococcal serine proteases. Finally, we compared the catalytic center of all four serine proteases from pneumococci by using computer-assisted analysis. The calculations were performed within the Multiple Sequence Viewer/Editor application in Maestro 2020-4 (Schrödinger, LLC, New York, NY, 2020) using an energy-based approach. The templates were obtained by BLAST search in the PDB database **8.4.2** (SFP: 4MZD; PrtA: 5FAX; HtrA: 5ZVJ; CbpG: 1P3C).

Chapter **8**

APPENDIX

8. Appendix

8.1 List of abbreviations

%	Percentage	IPD	Invasive pneumococcal disease
(v/v)	Volume percent, volume per volume	IPTG	Isopropyl-β-D-thiogalactopyranosid
(w/v)	Weight, weight per volume	kb	Kilo base pair
®	Registered		
6-PGD	6-phosphogluconate dehydrogenases	kDa	Kilo-Dalton
α	Alpha	kDa	Kilo-Dalton
β	Beta	Km^R	Resistance to kanamycin
°C	Degree Celsius	l	Liter
A	Absorbance	LB	Luria Bertani broth
aa	Amino acid	Lgt	Lipoprotein diacylglycerol transferase
A. dest	Distilled water	LPxTG proteins	Sortase-anchored proteins
AJ	Adhesion junction	Lsp	Lipoprotein signal peptidase II
AOM	Acute otitis media	LTA	Lipoteichoic acid
ATP	Adenosine triphosphate	LytA	Autolysin A
BSA	Bovine serum albumin	m	Milli (10^{-3})
CAM	Cell adhesion molecule superfamily	M	Molar (mol/l)
CLSM	confocal laser scanning microscopy	MMPs	Matrix metalloproteinases
CDM	chemically defined medium	MW	Molecular weight
			Milligram
CAP	Community-acquired pneumonia	mg	
			Minute
CBB	Coomassie Brilliant Blue	min	
CBD	Choline-binding domain	ml	Milliliter
CBM	Choline-binding module	Mm	Millimolar
CBP	Choline-binding protein	MSCRAMMS	Microbial surface components recognizing adhesive matrix molecules
CbpG	Choline-binding-protein G	MurNAc	N-acetylmuramic acid
CbpA	Pneumococcal surface protein C (PspC)	NanA	Neuraminidase A
CFU	Colony forming units	NCBI	National Center for Biotechnology Information
CiaRH	Response regulator/histidine kinase protein	NCSP	Non-classical surface protein
ClpP	Caseinolytic Protease P	NESp	Nonencapsulated *Streptococcus pneumoniae*
CPS	Capsular polysaccharide	NGS	Next-generation sequence

CSP	Competence stimulating peptide	ns	Nor significant
C-terminal	Carboxy-terminus of proteins	N-terminal	Amino-terminus of proteins
Da	Dalton	OD	Optical density
DIF	Double immunofluorescence	p	p-value
DMSO	Dimethyl sulfoxide	PAGE	Polyacrylamide gel electrophoresis
DNA	Deoxyribonucleic acid	PavA	Pneumococcal adhesion and virulence factor A
DNase	Deoxyribonuclease	PBS	Phosphate-buffered saline
DegP	Degradation of periplasmic proteins	PCho	Phosphorylcholine
dNTP/s	Deoxyribonucleic triphosphate	PCR	Polymerase chain reaction
E-cadherin	Epithelial cadherin	PCV	Pneumococcal conjugate vaccines
E. coli	*Escherichia coli*	PepN	Aminopeptidase N
ECDC	European Centre for Disease Prevention and Control	PFA	Paraformaldehyde
ECM	Extracellular matrix	Pfu	Pyrococcus furiosus
Eno	Enolase	Plg	Plasminogen
EPS	Extracellular polymeric substance	pIgR	Polymeric immunoglobulin receptoR
Erm^R	Resistance to erythromycin	Ply	Pneumolysin
et al.	And others	PrtA	Cell wall-associated serine proteinase A
FCS	Fetal Calf Serum	PspA	Pneumococcal surface protein A
FM	Functional module	PspC	Pneumococcal surface protein C (=CbpA)
Fn	Fibronectin	PhtA	pneumococcal histidine triad protein A
GAPDH	Glycerate-3-phosphate dehydrogenase	Pce	Phosphorylcholine esterase
GlcNAc	N-acetylglucosamine	TCS	pneumococcal two-component
h	Hours	R	Resistance
H2O2	Hydrogen peroxide	Ref.	Reference
His	Histidine	rpm	Revolutions per minute
His$_6$-Tag	Histidine tagged	RT	Room temperature
HtrA	High-temperature requirement A	S1	First strand
i.n.	Intranasal	sec	Second
i.p.	Intraperitoneal	SEM	Scanning electron microscopy
i.v.	Intravenous	*S. p.*	*Streptococcus pneumoniae*
IgA	Immunglobulin A	SprE	secreted serine protease
IgG	Immunglobulin G		

8.2 Chemicals, laboratory equipment, consumables and software

Supplementary Table 8.1: Laboratory equipment

Equipment	Provider
Agarose gel electrophoresis chambers	BioRad
ÄKTApurifier™-900	GE Healthcare
Autoclave VX-100 and VX-150	Systec
Balance Sartorius Basic	Sartorius
Bunsen burner	WLD-Tec
Centrifuge 5417R	Eppendorf
Centrifuge Pico17	ThermoFisher Scientific
CO_2 incubators Heracell 150 Heraeus	ThermoFisher Scientific
Electric pipetting aids accu-jet® pro	Brand
Deep freezer, -80°C	ThermoFisher Scientific
Heating block, ThermoMixer compact	Eppendorf
Ice machine	Scotsman
Incubator	ThermoFisher Scientific
IVIS® Lumina Imaging System	Xenogen Corporation, Caliper Life Sciences
INTAS GDS®, gel documentation system	INTAS
Magnetic plate separator	Greiner
Magnetic stirrer, RCT basic	IKA
Microscopes:	Microscopes:
Confocal Laser Scanning Microscope Zeiss LSM510 META	Zeiss
Microwave	LG
Multichannel pipette manual/ electronic	Eppendorf/ RAININ
Neubauer counting chamber	Brand
PCR-Thermocycler T3 Thermocycler	Biometra
pH-Meter	WTW
Photo developing machine	Fuji
Pipettes, manual/ electronic	Gilson, Eppendorf/ RAININ
Power Supply Power Pac 200 und 300	BioRad
Refrigerator, 4 °C	Liebherr
Rolling incubator	Schmidt GmbH
Scanner jet 8200	HP
SDS-Gel electrophoresis chamber	BioRad
Sensitive balance	OHAUS
Shaking incubator, VKS-75 control	Edmund Bühler
Spectrophotometer, NanoDrop ND-1000	peQLab
Sterile work benches for cell culture, MSC-Advantage HeraSafe	ThermoFisher Scientific
Sterile workbenches for bacterial culture	Heraeus Instruments
Timer	Roth
Transblot SD Semidry Transfer Cell	BioRad
Ultrasonic bath, Ultrasonic cleaner	Unisonic
Ultrasound device, Sonifier 250	Branson
Vacuum pump	KNF
Vortex	IKA
Water bath	GFL

Supplementary Table 8.2: Consumable items

Consumable	Provider
24-Cell culture plates, flat bottom, Cellstar bio-one	Greiner
Coverslips, 12mm	Hartenstein
Centrifuge tubes 15ml, 50ml	Sarstedt
Cryo tubes	Sarstedt
Disposable gloves	Hartmann
Dialysis bag 12-14,000 MWCO	neoLab Laborbedarf
Disposable serological pipettes, steril, 5 ml, 10 ml, 25 ml	Sarstedt
Eppendorf tubes 0.5ml, 1ml, 2ml	Sarstedt
FLOQSwab™	Copan diagnostics
Forceps, 130 mm	Hartenstein
Nitro cellulose membranes Protran® Whatman	Hartenstein
Parafilm	Hartenstein
Pasteur pipettes	Hartenstein
PCR reaction tubes, 0.2 ml	Sarstedt
Petri dishes	Greiner
Pipette tips, 10 µl, 200 µl, 1000 µl	Sarstedt
Plastic cuvettes	Sarstedt
Cotton tipped applicator	Henry Schein Inc
Screw cap tubes 2ml	Sarstedt
Scalpel	Braun
Scissors, curved 105 mm	Hartenstein
Scissors, straight 105 mm	Hartenstein
Slides	Hartenstein
Syringe filter, 0.22 µm, 0.45 µm	Roth

Supplementary Table 8.3: Software

Software	Provider
Adobe Acrobat Pro 2017	Adobe Systems
Adobe Illustrator CS5	Adobe Systems
Adobe Photoshop Elements 8.0	Adobe Systems
BioRender	BioRender Online Software
EndNote X7	Thomson Reuters
GraphPad Prism 5	GraphPad
Living Image IgorPro 4.0 software package	Xenogen Corporation
MS Office Professional Plus 2016	Microsoft
Maestro 2020-4	Schrödinger, New York, 2020
Serial Cloner 2.6	Serialbasics

Supplementary Table 8.4: Databases and online tools

Database and online tool	Source
BlastP	https://blast.ncbi.nlm.nih.gov/Blast.cgi
Clustal Omega	https://www.ebi.ac.uk/Tools/msa/clustalo/
KEGG database	http://www.genome.jp/kegg/
NCBI database	http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi
TMHMM 2.0 algorithm	http://www.cbs.dtu.dk/services/TMHMM/
SignalP 4.0	http://www.cbs Basis.dtu.dk/services/
PSORT db 3.0	https://www.psort.org/psortb/
Pairwise sequence	https://www.ebi.ac.uk/Tools/psa/emboss_water/

8.3 Pneumococcal serine protease amino acid sequences

8.3.1 HtrA amino acid sequences

S. pneumoniae HtrA *sp_2239* in *TIGR4* and *EF3030_11105* in 19F proteins are identical (protein accession no. AAK76286.1 and QBF69928.1). The signal peptide sequence (31 aa) is marked in blue, the serine protease catalytic domain (182 aa) position 96-277 is marked in red, PDZ domain (87 aa) position 289-375 is marked in green.

```
MKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYKNENSTTQAVNKV
KDAVVSVITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAYIVTNNHVINGASKVDIRLSD
GTKVPGEIVGADTFSDIAVVKISSEKVTTVAEFGDSSKLTVGETAIAIGSPLGSEYANTVTQGIVSSL
NRNVSLKSEDGQAISTKAIQTDTAINPGNSGGPLINIQGQVIGITSSKIATNGGTSVEGLGFAIPAND
AINIIEQLEKNGKVTRPALGIQMVNLSNVSTSDIRRLNIPSNVTSGVIVRSVQSNMPANGHLEKYDVI
TKVDDKEIASSTDLQSALYNHSIGDTIKITYYRNGKEETTSIKLNKSSGDLES
```

8.3.2 PrtA amino acid sequences

S. pneumoniae PrtA *sp_0641* in TIG4 and *EF3030_03025* in 19F proteins are identical (protein accession no. AAK74791.1 and QBF68585.1). The signal peptide sequence (27 aa) is marked in blue, the serine protease catalytic domain (542 aa) position 223-764 is marked in red, DUF 1034 domain (140 aa) position 795-934 is marked in green. The C-terminal LPKTG anchoring motif (42) is marked in yellow.

```
MKKSTVLSLTTAAVILAAYAPNEVVLADTSSSEDALNISDKEKVAENKEKHENIHSAMETSQDFKEKK
TAVIKEKEVVSKNPVIDNNTSNEEAKIKEENSNKSQGDYTDSFVNKNTENPKKEDKVVYIAEFKDKES
GEKAIKELSSLKNTKVLYTYDRIFNGSAIETTPDNLDKIKQIEGISSVERAQKVQPMMNHARKEIGVE
EAIDYLKSINAPFGKNFDGRGMVISNIDTGTDYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKI
PHAFNYYNGGKITVEKYDDGRDYFDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDA
GSGFAGDETMFHAIEDSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASS
SSWDLVANNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKSKITTN
EDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIMVVNTVNYYN
RDNWTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPDKKTEVKRNNKEDFKDKLEQYYPIDMESF
NSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRIDLLLKPDVSAPGKNIKSTLNVINGK
STYGYMSGTSMATPIVAASTVLIRPKLKEMLERPVLKNLKGDDKIDLTSLTKIALQNTARPMMDATSW
KEKSQYFASPRQQGAGLINVANALRNEVVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSN
RPLTFKVSASAITTDSLTDRLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFD
LNAVINVGEAKNKNKFVESFIHFESVEEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWE
EGSRSKTLGGYDDDGKPKIPGTLNKGIGGEHGIDKFNPAGVIQNRKDKNTTSLDQNPELFAFNNEGIN
APSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNTNKEGENQRDLKVISREHF
IRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKDNQDPATKIRGQFEPIAEGQYF
YKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEKIKLITKDTYHKVKDQYKNETLFARDQKEH
PEKFDEIANEVWYAGAALVNEDGEVEKNLEVTYAGEGQGRNRKLDKDGNTIYEIKGAGDLRGKIIEVI
ALDGSSNFTKIHRIKFANQADEKGMISYYLVDPDQDSSKYQKLGEIAESKFKNLGNGKEGSLKKDTTG
VEHHHQENEESIKEKSSFTIDRNISTIRDFENKDLKKLIKKKFREVDDFTSETGKRMEEYDYKYDDKG
NIIAYDDGTDLEYETEKLDEIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNNYKSIEIKATKY
DFHSKTMTFDLYANINDIVDGLAFAGDMRLFVKDNDQKKAEIKIRMPEKIKETKSEYPYVSSYGNVIE
LGEGDLSKNKPDNLTKMESGKIYSDSEKQQYLLKDNIILRKGYALKVTTYNPGKTDMLEGNGVYSKED
IAKIQKANPNLRALSETTIYADSRNVEDGRSTQSVLMSALDGFNIIRYQVFTFKMNDKGEAIDKDGNL
```

```
VTDSSKLVLFGKDDKEYTGEDKFNVEAIKEDGSMLFIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYL
RGKISDKGGFNWELRVNESVVDNYLIYGDLHIDNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTD
MDGNVYLQTGYSDLNAKAVGVHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNNGFDGEI
QEQHIYINGKEYTSFNDIKQIIDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNG
KEMSSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEEKKEEE
NKPTFDVSKKKDNPQVNHSQLNESHRKEDLQREEHSQKSDSTKDVTATVLDKNNISSK**STTNNPNKLP**
**KTG**TASGAQTLLAAGIMFIVGIFLGLKKKNQD
```

8.3.3 SFP amino acid sequences

SFP amino acid sequences of *S. pneumoniae sp_1954* in TIGR4 (protein accession number ABC75782.1). The signal peptide sequence (22 aa) is marked in blue, the serine protease catalytic domain (295 aa) position 167-461 is marked in red.

```
MKKKYWTLAILFFCLFNNSVTAQEIPKNLDGNITHTQTSESFSESDEKQVDYSNKNQEEVDQNKFRIQ
IDKTELFVTTDKHLEKNCCKLELEPQINNDIVNSESNNLLGEDNLDNKIKENVSHLDNRGGNIEHDKD
NLESSIVRKYEWDIDKVTGGGESYKLYSKSNSKVSIAILDSGVDLQNTGLLKNLSNHSKNYVPNKGYL
GKEEGEEGIISDIQDRLGHGTAVVAQIVGDDNINGVNPHVNINVYRIFGKSSASPDWIVKAIFDAVDD
GNDIINLSTGQYLMIDGEYEDGTNDFETFLKYKKAIDYANQKGVIIVAALGNDSLNVSNQSDLLKLIS
SRKKVRKPGLVVDVPSYFSSTISVGGIDRLGNLSDFSNKGDSDAIYAPAGSTLSLSELGLNNFINAEK
YKEDWIFSATLGGYTYLYGNSFAAPKVSGAIAMIIDKYKLKDQPYNYMFVKKFWKKHYQ
```

8.3.4 CbpG amino acid sequences

CbpG amino acid sequences of *S. pneumoniae sp_0390 in* TIGR4 and *EF3030_01920* in 19F protein are highly homologous in TIGR4 and 19F (protein accession no. AAK74556.1 and QBF69943.1). The trypsin-like serine protease catalytic domain (184 aa) position 14-197 is marked in red, the repeats of the CBM (CW1 position 207-221), (CW2 position 226-245), and (CW3 position 246-265) are marked in orange. At the C-terminal region (in light blue, position 267-285), it is probably involved in binding to choline resides of teichoic acids.

```
MVLSKYYGVADGMNVEGRGSANFIKDNVLITAAHNYYRHDYGKEADDIYVLPAVSPSQEPFGKIKVKE
VRYLKEFRNLNSKDAREYDLALLILEEPIGAKLGTLGLPTSQKNLTGITVTITGYPSYNFKIHQMYTD
KKQVLSDDGMFLDYQVDTLEGSSGSTVYDASHRVVGVHTLGDGANQINSAVKLNERNLPFIYSVLKGY
SLEGWKKINGSWYHYRQHDKQTGWQEINDTWYYLDSSGKMLTDWQKVNGKWYYLNSNGAMVTGSQTID
GKVYNFASSGEWI
```

8.4 Amino acid sequence comparisons

8.4.1 Serine proteases amino acid sequence comparisons in *S. pneumoniae*

Multiple sequence alignment comparisons of serine proteases were performed using the bioinformatic tool "Clustal Omega" available on the following webpage: https://www.ebi.ac.uk/Tools/msa/clustalo/. The signal peptide was detected by using the software tool SignalP 4.0 (http://www.cbs.dtu.dk/services/SignalP-4.0/) marked in red color.

8.4.1.1 Amino acid sequence comparisons of CbpG in *S. pneumoniae*

```
spn_ST556_snd:MYY_0470_(19F)        MKKTVYKKLGISIIASTLLASQLSTVSALSVISSTGEEYEVSETLQDSPGTNNFSLPNVS    60
sp_Hungary19A-6_SPH_0499_(19A)      ------------------------------------------------------------    0
spn_D39_spd_0356_(2)                ------------------------------------------------------------    0
spn_R6_spr0349_(2)                  ------------------------------------------------------------    0
spn_ST81_spn23F03640_(23F)          ------------------------------------------------------------    0
spn_JJA_spj_0378_(14)               ------------------------------------------------------------    0
spn_TIGR4_sp_0390_(4)               ------------------------------------------------------------    0
spn_G54_spg_0356_(19F)              ------------------------------------------------------------    0
spn_EF3030_01920_(19F)              ------------------------------------------------------------    0


spn_ST556_snd:MYY_0470_(19F)        PTYGQYYTNQSEVIIGKNDLVKVKNTLQYPYSTSAYVESEFRGVKNGKDVTYRGSANFIK  120
sp_Hungary19A-6_SPH_0499_(19A)      ---------------------------------------------MNVEGRGSVNFIK   13
spn_D39_spd_0356_(2)                ------------------------------MVLSKYYGVADGMNVEGRGSANFIK   25
spn_R6_spr0349_(2)                  ------------------MIKIDNTLQYPYSTSAMVLSKYYGVADGMNVEGRGSANFIK   41
spn_ST81_spn23F03640_(23F)          ---------------------------------------------MNVEGRGSANFIK   13
spn_JJA_spj_0378_(14)               ---------------------------------------------MNVEGRGSANFIK   13
spn_TIGR4_sp_0390_(4)               ------------------------------MVLSKYYGVADGMNVEGRGSANFIK   25
spn_G54_spg_0356_(19F)              ---------------------------------------------MNVEGRGSANFIK   13
spn_EF3030_01920_(19F)              ---------------------------------------------MNVEGRGSANFIK   13
                                                                                 :*   ***.****


spn_ST556_snd:MYY_0470_(19F)        DNVLITAAHNIYKHEFGKDAEELYVTPARTPNSVPFGRIKVKEVRYLKEFRNLNSKDARE  180
sp_Hungary19A-6_SPH_0499_(19A)      DNVLITAAHNYYRHDYGKEADDIYVLPAVSPSQELFGKIKVKEVCYLKEFRNLNSKDARE   73
spn_D39_spd_0356_(2)                DNVLITAAHNYYRHDYGKEADDIYVLPAVSPSQELFGKIKVKEVRYLKEFRNLNSKDARE   85
spn_R6_spr0349_(2)                  DNVLITAAHNYYRHDYGKEADDIYVLPAVSPSQELFGKIKVKEVRYLKEFRNLNSKDARE  101
spn_ST81_spn23F03640_(23F)          DNVLITAAHNYYRHDYGKEADDIYVLPAVSPSQELFGKIKVKEVRYLKEFRNLNSKNARE   73
spn_JJA_spj_0378_(14)               DNVLITAAHSYYRHDYGKEADDIYVLPAVSPSQEPFGKIKVKEVRYLKEFRNLNSKDARE   73
spn_TIGR4_sp_0390_(4)               DNVLITAAHNYYRHDYGKEADDIYVLPAVSPSQEPFGKIKVKEVRYLKEFRNLNSKDARE   85
spn_G54_spg_0356_(19F)              DNVLITAAHNYYRHDYGKEADDIYVLPAVSPSQEPFGKIKVKEVRYLKEFRNLNSKDARE   73
spn_EF3030_01920_(19F)              DNVLITAAHNYYRHDYGKEADDIYVLPAVSPSQELFGKIKVKEVRYLKEFRNLNSKDARE   73
                                    ********.  *:*::**:*::;** ** :*..  **:****** ***********:***


spn_ST556_snd:MYY_0470_(19F)        YDLALLILEEPIGAKLGTLGLPTSQKNLTGITVTITGYPSYNFKIHQMYTDKKQVLSDDG  240
sp_Hungary19A-6_SPH_0499_(19A)      YDLALLILEEPIGAKLGTLGLPTSQKNLTGITVTITGYPSYNFKIHQMYTDKKQVLSDDG  133
spn_D39_spd_0356_(2)                YDLALLILEKPIGAKLGTLGLPTSQKNLTGITVTITGYPSYNFKIHQMYTDKKQVLSDDG  145
spn_R6_spr0349_(2)                  YDLALLILEKPIGAKLGTLGLPTSQKNLTGITVTITGYPSYNFKIHQMYTDKKQVLSDDG  161
spn_ST81_spn23F03640_(23F)          YDLALLILEEPIGAKLGTLGLPTSQKNLTGITVTITGYLSYNFKIHQMYTDKKQVLSDDG  133
spn_JJA_spj_0378_(14)               YDLALLILEEPIGAKLGTLGLPTSQKNLTGITVTITGYPSYNFKVHQMYTDKKQVLSDDG  133
spn_TIGR4_sp_0390_(4)               YDLALLILEEPIGAKLGTLGLPTSQKNLTGITVTITGYPSYNFKIHQMYTDKKQVLSDDG  145
spn_G54_spg_0356_(19F)              YDLALLILEEPIGAKLGTLGLPTSQKNLTGITVTITGYPSYNFKIHQMYTDKKQVLSDDG  133
spn_EF3030_01920_(19F)              YDLALLILEEPIGAKLGTLGLPTSQKNLTGITVTITGYPSYNFKIHQMYTDKKQVLSDDG  133
                                    *********:*********************** *****:****************


spn_ST556_snd:MYY_0470_(19F)        MFLDYQVDTLEGASGSTVYDASNRVVGVHTLGDGANQINSAVKLNERNLPFIYSVLKGYS  300
sp_Hungary19A-6_SPH_0499_(19A)      MFLDYQVDTLEGSSGSTVYDAS-RVVGVHTLGDGANQINSAVKLNERNLPFIYLFIRFLK  192
spn_D39_spd_0356_(2)                MFLDYQVDTLEGSSGSTVYDASHRVVGVHTLGDGANQINSAVKLNERNLPFIYLFGS---  202
spn_R6_spr0349_(2)                  MFLDYQVDTLEGSSGSTVYDASHRVVGVHTLGDGANQINSAVKLNERNLPFIYLFGS---  218
spn_ST81_spn23F03640_(23F)          MFLDYQVDTLEGSSGSAVYDASHRVVGVHTLGDGANQINSAVKLNERNLPFIYSVLKGYS  193
spn_JJA_spj_0378_(14)               MFLDYQVDTLEGSSGSTVYDASHRVVGVHTLGDGANQINSAVKLNERNLPFIYSVLKGYS  193
spn_TIGR4_sp_0390_(4)               MFLDYQVDTLEGSSGSTVYDASHRVVGVHTLGDGANQINSAVKLNERNLPFIYSVLKGYS  205
spn_G54_spg_0356_(19F)              MFLDYQVDTLEGSSGSTVYDASHRVVGVHTLGDGANQINSAVKLNERNLPFIYSVLKGYS  193
spn_EF3030_01920_(19F)              MFLDYQVDTLEGSSGSTVYDASHRVVGVHTLGDGANQINSAVKLNERNLPFIYSVLKGYS  193
                                    ************:***.***** ****************************** .


spn_ST556_snd:MYY_0470_(19F)        LEGW--------------------------------------------------------  304
sp_Hungary19A-6_SPH_0499_(19A)      VTLLKDGRK---------------------------------------------------  201
spn_D39_spd_0356_(2)                ------------------------------------------------------------  202
spn_R6_spr0349_(2)                  ------------------------------------------------------------  218
spn_ST81_spn23F03640_(23F)          LEGWKKINGSWYYRQHDKQTGWQEINDTWYYLDSSGKMLTDWQKVNGKWYYLNSNRAMV  253
spn_JJA_spj_0378_(14)               LEGWKKINGSWYHYRQHDKQTGWQEINDTWYYLDSSGKMLTDWQKVNGKWYYLNSNGAMV  253
spn_TIGR4_sp_0390_(4)               LEGWKKINGSWYHYRQHDKQTGWQEINDTWYYLDSSGKMLTDWQKVNGKWYYLNSNGAMV  265
spn_G54_spg_0356_(19F)              LEGWKKINGSWYHYRQHDKQTGWQEINDTWYYLDSSGKMLTDWQKVNGKWYYLNSNGAMV  253
spn_EF3030_01920_(19F)              LEGWKKINGSWYHYRQHDKQTGWQEINDTWYYLDSSGKMLTDWQKVNGKWYYLNSNGAMV  253
                                         CBD_1                CBD_2            CBD_3


spn_ST556_snd:MYY_0470_(19F)        ---------------------  304
sp_Hungary19A-6_SPH_0499_(19A)      ---------------------  201
spn_D39_spd_0356_(2)                ---------------------  202
spn_R6_spr0349_(2)                  ---------------------  218
spn_ST81_spn23F03640_(23F)          TGSQTIDGKVYNFASSGEWI  273
spn_JJA_spj_0378_(14)               TGSQTIDGKVYNFASSGEWI  273
spn_TIGR4_sp_0390_(4)               TGSQTIDGKVYNFASSGEWI  285
spn_G54_spg_0356_(19F)              TGSQTIDGKVYNFASSGEWI  273
spn_EF3030_01920_(19F)              TGSQTIDGKVYNFASSGEWI  273
                                        CBD_?
```

*The repeats of the CBM at the C-terminal region are highlighted by different colors.

8.4.1.2 Amino acid sequence comparisons of SFP in *S. pneumoniae*

```
spn_TIGR4_sp_1954_(4)        ---MKKKYWTLAILFFCLFNNSVTAQEIPKNLDGNITHTQTSESFSESDEKQVDYSNKNQ    57
spn_D39_spd_1753_(2)         MKIMKKKYWTLAILFFCLFNNSVTAQEIPKNLDGNITHTQTSESFSESDEKQVDYSNKNQ    60
spn_spn23F19760_ST81_(23F)   MKIMKKKYWTLAILFFCLFNNSVTAQEIPKNLDGNITHTQTSESFSESDEKQVDYSNKNQ    60
spn_spj_1948_JJA_(14)        MKIMKKKYWTLAILFFCLFNNSVTAQEIPKNLDGNITHTQTSESFSESDEKQVDYSNKNQ    60
spn_spr1771_R6_(2)           MKIMKKKYWTLAILFFCLFNNSVTAQEIPKNLDGNITHTQTSESFSESDEKQVDYSNKNQ    60
sp_E5Q10_09305_R6_CIB17      MKIMKKKYWTLAILFFCLFNNSVTAQEIPKNLDGNITHTQTSESFSESDEKQVDYSNKNQ    60
                             ***********************************************************

spn_TIGR4_sp_1954_(4)        EEVDQNKFRIQIDKTELFVTTDKHLEKNCCKLELEPQINNDIVNSESNNLLGEDNLDNKI   117
spn_D39_spd_1753_(2)         EEVDQNKFRIQIDKTELFVTTDKHLEKNCCKLELEPQINNDIVNSESNNLLGEDNLDNKI   120
spn_spn23F19760_ST81_(23F)   EEVDQNKFRIQIDKTELFVTTDKHLEKNCCKLELEPQINNDIVNSESNNLLGEDNLDNKI   120
spn_spj_1948_JJA_(14)        EEVDQNKFRIQIDKTELFVTTDKHLEKNCCKLELEPQINNDIVNSESNNLLGEDNLDNKI   120
spn_spr1771_R6_(2)           EEVDQNKFRIQIDKTELFVTTDKHLEKNCCKLELEPQINNDIVNSESNNLLGEDNLDNKI   120
sp_E5Q10_09305_R6_CIB17      EEVDQNKFRIQIDKTELFVTTDKHLEKNCCKLELEPQINNDIVNSESNNLLGEDNLDNKI   120
                             ***********************************************************

spn_TIGR4_sp_1954_(4)        KENVSHLDNRGGNIEHDKDNLESSIVRKYEWDIDKVTGGGESYKLYSKSNSKVSIAILDS   177
spn_D39_spd_1753_(2)         KENVSHLDNRGGNIEHDKDNLESSIVRKYEWDIDKVTGGGESYKLYSKSNSKVSIAILDS   180
spn_spn23F19760_ST81_(23F)   KENVSHLDNRGGNIEHDKDNLESSIVRKYEWDIDKVTGGGESYKLYSKSNSKVSIAILDS   180
spn_spj_1948_JJA_(14)        KENVSHLDNRGGNIEHDKDNLESSIVRKYEWDIDKVTGGGESYKLYSKSNSKVSIAILDS   180
spn_spr1771_R6_(2)           KENVSHLDNRGGNIEHDKDNLESSIVRKYEWDIDKVTGGGESYKLYSKSNSKVSIAILDS   180
sp_E5Q10_09305_R6_CIB17      KENVSHLDNRGGNIEHDKDNLESSIVRKYEWDIDKVTGGGESYKLYSKSNSKVSIAILDS   180
                             ***********************************************************

spn_TIGR4_sp_1954_(4)        GVDLQNTGLLKNLSNHSKNYVPNKGYLGKEEGEEGIISDIQDRLGHGTAVVAQIVGDDNI   237
spn_D39_spd_1753_(2)         GVDLQNTGLLKNLSNHSKNYVPNKGYLGKEEGEEGIISDIQDRLGHGTAVVAQIVGDDNI   240
spn_spn23F19760_ST81_(23F)   GVDLQNTGLLKNLSNHSKNYVPNKGYLGKEEGEEGIISDIQDRLGHGTAVVAQIVGDDNI   240
spn_spj_1948_JJA_(14)        GVDLQNTGLLKNLSNHSKNYVPNKGYLGKEEGEEGIISDIQDRLGHGTAVVAQIVGDDNI   240
spn_spr1771_R6_(2)           GVDLQNTGLLKNLSNHSKNYVPNKGYLGKEEGEEGIISDIQDRLGHGTAVVAQIVGDDNI   240
sp_E5Q10_09305_R6_CIB17      GVDLQNTGLLKNLSNHSKNYVPNKGYLGKEEGEEGIISDIQDRLGHGTAVVAQIVGDDNI   240
                             ***********************************************************

spn_TIGR4_sp_1954_(4)        NGVNPHVNINVYRIFGKSSASPDWIVKAIFDAVDDGNDIINLSTGQYLMIDGEYEDGTND   297
spn_D39_spd_1753_(2)         NGVNPHVNINVYRIFGKSSASPDWIVKAIFDAVDDGNDIINLSTGQYLMIDGEYEDGTND   300
spn_spn23F19760_ST81_(23F)   NGVNPHVNINVYRIFGKSSASPDWIVKAIFDAVDDGNDIINLSTGQYLMIDGEYEDGTND   300
spn_spj_1948_JJA_(14)        NGVNPHVNINVYRIFGKSSASPDWIVKAIFDAVDDGNDIINLSTGQYLMIDGEYEDGTND   300
spn_spr1771_R6_(2)           NGVNPHVNINVYRIFGKSSASPDWIVKAIFDAVDDGNDIINLSTGQYLMIDGEYEDGTND   300
sp_E5Q10_09305_R6_CIB17      NGVNPHVNINVYRIFGKSSASPDWIVKAIFDAVDDGNDIINLSTGQYLMIDGEYEDGTND   300
                             ***********************************************************

spn_TIGR4_sp_1954_(4)        FETFLKYKKAIDYANQKGVIIVAALGNDSLNVSNQSDLLKLISSRKKVRKPGLVVDVPSY   357
spn_D39_spd_1753_(2)         FETFLKYKKAIDYANQKGVIIVAALGNDSLNVSNQSDLLKLISSRKKVRKPGLVVDVPSY   360
spn_spn23F19760_ST81_(23F)   FETFLKYKKAIDYANQKGVIIVAALGNDSLNVSNQSDLLKLISSRKKVRKPGLVVDVPSY   360
spn_spj_1948_JJA_(14)        FETFLKYKKAIDYANQKGVIIVAALGNDSLNVSNQSDLLKLISSRKKVRKPGLVVDVPSY   360
spn_spr1771_R6_(2)           FETFLKYKKAIDYANQKGVIIVAALGNDSLNVSNQSDLLKLISSRKKVRKPGLVVDVPSY   360
sp_E5Q10_09305_R6_CIB17      FETFLKYKKAIDYANQKGVIIVAALGNDSLNVSNQSDLLKLISSRKKVRKPGLVVDVPSY   360
                             ***********************************************************

spn_TIGR4_sp_1954_(4)        FSSTISVGGIDRLGNLSDFSNKGDSDAIYAPAGSTLSLSELGLNNFINAEKYKEDWIFSA   417
spn_D39_spd_1753_(2)         FSSTISVGGIDRLGNLSDFSNKGDSDAIYAPAGSTLSLSELGLNNFINAEKYKEDWIFSA   420
spn_spn23F19760_ST81_(23F)   FSSTISVGGIDRLGNLSDFSNKGDSDAIYAPAGSTLSLSELGLNNFINAEKYKEDWIFSA   420
spn_spj_1948_JJA_(14)        FSSTISVGGIDRLGNLSDFSNKGDSDAIYAPAGSTLSLSELGLNNFINAEKYKEDWIFSA   420
spn_spr1771_R6_(2)           FSSTISVGGIDRLGNLSDFSNKGDSDAIYAPAGSTLSLSELGLNNFINAEKYKEDWIFSA   420
sp_E5Q10_09305_R6_CIB17      FSSTISVGGIDRLGNLSDFSNKGDSDAIYAPAGSTLSLSELGLNNFINAEKYKEDWIFSA   420
                             ***********************************************************

spn_TIGR4_sp_1954_(4)        TLGGYTYLYGNSFAAPKVSGAIAMIIDKYKLKDQPYNYMFVKKFWKKHYQ----------   467
spn_D39_spd_1753_(2)         TLGGYTYLYGNSFAAPKVSGAIAMIIDKYKLKDQPYNYMFVKKILEETLPVKNGIKVLNI   480
spn_spn23F19760_ST81_(23F)   TLGGYTYLYGNSFAAPKVSGAIAMIIDKYKLKDQPYNYMFVKKILEETLPVKNGIKVLNI   480
spn_spj_1948_JJA_(14)        TLGGYTYLYGNSFAAPKVSGAIAMIIDKYKLKDQPYNYMFVKKILEETLPVKNGIKVLNI   480
spn_spr1771_R6_(2)           TLGGYTYLYGNSFAAPKVSGAIAMIIDKYKLKDQPYNYMFVKKILEETLPVKNGIKVLNI   480
sp_E5Q10_09305_R6_CIB17      TLGGYTYLYGNSFAAPKVSGAIAMIIDKYKLKDQPYNYMFVKKILEETLPVKNGIKVLNI   480
                             *********************************************:  ::

spn_TIGR4_sp_1954_(4)        ------------------------------------------------------------   467
spn_D39_spd_1753_(2)         PNVLRYDLNMLQLEYKNEQSWDSFIDNVNLIELEERIQTTIGIKQINTHNIITIAREGYS   540
spn_spn23F19760_ST81_(23F)   PNVLRYDLNMLQLEYKNEQSWDSFIDNVNLIELEERIQTTIGIKQINTHNIITIAREGYS   540
spn_spj_1948_JJA_(14)        PNVLRYDLNMLQLEYKNEQSWDSFIDNVNLIELEERIQTTIGIKQINTHNIITIAREGYS   540
spn_spr1771_R6_(2)           PNVLRYDLNMLQLEYKNEQSWDSFIDNVNLIELEERIQTTIGIKQINTHNIITIAREGYS   540
sp_E5Q10_09305_R6_CIB17      PNVLRYDLNMLQLEYKNEQSWDSFIDNVNLIELEERIQTTIGIKQINTHNIITIAREGYS   540

spn_TIGR4_sp_1954_(4)        ----------------------------------------        467
spn_D39_spd_1753_(2)         QNYLPNTSENTYNSLQVSLVGVLLLFISMVNILWAKKSK        579
spn_spn23F19760_ST81_(23F)   QNYLPNTSENTYNSLQVSLVGVLLLFISMVNILWAKKSK        579
spn_spj_1948_JJA_(14)        QNYLPNTSENTYNSLQVSLVGVLLLFISMVNILWAKKSK        579
spn_spr1771_R6_(2)           QNYLPNTSENTYNSLQVSLVGVLLLFISMVNILWAKKSK        579
sp_E5Q10_09305_R6_CIB17      QNYLPNTSENTYNSLQVSLVGVLLLFISMVNILWAKKSK        579
```

8.4.1.3 Amino acid sequence comparisons of PrtA in *S. pneumoniae*

```
spn_ST556_snd:MYY_0688_(19F)      MKKSTVLSLTTAAVILAAYAPNEVVLADTSNSEDALNISDKEKVVLDKETENKEKHEDIH     60
spn_G54_spg_0584_(19F)            MKKSTVLSLTTAAVILAAYAPNEVVLADTSNSEDALSISDKEKVVVDKETENKEKHKDIH     60
spn_D39_spd_0558_(2)              MKKSTVLSLTTAAVILAAYAPNEVVLADTSSSEDALSISDKEKVVVDKETENKEKHKDIH     60
spn_R6_spr0561_(2)                MKKSTVLSLTTAAVILAAYAPNEVVLADTSSSEDALSISDKEKVVVDKETENKEKHKDIH     60
sp_Hungary19A-6_sph_0733_(19A)    MKKSTVLSLTTAAVILAAYAPHEVVLADTSNSEDALSISDKEKVA-----ENKEKHENIH     55
spn_TIGR4_sp_0641_(4)             MKKSTVLSLTTAAVILAAYAPNEVVLADTSSSEDALNISDKEKVA-----ENKEKHENIH     55
spn_ST81_spn23F05790_(23F)        MKKSTVLSLTTAAVILAAYAPNEVVLADTSSSEDALNISDKEKVVLDKETENKEKHEDIH     60
spn_EF3030_EF3030_03025_(19F)     MKKSTVLSLTTAAVILAAYAPNEVVLADTSSSEDALSISDKEKVA-----ENKEKHKDIH     55
spn_JJA_spj_0592_(14)             MKKSTVLSLTTAAVILAAYAPNEVVLADTSSSEDALSISDKEKVA-----ENKEKHEDIH     55
                                  ********************:********.*****.*******.      ******:.:**


spn_ST556_snd:MYY_0688_(19F)      NAIETSQDFKEKKTAVIEGKAVVSKNPVTDNNTSNEEARIKE-DSNQSQGDHTDSFVNKN    119
spn_G54_spg_0584_(19F)            NAIETSKDTEEKKTTIIEEKEVVSKNPVIDNNTSNEEARIKE-DSNQSQGDNAHSSANKN    119
spn_D39_spd_0558_(2)              NAIETSKDTEEKKTTIIEEKEVVSKNPVIDTKTSNEEAKIKEENSNQSQGDHTDSFVNKN    120
spn_R6_spr0561_(2)                NAIETSKDTEEKKTTIIEEKEVVSKNPVIDTKTSNEEAKIKEENSNQSQGDHTDSFVNKN    120
sp_Hungary19A-6_sph_0733_(19A)    SAIETSQDFKEKKTAVIEGKEVVSKNPVTDNNTSNEEARIKE-DSNQSQGDHTDSFVNKN    114
spn_TIGR4_sp_0641_(4)             SAMETSQDFKEKKTAVIKEKEVVSKNPVIDNNTSNEEAKIKEENSNKSQGDYTDSFVNKN    115
spn_ST81_spn23F05790_(23F)        NAIETSKDTEEKKTTVIEEKEVVSKNPVIDTKTSNEEARIKED-SNQSQGDNAHSSANKG    119
spn_EF3030_EF3030_03025_(19F)     NAIETSKDTEEKKTTIIEEKEVVSKNPVIDTKTSNEEAKIKEENSNKSQGDHTDSFVNKN    115
spn_JJA_spj_0592_(14)             SAMETSQDTEEKKTTIIEEKEVVSKNPVIDTKTSNEEAKIKEENSNKSQGDHTDSFVNKN    115
                                  .*:***:* :****::*: * ******* *.:******:*** **:**** :.* .**.


spn_ST556_snd:MYY_0688_(19F)      TENPKKEDKVVYIAEFKDKESGEKAIKQLSNLKNTKVLYTYDRIFNGSAIETTPDNLDKI    179
spn_G54_spg_0584_(19F)            TENPKKEDKVVYIAEFKDKESGEKAIKQLSSLKNTKVLYTYDRIFNGSAIETTPDNLDKI    179
spn_D39_spd_0558_(2)              TENPKKEDKVVYIAEFKDKESGEKAIKGLSNLKNTKVLYTYDRIFNGSAIETTPDNLDKI    180
spn_R6_spr0561_(2)                TENPKKEDKVVYIAEFKDKESGEKAIKGLSNLKNTKVLYTYDRIFNGSAIETTPDNLDKI    180
sp_Hungary19A-6_sph_0733_(19A)    TENPKKEDKVVYIAEFKDKESGEKAIKQLSNLKNTKVLYTYDRIFNGSAIETTPDNLDKI    174
spn_TIGR4_sp_0641_(4)             TENPKKEDKVVYIAEFKDKESGEKAIKELSSLKNTKVLYTYDRIFNGSAIETTPDNLDKI    175
spn_ST81_spn23F05790_(23F)        TENPKKEDRLVYIAEFKDKESGEKAIKQLSNLKNTKVLYTYDRIFNGSAIETTPDNLDKI    179
spn_EF3030_EF3030_03025_(19F)     TENPKKEDKVVYIAEFKDKESGSKAIKELSSLKNTKVLYTYDRIFNGGAIETTQDNLNKI    175
spn_JJA_spj_0592_(14)             TENPKKEDKVVYIAEFKDKESGEKAIKELSSLKNTKVLYTYDRIFNGGAIETTQDNLNKI    175
                                  ********::************.**** **.****************.***** ***:**


spn_ST556_snd:MYY_0688_(19F)      KQIEGITSVERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGT    239
spn_G54_spg_0584_(19F)            KQIEGISSVERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGT    239
spn_D39_spd_0558_(2)              KQIEGISSIERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGT    240
spn_R6_spr0561_(2)                KQIEGISSIERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGT    240
sp_Hungary19A-6_sph_0733_(19A)    KQIEGITSVERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGT    234
spn_TIGR4_sp_0641_(4)             KQIEGISSVERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGT    235
spn_ST81_spn23F05790_(23F)        KQIEGITSVERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGT    239
spn_EF3030_EF3030_03025_(19F)     KQIEGITSVERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGT    235
spn_JJA_spj_0592_(14)             KQIEGITSVERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGT    235
                                  ******:*:*************************************************


spn_ST556_snd:MYY_0688_(19F)      DYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDY    299
spn_G54_spg_0584_(19F)            DYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDY    299
spn_D39_spd_0558_(2)              DYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDY    300
spn_R6_spr0561_(2)                DYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDY    300
sp_Hungary19A-6_sph_0733_(19A)    DYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDY    294
spn_TIGR4_sp_0641_(4)             DYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDY    295
spn_ST81_spn23F05790_(23F)        DYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDY    299
spn_EF3030_EF3030_03025_(19F)     DYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDY    295
spn_JJA_spj_0592_(14)             DYRHKAMRIDDDAKASMRFKKEDLKGTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDY    295
                                  ************************************************************


spn_ST556_snd:MYY_0688_(19F)      FDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIE    359
spn_G54_spg_0584_(19F)            FDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIE    359
spn_D39_spd_0558_(2)              FDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIE    360
spn_R6_spr0561_(2)                FDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIE    360
sp_Hungary19A-6_sph_0733_(19A)    FDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIE    354
spn_TIGR4_sp_0641_(4)             FDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIE    355
spn_ST81_spn23F05790_(23F)        FDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIE    359
spn_EF3030_EF3030_03025_(19F)     FDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIE    355
spn_JJA_spj_0592_(14)             FDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIE    355
                                  ************************************************************


spn_ST556_snd:MYY_0688_(19F)      DSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVA    419
spn_G54_spg_0584_(19F)            DSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVA    419
spn_D39_spd_0558_(2)              DSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVA    420
spn_R6_spr0561_(2)                DSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVA    420
sp_Hungary19A-6_sph_0733_(19A)    DSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVA    414
spn_TIGR4_sp_0641_(4)             DSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVA    415
spn_ST81_spn23F05790_(23F)        DSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVA    419
spn_EF3030_EF3030_03025_(19F)     DSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVA    415
spn_JJA_spj_0592_(14)             DSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVA    415
                                  ************************************************************
```

8.4.1.3 Continued:

```
spn_ST556_snd:MYY_0688_(19F)        NNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKNKITT   479
spn_G54_spg_0584_(19F)              NNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKNKITT   479
spn_D39_spd_0558_(2)                NNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKNKITT   480
spn_R6_spr0561_(2)                  NNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKNKITT   480
sp_Hungary19A-6_sph_0733_(19A)      NNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGQSFKYRNIGAFFDKNKITT   474
spn_TIGR4_sp_0641_(4)               NNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKSKITT   475
spn_ST81_spn23F05790_(23F)          NNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKNKITT   479
spn_EF3030_EF3030_03025_(19F)       NNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKNKITT   475
spn_JJA_spj_0592_(14)               NNHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKNKITT   475
                                    **********************************:*************.****

spn_ST556_snd:MYY_0688_(19F)        NEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIM   539
spn_G54_spg_0584_(19F)              NEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIM   539
spn_D39_spd_0558_(2)                NEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIM   540
spn_R6_spr0561_(2)                  NEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIM   540
sp_Hungary19A-6_sph_0733_(19A)      NEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIM   534
spn_TIGR4_sp_0641_(4)               NEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIM   535
spn_ST81_spn23F05790_(23F)          NEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIM   539
spn_EF3030_EF3030_03025_(19F)       NEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIM   535
spn_JJA_spj_0592_(14)               NEDGTKAPSKLKFVYIGKGQDQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIM   535
                                    ************************************************************

spn_ST556_snd:MYY_0688_(19F)        VVNTVNYYNRDNWKELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPDKKTEVKRNNKE   599
spn_G54_spg_0584_(19F)              VVNTVNYYNRDNWTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPDKKTEVKRNNKE   599
spn_D39_spd_0558_(2)                VVNTVNYYNRDNWTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPDKKTEVKRNNKE   600
spn_R6_spr0561_(2)                  VVNTVNYYNRDNWTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPDKKTEVKRNNKE   600
sp_Hungary19A-6_sph_0733_(19A)      VVNTVNYYNRDNWTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPNKKTEVKRNNKE   594
spn_TIGR4_sp_0641_(4)               VVNTVNYYNRDNWTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPDKKTEVKRNNKE   595
spn_ST81_spn23F05790_(23F)          VVNTVNYYNRDNWTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPNKKTEVKRNNKE   599
spn_EF3030_EF3030_03025_(19F)       VVNTVNYYNRDNWTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPNKKTEVKRNNKE   595
spn_JJA_spj_0592_(14)               VVNTVNYYNRDNWTELPAMGYEADEGTKSQVFSISGDDGVKLWNMINPNKKTEVKRNNKE   595
                                    *************.************************************:**********

spn_ST556_snd:MYY_0688_(19F)        DFKDKLEQYYPIDMESFNSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRI   659
spn_G54_spg_0584_(19F)              DFKDKLEQYYPIDMESFNSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRI   659
spn_D39_spd_0558_(2)                DFKDKLEQYYPIDMESFNSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRI   660
spn_R6_spr0561_(2)                  DFKDKLEQYYPIDMESFNSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRI   660
sp_Hungary19A-6_sph_0733_(19A)      DFKDKLEQYYPIDMESFNSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRI   654
spn_TIGR4_sp_0641_(4)               DFKDKLEQYYPIDMESFNSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRI   655
spn_ST81_spn23F05790_(23F)          DFKDKLEQYYPIDMESFNSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRI   659
spn_EF3030_EF3030_03025_(19F)       DFKDKLEQYYPIDMESFNSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRI   655
spn_JJA_spj_0592_(14)               DFKDKLEQYYPIDMESFNSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTSWGPRI   655
                                    ************************************************************

spn_ST556_snd:MYY_0688_(19F)        DLLLKPDVSAPGKNIKSTLNVINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPV   719
spn_G54_spg_0584_(19F)              DLLLKPDVSAPGKNIKSTLNVINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPV   719
spn_D39_spd_0558_(2)                DLLLKPDVSAPGKNIKSTLNVINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPV   720
spn_R6_spr0561_(2)                  DLLLKPDVSAPGKNIKSTLNVINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPV   720
sp_Hungary19A-6_sph_0733_(19A)      DLLLKPDVSAPGKNIKSTLNVINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPV   714
spn_TIGR4_sp_0641_(4)               DLLLKPDVSAPGKNIKSTLNVINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPV   715
spn_ST81_spn23F05790_(23F)          DLLLKPDVSAPGKNIKSTLNVINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPV   719
spn_EF3030_EF3030_03025_(19F)       DLLLKPDVSAPGKNIKSTLNVINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPV   715
spn_JJA_spj_0592_(14)               DLLLKPDVSAPGKNIKSTLNVINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPV   715
                                    ************************************************************

spn_ST556_snd:MYY_0688_(19F)        LKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQQGAGLINVANALRNE   779
spn_G54_spg_0584_(19F)              LKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQQGAGLINVANALRNE   779
spn_D39_spd_0558_(2)                LKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQQGAGLINVANALRNE   780
spn_R6_spr0561_(2)                  LKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQQGAGLINVANALRNE   780
sp_Hungary19A-6_sph_0733_(19A)      LKNLTGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQQGAGLINVANALRNE   774
spn_TIGR4_sp_0641_(4)               LKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQQGAGLINVANALRNE   775
spn_ST81_spn23F05790_(23F)          LKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQQGAGLINVANALRNE   779
spn_EF3030_EF3030_03025_(19F)       LKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQQGAGLINVANALRNE   775
spn_JJA_spj_0592_(14)               LKNLKGDDKIDLTSLTKIALQNTARPMMDATSWKEKSQYFASPRQQGAGLINVANALRNE   775
                                    ****.*******************************************************

spn_ST556_snd:MYY_0688_(19F)        VVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAVTTDSLTD   839
spn_G54_spg_0584_(19F)              VVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTD   839
spn_D39_spd_0558_(2)                VVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTD   840
spn_R6_spr0561_(2)                  VVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTD   840
sp_Hungary19A-6_sph_0733_(19A)      VVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTD   834
spn_TIGR4_sp_0641_(4)               VVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTD   835
spn_ST81_spn23F05790_(23F)          VVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTD   839
spn_EF3030_EF3030_03025_(19F)       VVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTD   835
spn_JJA_spj_0592_(14)               VVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTIKLHNTSNRPLTFKVSASAITTDSLTD   835
                                    **********************************************:*.******
```

8.4.1.3 Continued:

```
spn_ST556_snd:MYY_0688_(19F)        RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAK    899
spn_G54_spg_0584_(19F)              RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAK    899
spn_D39_spd_0558_(2)                RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHGTFTIGANSSFDLNAVINVGEAK    900
spn_R6_spr0561_(2)                  RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHGTFTIGANSSFDLNAVINVGEAK    900
sp_Hungary19A-6_sph_0733_(19A)      RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAK    894
spn_TIGR4_sp_0641_(4)               RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAK    895
spn_ST81_spn23F05790_(23F)          RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAK    899
spn_EF3030_EF3030_03025_(19F)       RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAK    895
spn_JJA_spj_0592_(14)               RLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVINVGEAK    895
                                    *********************************.*********************

spn_ST556_snd:MYY_0688_(19F)        NKNKFVESFIHFESVEEMEALSSNGKKTDFQPSLSMPLMGFAGNWNHEPILDKWAWEEGS    959
spn_G54_spg_0584_(19F)              NKNKFVESFIHFESVEEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGS    959
spn_D39_spd_0558_(2)                NKNKFVESFIHFESVEEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGS    960
spn_R6_spr0561_(2)                  NKNKFVESFIHFESVEEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGS    960
sp_Hungary19A-6_sph_0733_(19A)      NKNKFVESFIHFESVEEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGS    954
spn_TIGR4_sp_0641_(4)               NKNKFVESFIHFESVEEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGS    955
spn_ST81_spn23F05790_(23F)          NKNKFVESFIHFESVEEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGS    959
spn_EF3030_EF3030_03025_(19F)       NKNKFVESFIHFESVEEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGS    955
spn_JJA_spj_0592_(14)               NKNKFVESFIHFESVEEMEALNSNGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGS    955
                                    ********************.***** :******************************

spn_ST556_snd:MYY_0688_(19F)        KSKTMEGYDDDGKPKIPGTLNKGIGGEHGIDKFNPAGVIQNRRDKNRTSLDQDPDLFAFN   1019
spn_G54_spg_0584_(19F)              RSKTLGGYDDDGKPKIPGTLNKGIGGEHGIDKFNPAGVIQNRRDKNRTSLDQDPDLFAFN   1019
spn_D39_spd_0558_(2)                RSKTLGGYDDDGKPKIPGTLNKGIGGEHGIDKFNPAGVIQNRRDKNTTSLDQNPELFAFN   1020
spn_R6_spr0561_(2)                  RSKTLGGYDDDGKPKIPGTLNKGIGGEHGIDKFNPAGVIQNRRDKNTTSLDQNPELFAFN   1020
sp_Hungary19A-6_sph_0733_(19A)      RSKTLGGYDDDGKPKIPGTLNKGIGGEHGIDKFNPAGVIQNRRDKNTTSLDQNPELFAFN   1014
spn_TIGR4_sp_0641_(4)               RSKTLGGYDDDGKPKIPGTLNKGIGGEHGIDKFNPAGVIQNRRDKNTTSLDQNPELFAFN   1015
spn_ST81_spn23F05790_(23F)          RSKTLGGYDDDGKPKIPGTLNKGIGGEHGIDKFNPAGVIQNRRDKNTTSLDQNPELFAFN   1019
spn_EF3030_EF3030_03025_(19F)       RSKTLGGYDDDGKPKIPGTLNKGIGGEHGIDKFNPAGVIQNRRDKNTTSLDQNPELFAFN   1015
spn_JJA_spj_0592_(14)               RSKTLGGYDDDGKPKIPGTLNKGIGGEHGIDKFNPAGVIQNRRDKNTTSLDQNPELFAFN   1015
                                    :***: *****************************************  *****:*:*****

spn_ST556_snd:MYY_0688_(19F)        NQGVHAESTSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNTNKEGE   1079
spn_G54_spg_0584_(19F)              NQGVHAESTSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNTNKEGE   1079
spn_D39_spd_0558_(2)                NQGINAPSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNTNKEGE   1080
spn_R6_spr0561_(2)                  NQGINAPSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNTNKEGE   1080
sp_Hungary19A-6_sph_0733_(19A)      NQGINAPSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNTNKEGE   1074
spn_TIGR4_sp_0641_(4)               NEGINAPSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNTNKEGE   1075
spn_ST81_spn23F05790_(23F)          NQGINAPSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNTNKEGE   1079
spn_EF3030_EF3030_03025_(19F)       NQGINAPSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNTNKEGE   1075
spn_JJA_spj_0592_(14)               NQGINAPSSSGSKIANIYPLDSNGNPQDAQLERGLTPSPLVLRSAEEGLISIVNTNKEGE   1075
                                    *:*::* *:************************************************

spn_ST556_snd:MYY_0688_(19F)        NQKDLKVVSREHFIRGILNSKGNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKD   1139
spn_G54_spg_0584_(19F)              NQKDLKVVSREHFIRGILNSKGNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKD   1139
spn_D39_spd_0558_(2)                NQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKD   1140
spn_R6_spr0561_(2)                  NQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKD   1140
sp_Hungary19A-6_sph_0733_(19A)      NQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKD   1134
spn_TIGR4_sp_0641_(4)               NQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKD   1135
spn_ST81_spn23F05790_(23F)          NQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKD   1139
spn_EF3030_EF3030_03025_(19F)       NQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKD   1135
spn_JJA_spj_0592_(14)               NQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRGREENAPESKD   1135
                                    **:****:************.***************************************

spn_ST556_snd:MYY_0688_(19F)        NQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEK   1199
spn_G54_spg_0584_(19F)              NQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEK   1199
spn_D39_spd_0558_(2)                NQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEK   1200
spn_R6_spr0561_(2)                  NQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEK   1200
sp_Hungary19A-6_sph_0733_(19A)      NQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEK   1194
spn_TIGR4_sp_0641_(4)               NQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEK   1195
spn_ST81_spn23F05790_(23F)          NQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEK   1199
spn_EF3030_EF3030_03025_(19F)       NQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEK   1195
spn_JJA_spj_0592_(14)               NQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEK   1195
                                    ********************************************************

spn_ST556_snd:MYY_0688_(19F)        IKLITKDTYHKVKDQYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEV   1259
spn_G54_spg_0584_(19F)              IKLITKDTYHKVKDQYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEV   1259
spn_D39_spd_0558_(2)                IKLITKDTYHKVKDQYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEV   1260
spn_R6_spr0561_(2)                  IKLITKDTYHKVKDQYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEV   1260
sp_Hungary19A-6_sph_0733_(19A)      IKLITKDTYHKVKDQYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEV   1254
spn_TIGR4_sp_0641_(4)               IKLITKDTYHKVKDQYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEV   1255
spn_ST81_spn23F05790_(23F)          IKLITKDTYHKVKDQYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEV   1259
spn_EF3030_EF3030_03025_(19F)       IKLITKDTYHKVKDQYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEV   1255
spn_JJA_spj_0592_(14)               IKLITKDTYHKVKDQYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEV   1255
                                    ************************************************************
```

8.4.1.3 Continued:

```
spn_ST556_snd:MYY_0688_(19F)    TYAGEGQGRNRKLDKDGNTIYEISGAGDLRGKIIEVIALDGASNFTKIHRIKFANQADEK    1319
spn_G54_spg_0584_(19F)          TYAGEGQGRNRKLDKDGNTIYEISGAGDLRGKIIEVIALDGASNFTKIHRIKFANQADEK    1319
spn_D39_spd_0558_(2)            TYAGEGQGRNRKLDKDGNTIYEIKGAGDLRGKIIEVIALDGSSNFTKIHRIKFADQADEK    1320
spn_R6_spr0561_(2)              TYAGEGQGRNRKLDKDGNTIYEIKGAGDLRGKIIEVIALDGSSNFTKIHRIKFADQADEK    1320
sp_Hungary19A-6_sph_0733_(19A)  TYAGEGQGRNRKLDKDGNTIYEIKGAGDLRGKIIEVIALDGSSNFTKIHRIKFANHADEK    1314
spn_TIGR4_sp_0641_(4)           TYAGEGQGRNRKLDKDGNTIYEIKGAGDLRGKIIEVIALDGSSNFTKIHRIKFANQADEK    1315
spn_ST81_spn23F05790_(23F)      TYAGEGQGRNRKLDKDGNTIYEIKGAGDLRGKIIEVIALDGSSNFTKIHRIKFANQADEK    1319
spn_EF3030_EF3030_03025_(19F)   TYAGEGQGRNRKLDKDGNTIYEIKGAGDLRGKIIEVIALDGSSNFTKIHRIKFANQADEK    1315
spn_JJA_spj_0592_(14)           TYAGEGQGRNRKLDKDGNTIYEIKGAGDLRGKIIEVIALDGSSNFTKIHRIKFANQADEK    1315
                                ***********************.******************:***********::.****

spn_ST556_snd:MYY_0688_(19F)    GMISYYLVDPDQDSSKYQKLGEIPESKFKNLKNVKDDSLNKETAEVENNLLVDNQSIEGK    1379
spn_G54_spg_0584_(19F)          GMISYYLVDPDQDSSKYQKLGEIPESKFKNLKNVKDDXLNKETAEVENNLLVDNQSIEGK    1379
spn_D39_spd_0558_(2)            GMISYYLVDPDKDASKYEKLGEISEDKLKNAKSPEENTNNNQA-KDEDSKPDEKSSVEGE    1379
spn_R6_spr0561_(2)              GMISYYLVDPDKDASKYEKLGEISEDKLKNAKSPEENTNNNQA-KDEDSKPDEKSSVEGE    1379
sp_Hungary19A-6_sph_0733_(19A)  GMISYYLVDPDKDASKYEKLGEISEDKLKNAKSPEENTNNNQA-KDEDSKPDEKSSVEGE    1373
spn_TIGR4_sp_0641_(4)           GMISYYLVDPDQDSSKYQKLGEIAESKFKNLGNGKEGSLKKDTTGVEHHHQENEESIKEK    1375
spn_ST81_spn23F05790_(23F)      GMISYYLVDPDQDSSKYQKLGEIAESKFKNLGNGKEGSLKKDTTGVEHHHQENEESIKEK    1379
spn_EF3030_EF3030_03025_(19F)   GMISYYLVDPDQDSSKYQKLGEIAESKFKNLGNGKEGSLKKDTTGVEHHHQENEESIKEK    1375
spn_JJA_spj_0592_(14)           GMISYYLVDPDQDSSKYQKLGEIAESKFKNLGNGKEGSLKKDTTEVENH-QGNEEATEEK    1374
                                ***********:*:***:***** *.*:**   . ::.  :::*    *.    ::.: : :

spn_ST556_snd:MYY_0688_(19F)    SLFNIHKTISTIRDFENKDLKKLIKKKYKQEDDFVN-GGTRTVERDYKYDDKGNIIAYDD    1438
spn_G54_spg_0584_(19F)          SLFNIHKTISTIRDFENKDLKKLIKKKYKQEDDFVN-GGTRTVERDYKYDDKGNIIAYDD    1438
spn_D39_spd_0558_(2)            ASLEINKTISTIREFENKDLKKLIKKKFREVNDFTSETGKRIEEYDYKYDDKGNIIAYDD    1439
spn_R6_spr0561_(2)              ASLEINKTISTIREFENKDLKKLIKKKFREVNDFTSETGKRIEEYDYKYDDKGNIIAYDD    1439
sp_Hungary19A-6_sph_0733_(19A)  ASLEINKTISTIRDFENKDLKKLIKKKFREVDDFTSETGKRIEEYDYKYDDKGNIIAYDD    1433
spn_TIGR4_sp_0641_(4)           SSFTIDRNISTIRDFENKDLKKLIKKKFREVDDFTSETGKRMEEYDYKYDDKGNIIAYDD    1435
spn_ST81_spn23F05790_(23F)      SSFTIDRNISTIRDFENKDLKKLIKKKFREVDDFTSETGKRIEEYDYKYDDKGNIIAYDD    1439
spn_EF3030_EF3030_03025_(19F)   SSFTIDRNISTIRDFENKDLKKLIKKKYKEEDDFVT-GGKRTVELDYKYDDKGNITAYED    1434
spn_JJA_spj_0592_(14)           SSFTIDRNISTIRDFENKDLKKLIKKKFREVEDFVT-GGKRTVELDYKYDDKGNITAYED    1433
                                : : *.:.*****:************::: :**..  *.*  * ********* **:*

spn_ST556_snd:MYY_0688_(19F)    GTDLEYETEKLDEIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNNYKSIEIKATK    1498
spn_G54_spg_0584_(19F)          GTDLEYETEKLDEIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNNYKSIEIKATK    1498
spn_D39_spd_0558_(2)            GSALQYETEKFDEIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNSYKSIEIKATK    1499
spn_R6_spr0561_(2)              GSALQYETEKFDEIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNSYKSIEIKATK    1499
sp_Hungary19A-6_sph_0733_(19A)  GSALQYETEKFDEIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNSYKSIEIKATK    1493
spn_TIGR4_sp_0641_(4)           GTDLEYETEKLDEIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNNYKSIEIKATK    1495
spn_ST81_spn23F05790_(23F)      GSALQYETEKLDEIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNSYKSIEIKATK    1499
spn_EF3030_EF3030_03025_(19F)   ESALEYETEKFDEIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNNYKFIEIKATK    1494
spn_JJA_spj_0592_(14)           ESALEYETEKFDEIKSKIYGVLSPSKDGHFEILGKISKVSKNAKVYYGNSYKSIEIKATK    1493
                                : *:*****:**************************:**********.** *******

spn_ST556_snd:MYY_0688_(19F)    YDFHSKTMTFDLYANINDIVDGLAFAGDMRLFVKDNDQKKAEIKIRMPEKIKETKSEYPY    1558
spn_G54_spg_0584_(19F)          YDFHSKTMTFDLYANINDIVDGLAFAGDMRLFVKDNDQKKAEIKIRMPEKIKETKSEYPY    1558
spn_D39_spd_0558_(2)            YDSHSKTMIFDLYANINDIVDGLAFAGDMRLFVKDDNQIKAETKIRMPEKNKETKAEYPY    1559
spn_R6_spr0561_(2)              YDSHSKTMIFDLYANINDIVDGLAFAGDMRLFVKDDNQIKAETKIRMPEKNKETKAEYPY    1559
sp_Hungary19A-6_sph_0733_(19A)  YDSHSKTMIFDLYANINDIVDGLAFAGDMRLFVKDDNQIKAETKIRMPEKNKETKAEYPY    1553
spn_TIGR4_sp_0641_(4)           YDFHSKTMTFDLYANINDIVDGLAFAGDMRLFVKDNDQKKAEIKIRMPEKIKETKSEYPY    1555
spn_ST81_spn23F05790_(23F)      YDSHSKTMTFDLYANINDIVDGLAFAGDMRFFVKDDNRIKAETKIRMPEKNKETKAEYPY    1559
spn_EF3030_EF3030_03025_(19F)   YDSHSKTMTFDLYANINDIVDGLAFAGDMRFFVKDDDRIKAETKIRMPEKNKETKAEYPY    1554
spn_JJA_spj_0592_(14)           YDSHSKTMTFDLYANINDIVDGLAFAGDMRFFVKDDDRIKAETKIRMPEKNKETKAEYPY    1553
                                ** ***** ********************:****::: *** ******* ****:****

spn_ST556_snd:MYY_0688_(19F)    VSSYGNVIELGEGDLSKNKPDNLTKMESGKIYSDSEKQQYLLKDNIILRKGYALKVTTYN    1618
spn_G54_spg_0584_(19F)          VSSYGNVIELGEGDLSKNKPDNLTKMESGKIYSDSEKQQYLLKDNIILRKGYALKVTTYN    1618
spn_D39_spd_0558_(2)            VSSYGNVIELGEGDLSKNKPDNLTKMESGKIYSDSEKQQYLLKDNIILRKGYALKVTTYN    1619
spn_R6_spr0561_(2)              VSSYGNVIELGEGDLSKNKPDNLTKMESGKIYSDSEKQQYLLKDNIILRKGYALKVTTYN    1619
sp_Hungary19A-6_sph_0733_(19A)  VSSYGNVIELGEGDLSKNKPDNLTKMESGKIYSDSEKQQYLLKDNIILRKGYALKVTTYN    1613
spn_TIGR4_sp_0641_(4)           VSSYGNVIELGEGDLSKNKPDNLTKMESGKIYSDSEKQQYLLKDNIILRKGYALKVTTYN    1615
spn_ST81_spn23F05790_(23F)      VSSYGNVIELGEGDLSKNKPDNLTKMESGKIYSDSEKQQYLLKDNIILRKGYALKVTTYN    1619
spn_EF3030_EF3030_03025_(19F)   VSSYGNVIELGEGDLSKNKPDNLTKMESGKIYSDSEKQQYLLKDNIILRKGYALKVTTYN    1614
spn_JJA_spj_0592_(14)           VSSYGNVIELGEGDLSKNKPDNLTKMESGKIYSDSEKQQYLLKDNIILRKGYALKVTTYN    1613
                                ************************************************************

spn_ST556_snd:MYY_0688_(19F)    PGKTDMLEGNGVYSKEDIAKIQKANPNLRVLSETTIYADSRNVEDGRSTQAVLMSALDGF    1678
spn_G54_spg_0584_(19F)          PGKTDMLEGNGVYSKEDIAKIQKANPNLRALSETTIYADSRNVEDGRSTQSVLMSALDGF    1678
spn_D39_spd_0558_(2)            PGKTDMLEGNGVYSKEDIAKIQKANPNLRVLSETTIYADSRNVEDGRSTQAVLMSALDGF    1679
spn_R6_spr0561_(2)              PGKTDMLEGNGVYSKEDIAKIQKANPNLRVLSETTIYADSRNVEDGRSTQAVLMSALDGF    1679
sp_Hungary19A-6_sph_0733_(19A)  PGKTDMLEGNGVYSKEDIAKIQKANPNLRVLSETTIYADSRNVEDGRSTQAVLMSALDGF    1673
spn_TIGR4_sp_0641_(4)           PGKTDMLEGNGVYSKEDIAKIQKANPNLRALSETTIYADSRNVEDGRSTQSVLMSALDGF    1675
spn_ST81_spn23F05790_(23F)      PGKTDMLEGNGVYSKEDIAKIQKANPNLRVLSETTIYADSRNVEDGRSTQAVLMSALDGF    1679
spn_EF3030_EF3030_03025_(19F)   PGKTDMLEGNGVYSKEDIAKIQKANPNLRVLSETTIYADSRNVEDGRSTQAVLMSALDGF    1674
spn_JJA_spj_0592_(14)           PGKTDMLEGNGVYSKEDIAKIQKANPNLRVLSETTIYADSRNVEDGRSTQAVLMSALDGF    1673
                                *****************************.********************:********
```

8.4.1.3 Continued:

```
spn_ST556_snd:MYY_0688_(19F)        NIIRYQVFTFKMNDKGEAIDKDGNLVTDSSKLVLFGKDDKEYTGEDKSNVEAIKEDGSML   1738
spn_G54_spg_0584_(19F)              NIIRYQVFTFKMNDKGEAIDKDGNLVTDSSKLVLFGKDDKEYTGEDKSNVEAIKEDGSML   1738
spn_D39_spd_0558_(2)                NIIRYQVFTFKMNDKGEAIDKDGNLVTDSSKLVLFGKDDKEYTGEDKSNVEAIKEDGSML   1739
spn_R6_spr0561_(2)                  NIIRYQVFTFKMNDKGEAIDKDGNLVTDSSKLVLFGKDDKEYTGEDKSNVEAIKEDGSML   1739
sp_Hungary19A-6_sph_0733_(19A)      NIIRYQVFTFKMNDKGEAIDKDGNLVTDSSKLVLFGKDDKEYTGEDKSNVEAIKEDGSML   1733
spn_TIGR4_sp_0641_(4)               NIIRYQVFTFKMNDKGEAIDKDGNLVTDSSKLVLFGKDDKEYTGEDKFNVEAIKEDGSML   1735
spn_ST81_spn23F05790_(23F)          NIIRYQVFTFKMNDKGEAIDKDGNLVTDSSKLVLFGKDDKEYTGEDKSNVEAIKEDGSML   1739
spn_EF3030_EF3030_03025_(19F)       NIIRYQVFTFKMNDKGEAIDKDGNLVTDSSKLVLFGKDDKEYTGEDKSNVEAIKEDGSML   1734
spn_JJA_spj_0592_(14)               NIIRYQVFTFKMNDKGEAIDKDGNLVTDSSKLVLFGKDDKEYTGEDKSNVEAIKEDGSML   1733
                                    ************************************************ ************


spn_ST556_snd:MYY_0688_(19F)        FIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLIY   1798
spn_G54_spg_0584_(19F)              FIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLXY   1798
spn_D39_spd_0558_(2)                FIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLIY   1799
spn_R6_spr0561_(2)                  FIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLIY   1799
sp_Hungary19A-6_sph_0733_(19A)      FIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLIY   1793
spn_TIGR4_sp_0641_(4)               FIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLIY   1795
spn_ST81_spn23F05790_(23F)          FIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLIY   1799
spn_EF3030_EF3030_03025_(19F)       FIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLIY   1794
spn_JJA_spj_0592_(14)               FIDTKPVNLSMDKNYFNPSKSNKIYVRNPEFYLRGKISDKGGFNWELRVNESVVDNYLIY   1793
                                    ********************************************************** *


spn_ST556_snd:MYY_0688_(19F)        GDLHIDNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTDMDGNVYLQTGYSDLNAKAV   1858
spn_G54_spg_0584_(19F)              GDLHIDNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTDMDGNVYXQTGYSDLNAKAV   1858
spn_D39_spd_0558_(2)                GDLHIDNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTDMDGNVYLQTGYSDLNAKAV   1859
spn_R6_spr0561_(2)                  GDLHIDNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTDMDGNVYLQTGYSDLNAKAV   1859
sp_Hungary19A-6_sph_0733_(19A)      GDLHIDNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTDMDGNVYLQTGYSDLNAKAV   1853
spn_TIGR4_sp_0641_(4)               GDLHIDNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTDMDGNVYLQTGYSDLNAKAV   1855
spn_ST81_spn23F05790_(23F)          GDLHIDNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTDMDGNVYLQTGYSDLNAKAV   1859
spn_EF3030_EF3030_03025_(19F)       GDLHIDNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTDMDGNVYLQTGYSDLNAKAV   1854
spn_JJA_spj_0592_(14)               GDLHIDNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTDMDGNVYLQTGYSDLNAKAV   1853
                                    ************************************************ ***********


spn_ST556_snd:MYY_0688_(19F)        GVHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNNGFDGEIQEQHIYVNGKE   1918
spn_G54_spg_0584_(19F)              GVHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNNGFDGEIQEXHIYINGKE   1918
spn_D39_spd_0558_(2)                GVHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNNGFDGEIQEQHIYVNGKE   1919
spn_R6_spr0561_(2)                  GVHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNNGFDGEIQEQHIYVNGKE   1919
sp_Hungary19A-6_sph_0733_(19A)      GVHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNNGFDGEIQEQHIYVNGKE   1913
spn_TIGR4_sp_0641_(4)               GVHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNNGFDGEIQEQHIYINGKE   1915
spn_ST81_spn23F05790_(23F)          GVHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNNGFDGEIQEQHIYVNGKE   1919
spn_EF3030_EF3030_03025_(19F)       GVHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNNGFDGEIQEQHIYVNGKE   1914
spn_JJA_spj_0592_(14)               GVHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNNGFDGEIQEQHIYVNGKE   1913
                                    *************************************************** ***:****


spn_ST556_snd:MYY_0688_(19F)        YTSFDDIKQITDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNGKEM   1978
spn_G54_spg_0584_(19F)              YTSFNDIKQIIDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNGKEM   1978
spn_D39_spd_0558_(2)                YTSFDDIKQITDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNGKEM   1979
spn_R6_spr0561_(2)                  YTSFDDIKQITDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNGKEM   1979
sp_Hungary19A-6_sph_0733_(19A)      YTSFDDIKQITDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNGKEM   1973
spn_TIGR4_sp_0641_(4)               YTSFNDIKQIIDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNGKEM   1975
spn_ST81_spn23F05790_(23F)          YTSFDDIKQITDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNGKEM   1979
spn_EF3030_EF3030_03025_(19F)       YTSFDDIKQITDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNGKEM   1974
spn_JJA_spj_0592_(14)               YTSFDDIKQITDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNGKEM   1973
                                    ****:***** *************************************************


spn_ST556_snd:MYY_0688_(19F)        SSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEE   2038
spn_G54_spg_0584_(19F)              SSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEE   2038
spn_D39_spd_0558_(2)                SSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEE   2039
spn_R6_spr0561_(2)                  SSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEE   2039
sp_Hungary19A-6_sph_0733_(19A)      SSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEE   2033
spn_TIGR4_sp_0641_(4)               SSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEE   2035
spn_ST81_spn23F05790_(23F)          SSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEE   2039
spn_EF3030_EF3030_03025_(19F)       SSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEE   2034
spn_JJA_spj_0592_(14)               SSTIVSEEDFILPVYKGELEKGYQFDGWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEE   2033
                                    ************************************************************


spn_ST556_snd:MYY_0688_(19F)        KKEEENKPTFDVSKKKDNPQVNHSQLNESHRKEDLQREDHSQKSDSTKDVTATVLDKNNI   2098
spn_G54_spg_0584_(19F)              KKEEENKPTFDVSKKKDNPQVNHSQLNESHRKEDLQREDHSQKSDSTKDVTATVLDKNNI   2098
spn_D39_spd_0558_(2)                KKEEENKPTFDVSKKKDNPQVNHSQLNESHRKEDLQREDHSQKSDSTKDVTATVLDKNNI   2099
spn_R6_spr0561_(2)                  KKEEENKPTFDVSKKKDNPQVNHSQLNESHRKEDLQREDHSQKSDSTKDVTATVLDKNNI   2099
sp_Hungary19A-6_sph_0733_(19A)      KKEEENKPTFDVSKKKDNPQVNHSQLNESHRKEDLQREDHSQKSDSTKDVTATVLDKNNI   2093
spn_TIGR4_sp_0641_(4)               KKEEENKPTFDVSKKKDNPQVNHSQLNESHRKEDLQREEHSQKSDSTKDVTATVLDKNNI   2095
spn_ST81_spn23F05790_(23F)          KKEEENKPTFDVSKKKDNPQVNHSQLNESHRKEDLQREDHSQKSDSTKDVTATVLDKNNI   2099
spn_EF3030_EF3030_03025_(19F)       KKEEENKPTFDVSKKKDNPQVNHSQLNESHRKEDLQREDHSQKSDSTKDVTATVLDKNNI   2094
spn_JJA_spj_0592_(14)               KKEEENKPTFDVSKKKDNPQVNHSQLNESHRKEDLQREDHSQKSDSTKDVTATVLDKNNI   2093
                                    *********************************:**********************


spn_ST556_snd:MYY_0688_(19F)        SSKSTTNNPNKLPKTGTASGAQTLLAAGIMFIVGIFLGLKKKNQD  2143
spn_G54_spg_0584_(19F)              SSKSTTNNPNKLPKTGTASGAQTLLAAGIMFIVGIFLGLKKKNQD  2143
spn_D39_spd_0558_(2)                SSKSTTNNPNKLPKTGTASGAQTLLAAGIMFIVGIFLGLKKKNQD  2144
spn_R6_spr0561_(2)                  SSKSTTNNPNKLPKTGTASGAQTLLAAGIMFIVGIFLGLKKKNQD  2144
sp_Hungary19A-6_sph_0733_(19A)      SSKSTTNNPNKLPKTGTASGAQTLLAAGIMFIVGIFLGLKKKNQ-  2137
spn_TIGR4_sp_0641_(4)               SSKSTTNNPNKLPKTGTASGAQTLLAAGIMFIVGIFLGLKKKNQD  2140
spn_ST81_spn23F05790_(23F)          SSKSTTNNPNKLPKTGTASGAQTLLAAGIMFIVGIFLGLKKKNQD  2144
spn_EF3030_EF3030_03025_(19F)       SSKSTTNNPNKLPKTGTASGAQTLLAAGIMFIVGIFLGLKKKNQD  2139
spn_JJA_spj_0592_(14)               SSKSTTNNPNKLPKTGTASGAQTLLAAGIMFIVGIFLGLKKKNQD  2138
                                    *********************************************
```

8.4.1.4 Amino acid sequence comparisons of HtrA in *S. pneumoniae*

```
spn_G54_spg_2188_(19F)           ----MKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYK    56
spn_D39_spd_2068_(2)             ----MKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYK    56
sp_Hungary19A-6_sph_2438_(19A)   ----MKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYK    56
spn_TIGR4_sp_2239_(4)            ----MKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYK    56
spn_ST556_snd:MYY_2162_(19F)     ----MKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYK    56
spn_R6_spr2045_(2)               MEANMKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYK    60
sp_R6_CIB17_E5Q10_10910_(2)      MEANMKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYK    60
spn_EF3030_EF3030_11105_(19F)    MEANMKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYK    60
spn_ST81_spn23F22720_(23F)       ----MKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYK    56
spn_JJA_spj_2269_(14)            ----MKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYK    56
                                     ********************************************************

spn_G54_spg_2188_(19F)           NENSTTQAVNKVKDAVVSVITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAY   116
spn_D39_spd_2068_(2)             NENSTTQAVNKVKD-VVSVITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAY   115
sp_Hungary19A-6_sph_2438_(19A)   NENSTTQAVNKVKDAVVSVITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAY   116
spn_TIGR4_sp_2239_(4)            NENSTTQAVNKVKDAVVSVITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAY   116
spn_ST556_snd:MYY_2162_(19F)     NENSTTQAVNKVKDAVVSVITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAY   116
spn_R6_spr2045_(2)               NENSTTQAVNKVKDAVVSVITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAY   120
sp_R6_CIB17_E5Q10_10910_(2)      NENSTTQAVNKVKDAVVSVITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAY   120
spn_EF3030_EF3030_11105_(19F)    NENSTTQAVNKVKDAVVSVITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAY   120
spn_ST81_spn23F22720_(23F)       NENSTTQAVNKVKDAVVSVITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAY   116
spn_JJA_spj_2269_(14)            NENSTTQAVNKVKDAVVSVITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAY   116
                                 *************** *********************************************

spn_G54_spg_2188_(19F)           IVTNNHVINGASKVDIRLSDGTKVPGEIVGADTFSDIAVVKISSEKVTTVAEFGDSSKLT   176
spn_D39_spd_2068_(2)             IVTNNHVINGASKVDIRLSDGTKVPGEIVGADTFSDIAVVKISSEKVTTVAEFGDSSKLT   175
sp_Hungary19A-6_sph_2438_(19A)   IVTNNHVINGASKVDIRLSDGTKVPGEIVGADTFSDIAVVKISSEKVTTVAEFGDSSKLT   176
spn_TIGR4_sp_2239_(4)            IVTNNHVINGASKVDIRLSDGTKVPGEIVGADTFSDIAVVKISSEKVTTVAEFGDSSKLT   176
spn_ST556_snd:MYY_2162_(19F)     IVTNNHVINGASKVDIRLSDGTKVPGEIVGADTFSDIAVVKISSEKVTTVAEFGDSSKLT   176
spn_R6_spr2045_(2)               IVTNNHVINGASKVDIRLSDGTKVPGEIVGADTFSDIAVVKISSEKVTTVAEFGDSSKLT   180
sp_R6_CIB17_E5Q10_10910_(2)      IVTNNHVINGASKVDIRLSDGTKVPGEIVGADTFSDIAVVKISSEKVTTVAEFGDSSKLT   180
spn_EF3030_EF3030_11105_(19F)    IVTNNHVINGASKVDIRLSDGTKVPGEIVGADTFSDIAVVKISSEKVTTVAEFGDSSKLT   180
spn_ST81_spn23F22720_(23F)       IVTNNHVINGASKVDIRLSDGTKVPGEIVGADTFSDIAVVKISSEKVTTVAEFGDSSKLT   176
spn_JJA_spj_2269_(14)            IVTNNHVINGASKVDIRLSDGTKVPGEIVGADTFSDIAVVKISSEKVTTVAEFGDSSKLT   176
                                 ************************************************************

spn_G54_spg_2188_(19F)           VGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLRSEDGQAISTKAIQTDTAINPGNSGG   236
spn_D39_spd_2068_(2)             VGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLKSEDGQAISTKAIQTDTAINPGNSGG   235
sp_Hungary19A-6_sph_2438_(19A)   VGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLRSEDGQAISTKAIQTDTAINPGNSGG   236
spn_TIGR4_sp_2239_(4)            VGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLKSEDGQAISTKAIQTDTAINPGNSGG   236
spn_ST556_snd:MYY_2162_(19F)     VGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLKSEDGQAISTKAIQTDTAINPGNSGG   236
spn_R6_spr2045_(2)               VGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLKSEDGQAISTKAIQTDTAINPGNSGG   240
sp_R6_CIB17_E5Q10_10910_(2)      VGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLKSEDGQAISTKAIQTDTAINPGNSGG   240
spn_EF3030_EF3030_11105_(19F)    VGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLKSEDGQAISTKAIQTDTAINPGNSGG   240
spn_ST81_spn23F22720_(23F)       VGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLKSEDGQAISTKAIQTDTAINPGNSGG   236
spn_JJA_spj_2269_(14)            VGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLKSEDGQAISTKAIQTDTAINPGNSGG   236
                                 **********************************:*************************

spn_G54_spg_2188_(19F)           PLINIQGQVIGITSSKIATNGGTSVEGLGFAIPANDAINIIEQLEKNGKVTRPALGIQMV   296
spn_D39_spd_2068_(2)             PLINIQGQVIGITSSKIATNGGTSVEGLGFAIPANDAINIIEQLEKNGKVTRPALGIQMV   295
sp_Hungary19A-6_sph_2438_(19A)   PLINIQGQVIGITSSKIATNGGTSVEGLGFAIPANDAINIIEQLEKNGKVTRPALGIQMV   296
spn_TIGR4_sp_2239_(4)            PLINIQGQVIGITSSKIATNGGTSVEGLGFAIPANDAINIIEQLEKNGKVTRPALGIQMV   296
spn_ST556_snd:MYY_2162_(19F)     PLINIQGQVIGITSSKIATNGGTSVEGLGFAIPANDAINIIEQLEKNGKVTRPALGIQMV   296
spn_R6_spr2045_(2)               PLINIQGQVIGITSSKIATNGGTSVEGLGFAIPANDAINIIEQLEKNGKVTRPALGIQMV   300
sp_R6_CIB17_E5Q10_10910_(2)      PLINIQGQVIGITSSKIATNGGTSVEGLGFAIPANDAINIIEQLEKNGKVTRPALGIQMV   300
spn_EF3030_EF3030_11105_(19F)    PLINIQGQVIGITSSKIATNGGTSVEGLGFAIPANDAINIIEQLEKNGKVTRPALGIQMV   300
spn_ST81_spn23F22720_(23F)       PLINIQGQVIGITSSKIATNGGTSVEGLGFAIPANDAINIIEQLEKNGKVTRPALGIQMV   296
spn_JJA_spj_2269_(14)            PLINIQGQVIGITSSKIATNGGTSVEGLGFAIPANDAINIIEQLEKNGKVTRPALGIQMV   296
                                 ************************************************************

spn_G54_spg_2188_(19F)           NLSNVSTSDIRRLNIPSNVTSGVVVRSVQSNMPANGHLEKYDVITKVDDKEIASSTDLQS   356
spn_D39_spd_2068_(2)             NLSNVSTSDIRRLNIPSNVTSGVIVRSVQSNMPANGHLEKYDVITKVDDKEIASSTDLQS   355
sp_Hungary19A-6_sph_2438_(19A)   NLSNVSTSDIRRLNIPSNVTSGVVVRSVQSNMPANGHLEKYDVITKVDDKEIASSTDLQS   356
spn_TIGR4_sp_2239_(4)            NLSNVSTSDIRRLNIPSNVTSGVIVRSVQSNMPANGHLEKYDVITKVDDKEIASSTDLQS   356
spn_ST556_snd:MYY_2162_(19F)     NLSNVSTSDIRRLNIPSNVTSGVVVRSVQSNMPANGHLEKYDVITKVDDKEIASSTDLQS   356
spn_R6_spr2045_(2)               NLSNVSTSDIRRLNIPSNVTSGVIVRSVQSNMPANGHLEKYDVITKVDDKEIASSTDLQS   360
sp_R6_CIB17_E5Q10_10910_(2)      NLSNVSTSDIRRLNIPSNVTSGVIVRSVQSNMPANGHLEKYDVITKVDDKEIASSTDLQS   360
spn_EF3030_EF3030_11105_(19F)    NLSNVSTSDIRRLNIPSNVTSGVVVRSVQSNMPANGHLEKYDVITKVDDKEIASSTDLQS   360
spn_ST81_spn23F22720_(23F)       NLSNVSTSDIRRLNIPSNVTSGVVVRSVQSNMPANGHLEKYDVITKVDDKEIASSTDLQS   356
spn_JJA_spj_2269_(14)            NLSNVSTSDIRRLNIPSNVTSGVVVRSVQSNMPANGHLEKYDVITKVDDKEIASSTDLQS   356
                                 ***********************:************************************

spn_G54_spg_2188_(19F)           ALYNHSIGDTIKITYYRNGKEETTSTNLTRVQVI---   390
spn_D39_spd_2068_(2)             ALYNHSIGDTIKITYYRNGKEETTSIKLNKSSGDLES   392
sp_Hungary19A-6_sph_2438_(19A)   ALYNHSIGDTIKITYYRNGKEETTSIKLNKSSGDLES   393
spn_TIGR4_sp_2239_(4)            ALYNHSIGDTIKITYYRNGKEETTSIKLNKSSGDLES   393
spn_ST556_snd:MYY_2162_(19F)     ALYNHSIGDTIKITYYRNGKEETTSIKLNKSSGDLES   393
```
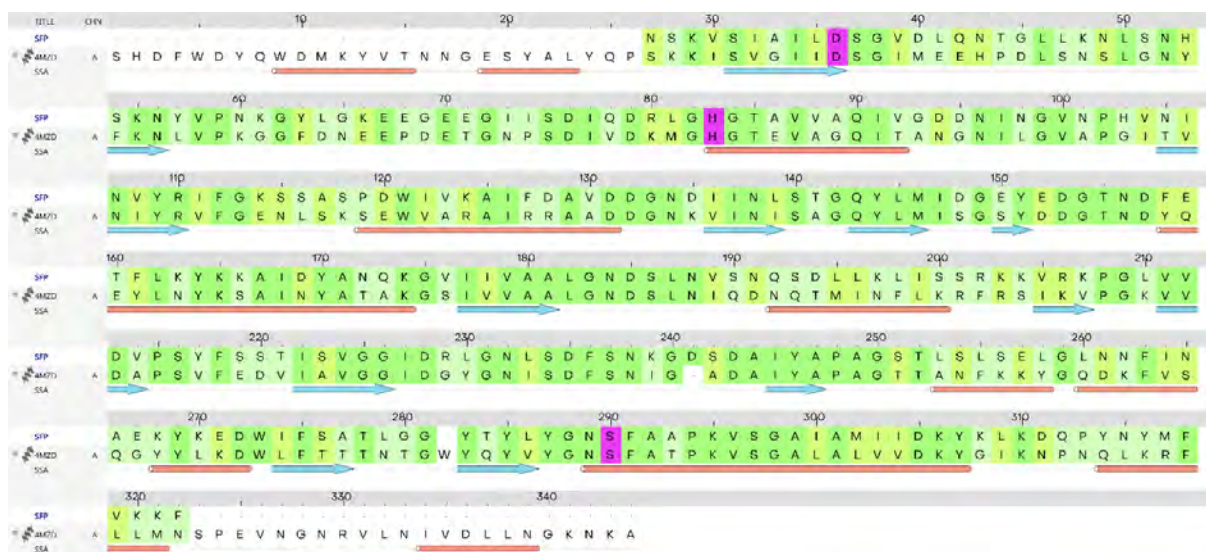
## 8.4.2 Sequence alignment of the serine proteases catalytic domain

Supplementary Figure 8.1: Sequence alignment of SFP catalytic domain

The sequence alignment of SFP catalytic domain (296 aa) with the corresponding template protein of the nisin leader peptidase NisP from *Lactococcus lactis* (4MZD) used for homology modeling (identical: green, similar: yellow, dissimilar: light green). The catalytic residues histidine (H), aspartate (D) and serine (S) are highlighted in purple. Red tubes and blue arrows depict helix and sheet secondary structures, respectively.
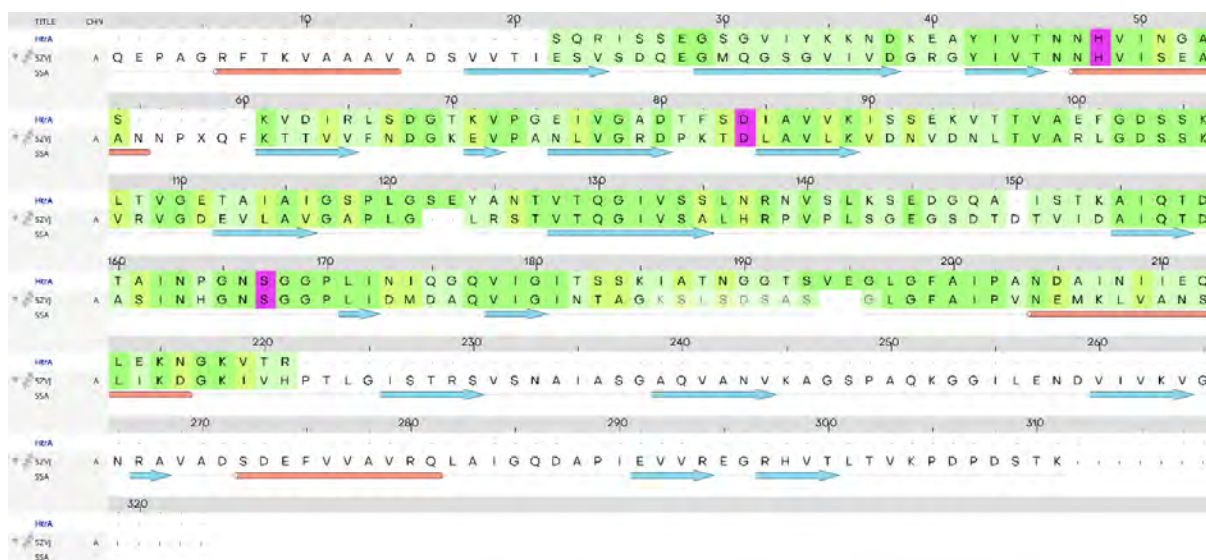
Supplementary Figure 8.2: Sequence alignment of HtrA catalytic domain

HtrA catalytic domain (182 aa) with the corresponding template protein of HtrA1 from *Mycobacterium tuberculosis* (5ZVJ) used for homology modeling (identical: green, similar: yellow, dissimilar: light green). The catalytic residues histidine (H), aspartate (D) and serine (S) are highlighted in purple. Red tubes and blue arrows depict helix and sheet secondary structures, respectively.
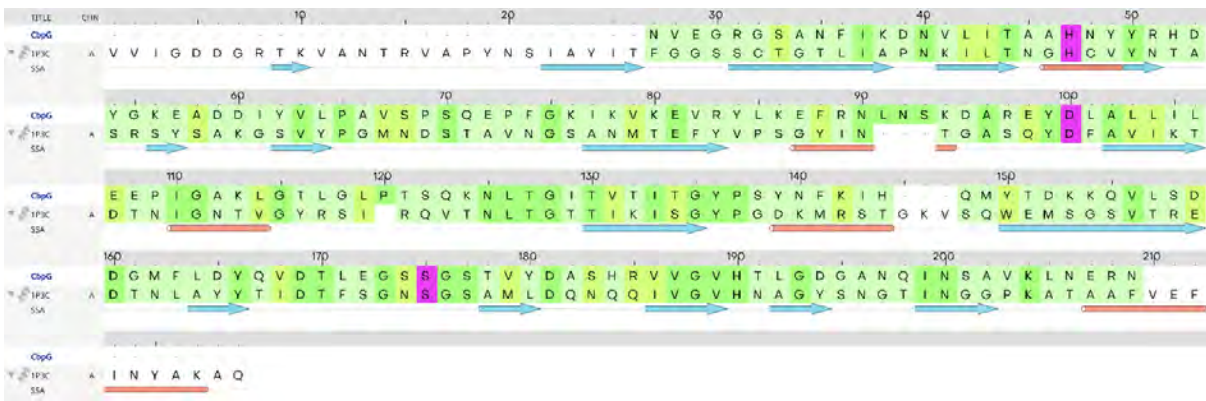
Supplementary Figure 8.3: Sequence alignment of CbpG catalytic domain

CbpG catalytic domain (184 aa) with the corresponding template protein glutamyl endopeptidase (1P3C) from *Bacillus intermedius* used for homology modeling (identical: green, similar: yellow, dissimilar: light green). The catalytic residues histidine (H), aspartate (D) and serine (S) are highlighted in purple. Red tubes and blue arrows depict helix and sheet secondary structures, respectively.
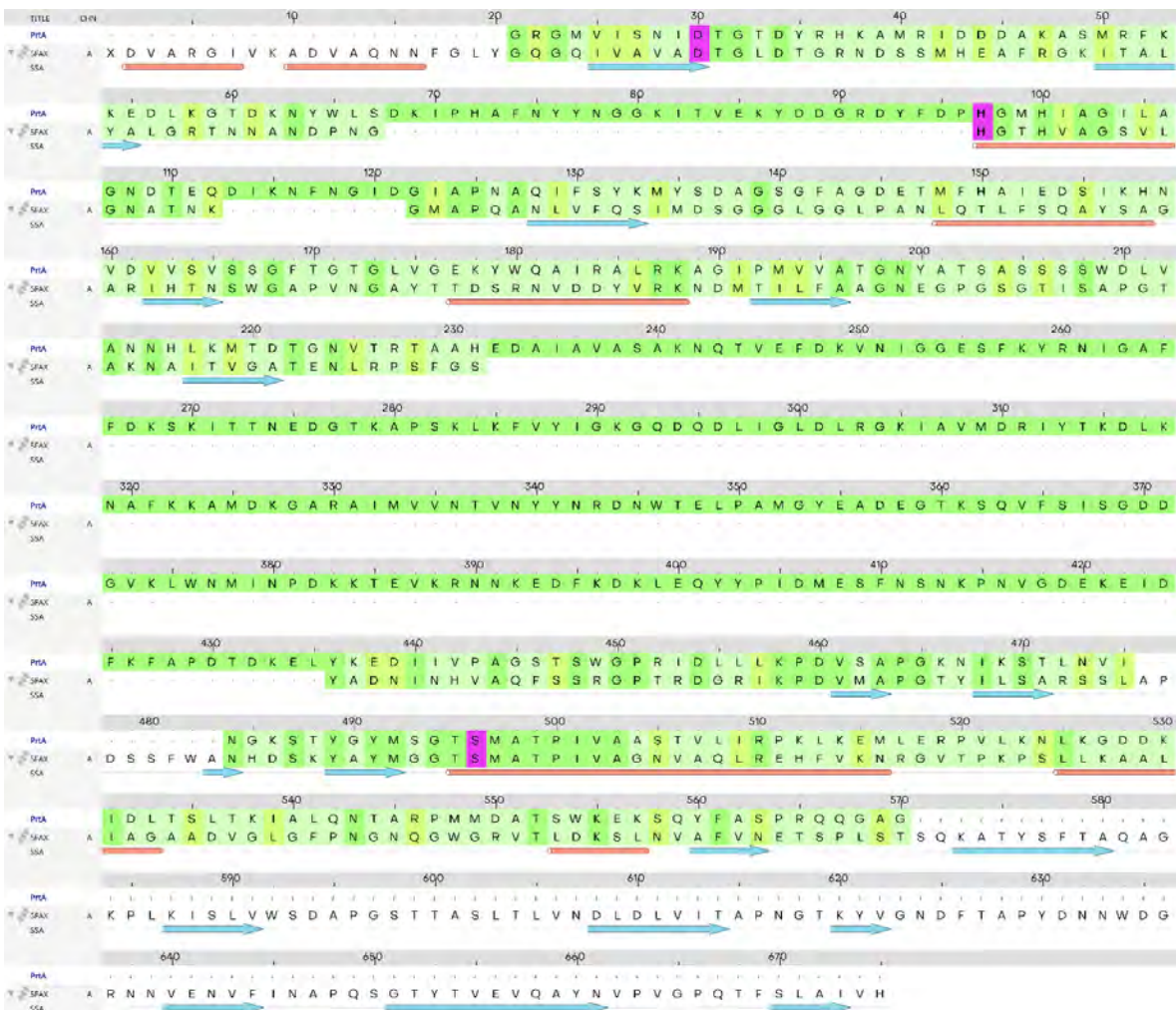
Supplementary Figure 8.4: Sequence alignment of PrtA catalytic domain

PrtA catalytic domain (542aa) with the corresponding template protein structure of subtilase SubHal from *Bacillus halmapalus* (5FAX) used for homology modeling (identical: green,

similar: yellow, dissimilar: light green). The catalytic residues histidine (H), aspartate (D) and serine (S) are highlighted in purple. Red tubes and blue arrows depict helix and sheet secondary structures, respectively.

8.5 Purification profile of serine proteases by affinity chromatography
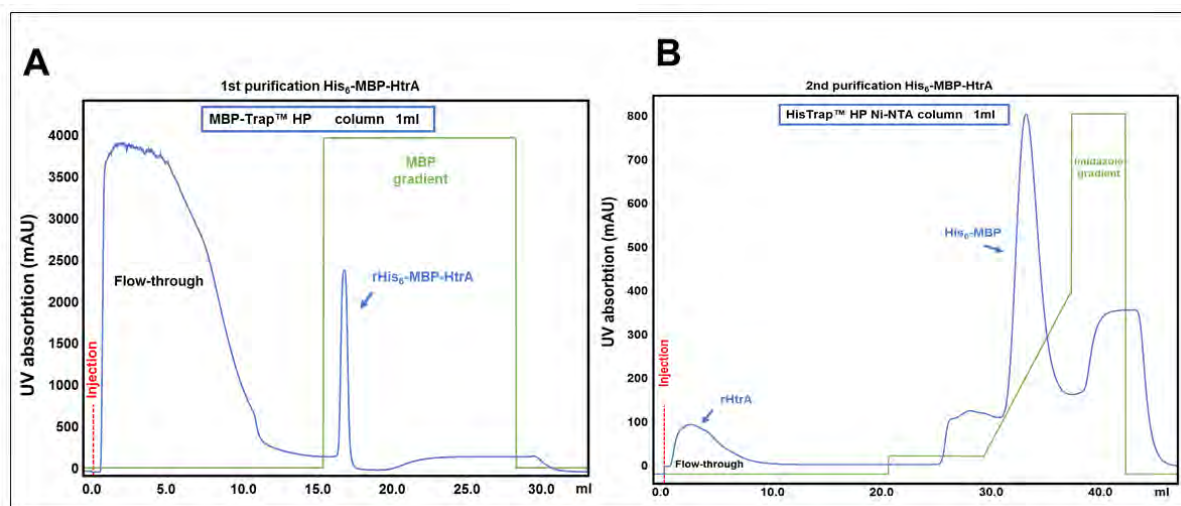
Supplementary Figure 8.5: Purification of His$_6$-MBP tagged HtrA by affinity chromatography.

(**A, B**) Elution profiles by Ni-affinity chromatography 1st with MBPTrap and the 2nd purification carried out with HisTrap column, purification was performed using ÄKTA purifier system.
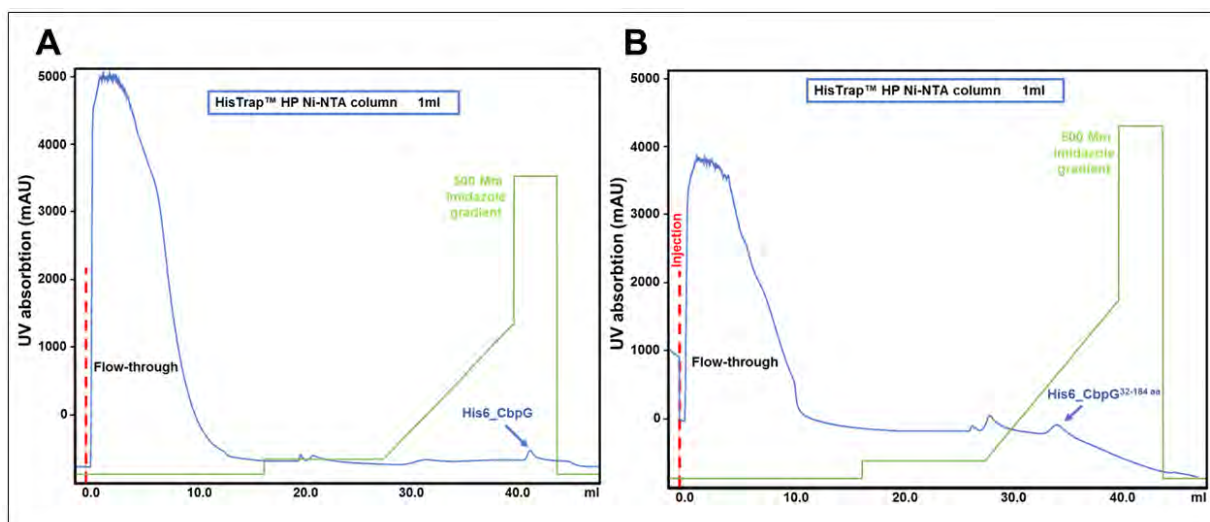
Supplementary Figure 8.6: Purification of His$_6$-CbpG by affinity chromatography.

His$_6$-tagged rCbpG (**A**) and rCbpG$^{32-184aa}$ (**B**) elution profile purification were purified by affinity liquid chromatography. Purification was performed by using the HisTrap™ Ni-NTA column with the ÄKTA purifier system. The extract protein was loaded on column and eluted with increased imidazole gradient (green line). The purified protein in the elution fractions was determined using an online UV absorption at 280 nm (blue line).
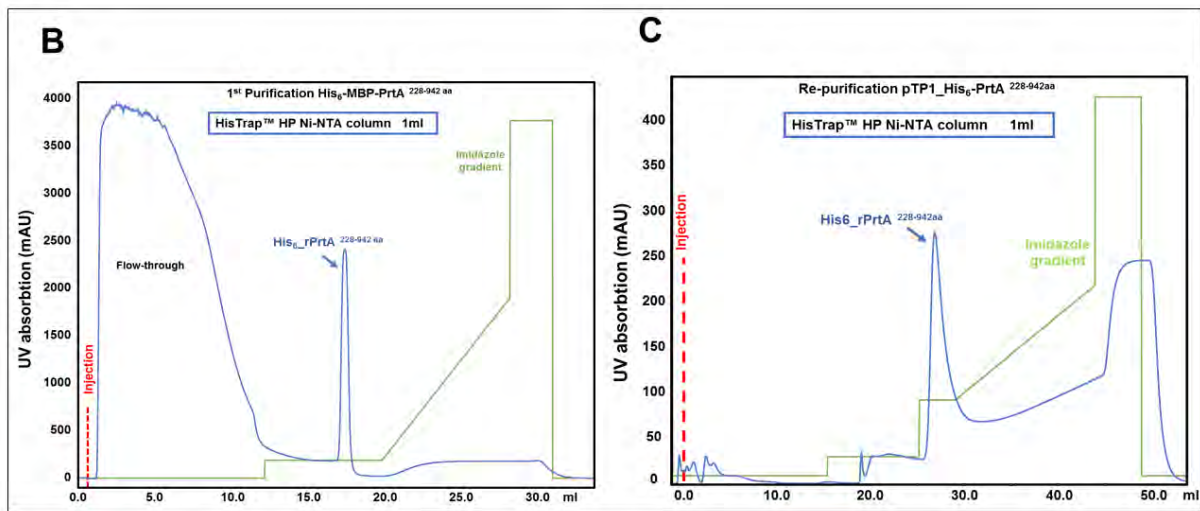
Supplementary Figure 8.7: Purification of His$_6$-PrtA.

Elution profiles for the 1st and 2nd protein purification by affinity chromatography. Purification was performed by a HisTrap™ HP Ni-NTA column using the ÄKTA purifier liquid chromatography system. His$_6$-PrtA$^{28-942aa}$ protein was eluted with increased imidazole gradient (green line). Purified protein in elution frictions was detected using the online UV absorption at 280 nm (blue line).
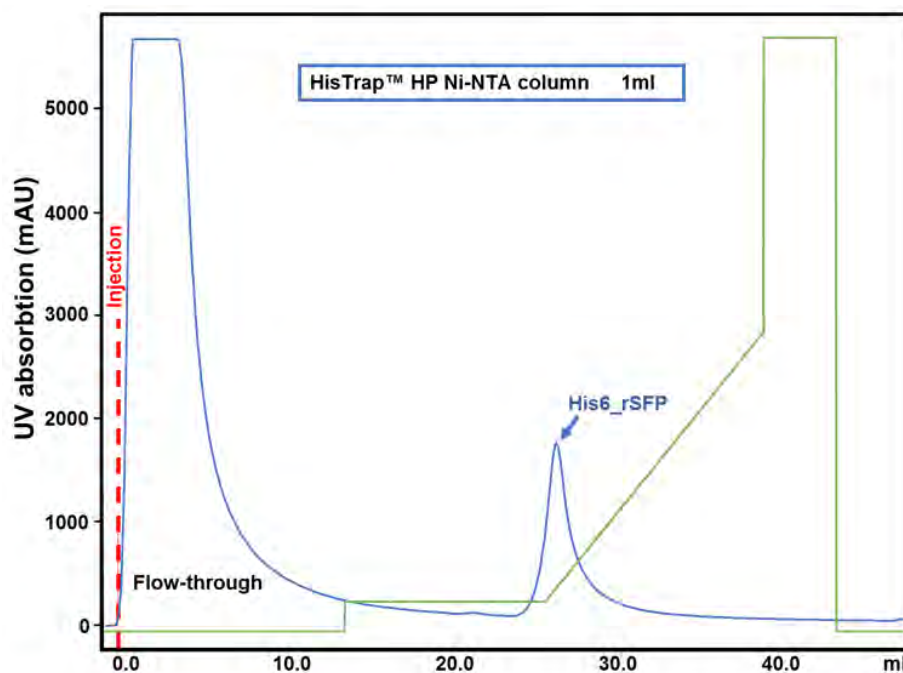
Supplementary Figure 8.8: Elution profile of SFP were purified by Ni affinity liquid chromatography.

Purification was performed by using the HisTrap™ Ni-NTA column with the "ÄKTA purifier" system. The extract protein was loaded on column and eluted with increased imidazole gradient (green line). The purified protein in the elution fractions was determined using an online UV absorption at 280 nm (blue line).

8.6 Supplementary tables related to bacterial fitness and growth behaviour

Supplementary Table 8.5: Growth rates of *S. pneumoniae* wild-type and isogenic mutants in THY and RPMI_modi medium.

Pneumococcal strain	Growth rate (μ) in THY [min⁻¹]	Growth rate (μ) in RPMI_modi [min⁻¹]
19F_EF3030 (wild type)	0.016	0.0031
19F_EF3030Δ*prtA*	0.015	0.0030
19F_EF3030Δ*htrA*	0.017	0.0028
19F_EF3030Δ*cbpG*	0.016	0.0028
19F_EF3030 (wild type)	0.0146	0.0045
19FΔ*htrA*Δ*cbpG* (*prtA*+)	0.0149	0.0045
19FΔ*prtA*Δ*cbpG* (*htrA*+)	0.0176	0.0048
19FΔ*prtA*Δ*htrA* (*cbpG*+)	0.0135	0.0041
19FΔ*htrA*Δ*cbpG*Δ*prtA* (*All proteases*)	0.0142	0.0052
TIGR4Δ*cps* (wild type)	0.0099	0.018
TIGR4Δ*cps*Δ*htrA*Δ*cbpG*Δ*sfp* (*prtA*+)	0.0096	0.013
TIGR4Δ*cps*Δ*prtA*Δ*cbpG*Δ*sfp* (*htrA*+)	0.0091	0.011
TIGR4Δ*cps*Δ*prtA*Δ*htrA*Δ*sfp* (*cbpG*+)	0.0090	0.009
TIGR4Δ*cps*Δ*htrA*Δ*prtA*Δ*cbpG* (*sfp*+)	0.0085	0.010
TIGR4*lux* (wild type)	0.0119	0.0041
TIGR4*lux*Δ*htrA*Δ*cbpG*Δ*sfp* (*prtA*+)	0.0123	0.0041
TIGR4*lux*Δ*prtA*Δ*cbpG*Δ*sfp* (*htrA*+)	0.0113	0.0041
TIGR4*lux*Δ*htrA*Δ*prtA*Δ*sfp* (*cbpG*+)	0.0081	0.0045
TIGR4*lux*Δ*htrA*Δ*prtA*Δ*cbpG* (*sfp*+)	0.0083	0.0045

8.7 Supplementary tables related to the impact of serine proteases Host-pathogen interaction

Supplementary Table 8.6: CFU of adherent pneumococci counted on blood agar plates after 4h of infection of Detroit-562 cells.

Pneumococcal strain	Adherent bacteria CFU\ml per 2.5 x 10⁵ cells	
	Mean ± SD	*P*-value
19F_EF3030 (wild type)	1162 ± 928.4	-
19FΔ*htrA*Δ*cbpG* (*prtA*+)	144.2 ± 140.1	0,0162
19FΔ*prtA*Δ*cbpG* (*htrA*+)	154.8 ± 151.4	0,0162
19FΔ*prtA*Δ*htrA* (*cbpG*+)	147.1 ± 175.3	0,0162
19FΔ*htrA*Δ*cbpG*Δ*prtA* (*non-function serine proteases*)	98.71 ± 109.6	0,0040
TIGR4Δ*cps* (wild type)	4870 ± 1294	-
TIGR4Δ*cps*Δ*htrA*Δ*cbpG*Δ*sfp* (*prtA*+)	2365 ± 1428	0.0591
TIGR4Δ*cps*Δ*prtA*Δ*cbpG*Δ*sfp* (*htrA*+)	1491 ± 1253	0.0294
TIGR4Δ*cps*Δ*prtA*Δ*htrA*Δ*sfp* (*cbpG*+)	1424 ± 8630	0.0294
TIGR4Δ*cps*Δ*htrA*Δ*prtA*Δ*cbp*G (*sfp*+)	8307 ± 2650	0.0294

*P value less than *0.05* was taken as statistically significant.

**Supplementary Table 8.7:** CFU from nasopharyngeal lavage counted on blood agar plates

Pneumococcal strains	CFU\ml in the nasopharyngeal lavage							
	Median Day 2	*P-* value	Median Day3	*P-* value	Median Day 7	*P-* value	Median Day 14	*P-* value
19F_EF3030 (wild type)	291000		210000		147000		23300	
19F*ΔhtrAΔcbpG* (*prtA*+)	31700	0,0004	27600	0,0004	25800	0,0005	7070	0,0163
19F*ΔprtAΔcbpG* (*htrA*+)	53700	0,0025	69200	0,0153	25200	0,0009	9950	0,0630
19F*ΔprtAΔhtrA* (*cbpG*+)	21600	0,0005	16700	0,0003	20000	0,0009	6060	0,0066
19F*ΔhtrAΔcbpGΔprtA*	33800	0,0005	35400	0,0012	36000	0,0081	1860	0,0015

*P value less than 0.05 was taken as statistically significant.

Supplementary Table 8.8: CFU from bronchoalveolar lavage counted on blood agar plates.

Pneumococcal strain	CFU\ml in the nasopharyngeal wash							
	Median Day 2	*P-* value	Median Day3	*P-* value	Median Day 7	*P-* value	Median Day 14	*P-* value
19F_EF3030 (wild type)	564,0		1795		6408		33,00	
19F*ΔhtrAΔcbpG* (*prtA*+)	15,00	0,1443	215,8	0,5753	30,00	0,0040	1,000	0,7819
19F*ΔprtAΔcbpG* (*htrA*+)	175,0	0,8519	50,00	0,0224	16,65	0,0852	233,0	0,5841
19F*ΔprtAΔhtrA* (*cbpG*+)	207,0	0,4327	108,0	0,1916	133,0	0,0733	1,000	0,0370
19F*ΔhtrAΔcbpGΔprtA*	1240	0,8813	425,0	0,6815	1,000	0,0911	1,000	0,0156

*P value less than 0.05 was taken as statistically significant.

Supplementary Table 8.9: CFU number of recovered intracellular pneumococci counted on blood agar plates after 30 min infection of J774 murine macrophages.

Pneumococcal strain	Recovered bacteria CFU\ml per 1 x 10^5 cells	
	Mean ± SD	*P*-value
TIGR4Δ*cps* (wild type)	10140 ± 761.8	-
TIGR4Δ*cps*Δ*htrA*Δ*cbpG*Δ*sfp* (*prtA*+)	310.3 ± 138.7	0.0162
TIGR4Δ*cps*Δ*prtA*Δ*cbpG*Δ*sfp* (*htrA*+)	269.0 ± 533.9	0.0040
TIGR4Δ*cps*Δ*prtA*Δ*htrA*Δ*sfp* (*cbpG*+)	9460 ± 391.4	0.4962
TIGR4Δ*cps*Δ*htrA*Δ*prtA*Δ*cbp*G (*sfp*+)	286.3 ± 710.7	0.0040

## 8.8 Supplementary tables related to the impact of serine biofilm formations
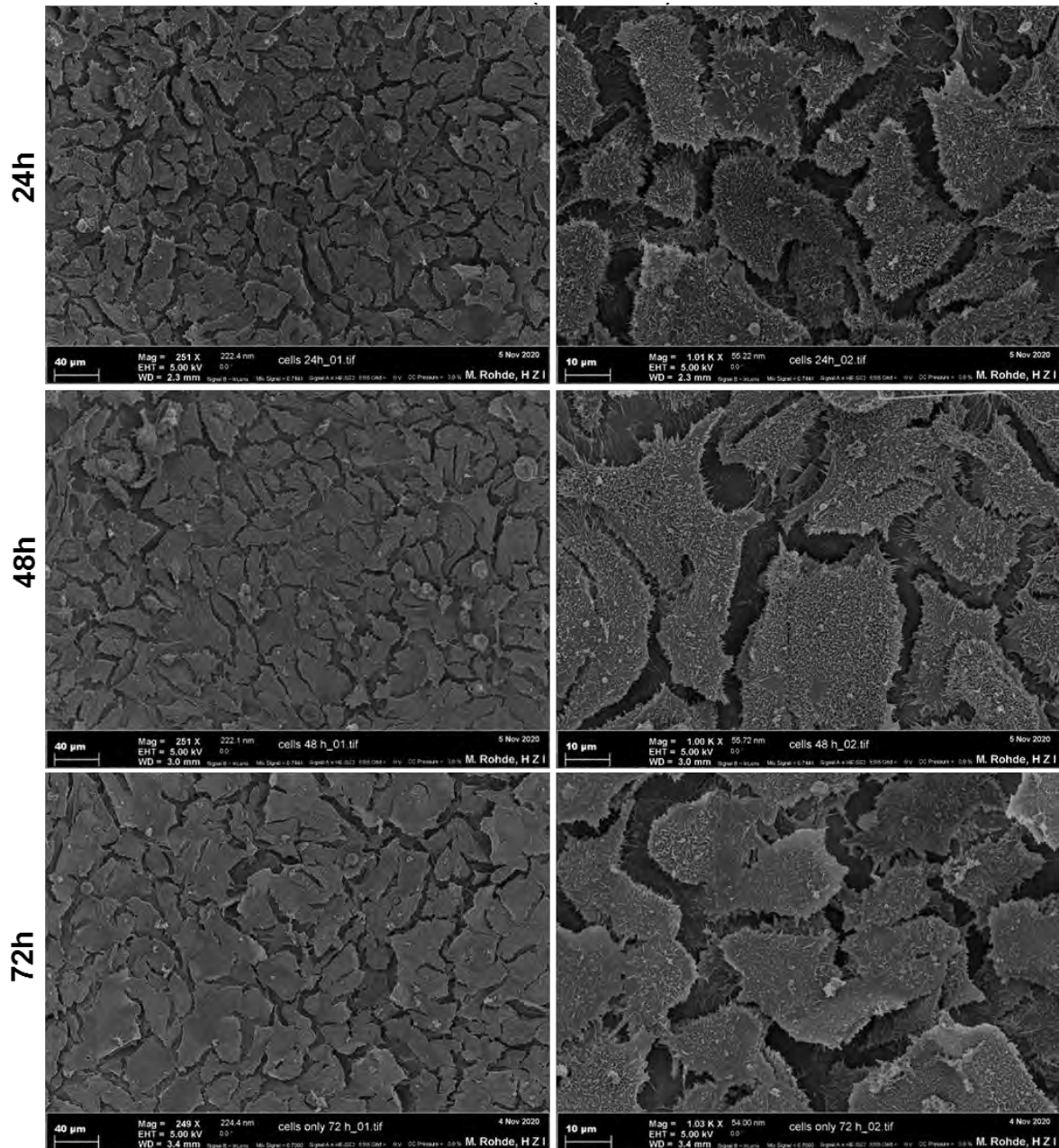
Supplementary Figure 8.9: Detroit-562 monolayer development during biofilm formation was monitored at both 24, 48 and 72h. Images were representative of 3 biological replicates. Each well was initially seeded with 2×10^5 cells/well.
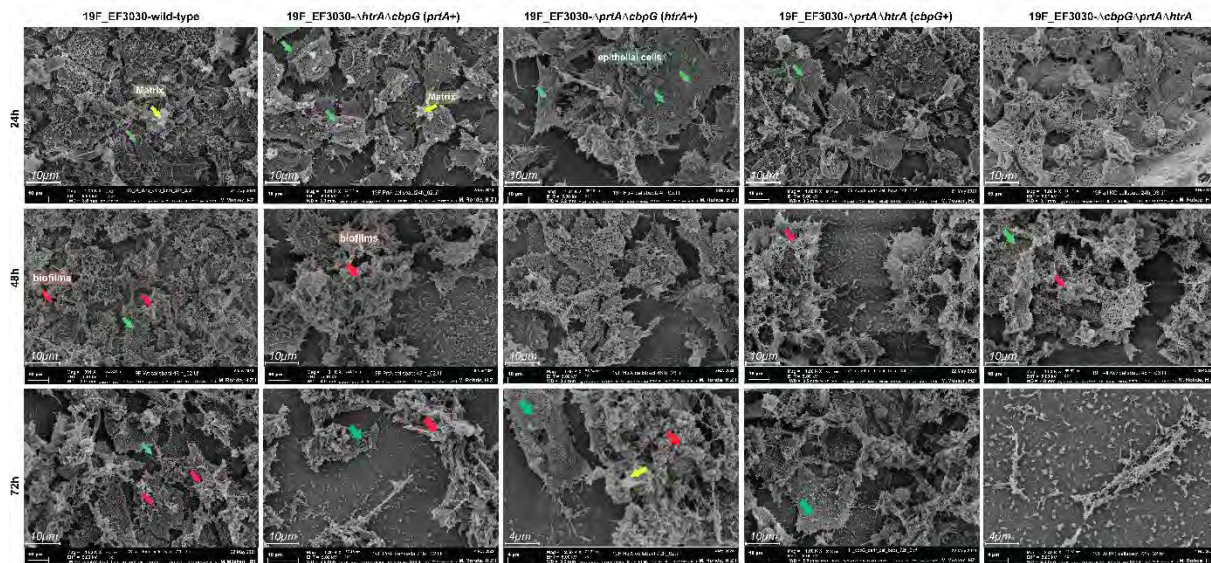
Supplementary Figure 8.10: Overview of the role of pneumococcal serine proteases on biofilms formation on living epithelial cells.

19F_EF3030 wild-type or mutant were seeded on living human epithelial cells Detroit-562. The low-magnification view demonstrated the architectural complexity of the biofilms formed after 24h. Biofilm development after 48 h and 72h were analyzed using SEM (bar =10 μm and 4 μm). The low-magnification view shows a high-density biofilms formation, matrix formation, and organization at 48h, and 72h (red arrow) and some detectable extracellular matrix formation covering the bacteria indicated by the yellow arrow. A green arrow indicates the underlying cells when they are visible. This figure shows a representative experiment of one infected culture.
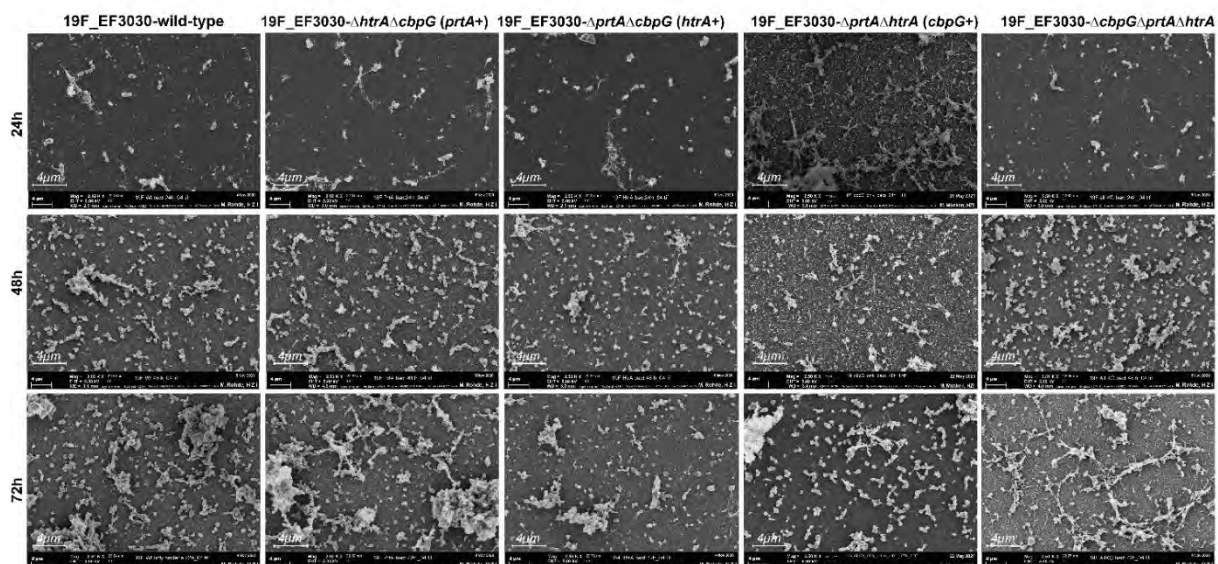
Supplementary Figure 8.11: Overview of the role of pneumococcal serine proteases on biofilms formation on abiotic glass surface.

19F_EF3030 wild-type or isogenic mutant strains were seeded on glass coverslips (diameter 12 mm) in wells of 24-well tissue culture plates and incubated for 24h. 48h and 72h. and prepared for SEM analysis (bar =20 µm). The low-magnification view shows the 19F_EF3030 wild-type and isogenic serine protease deletion strains attachment and matrix formation at 24 h. Higher biofilm organization could be detectable after 48h and 72h.

8.9 References

1. Pericone, C.D., Overweg, K., Hermans, P.W. & Weiser, J.N. Inhibitory and bactericidal effects of hydrogen peroxide production by Streptococcus pneumoniae on other inhabitants of the upper respiratory tract. *Infect Immun* **68**, 3990-7 (2000).
2. Regev-Yochay, G., Trzcinski, K., Thompson, C.M., Malley, R. & Lipsitch, M. Interference between Streptococcus pneumoniae and Staphylococcus aureus: In vitro hydrogen peroxide-mediated killing by Streptococcus pneumoniae. *J Bacteriol* **188**, 4996-5001 (2006).
3. Reid, S.D. et al. Streptococcus pneumoniae forms surface-attached communities in the middle ear of experimentally infected chinchillas. *J Infect Dis* **199**, 786-94 (2009).
4. Song, J.Y., Nahm, M.H. & Moseley, M.A. Clinical implications of pneumococcal serotypes: invasive disease potential, clinical presentations, and antibiotic resistance. *J Korean Med Sci* **28**, 4-15 (2013).
5. Geno, K.A. et al. Pneumococcal Capsules and Their Types: Past, Present, and Future. *Clin Microbiol Rev* **28**, 871-99 (2015).
6. Bentley, S.D. et al. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet* **2**, e31 (2006).
7. Hyams, C., Camberlein, E., Cohen, J.M., Bax, K. & Brown, J.S. The Streptococcus pneumoniae capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. *Infect Immun* **78**, 704-15 (2010).
8. Hathaway, L.J. et al. Capsule type of Streptococcus pneumoniae determines growth phenotype. *PLoS Pathog* **8**, e1002574 (2012).
9. Calix, J.J. et al. Biochemical, genetic, and serological characterization of two capsule subtypes among Streptococcus pneumoniae Serotype 20 strains: discovery of a new pneumococcal serotype. *J Biol Chem* **287**, 27885-94 (2012).
10. Calix, J.J. & Nahm, M.H. A new pneumococcal serotype, 11E, has a variably inactivated wcjE gene. *J Infect Dis* **202**, 29-38 (2010).
11. Bogaert, D., De Groot, R. & Hermans, P.W. Streptococcus pneumoniae colonisation: the key to pneumococcal disease. *Lancet Infect Dis* **4**, 144-54 (2004).
12. Marks, L.R., Parameswaran, G.I. & Hakansson, A.P. Pneumococcal interactions with epithelial cells are crucial for optimal biofilm formation and colonization in vitro and in vivo. *Infect Immun* **80**, 2744-60 (2012).
13. Weiser, J.N., Ferreira, D.M. & Paton, J.C. Streptococcus pneumoniae: transmission, colonization and invasion. *Nat Rev Microbiol* **16**, 355-367 (2018).
14. Bradshaw, J.L., Rafiqullah, I.M., Robinson, D.A. & McDaniel, L.S. Transformation of nonencapsulated Streptococcus pneumoniae during systemic infection. *Scientific Reports* **10**, 18932 (2020).
15. Bauer, T.T. et al. Cost analyses of community-acquired pneumonia from the hospital perspective. *Chest* **128**, 2238-46 (2005).
16. File, T.M. Community-acquired pneumonia. *Lancet* **362**, 1991-2001 (2003).
17. Yahiaoui, R.Y. et al. Prevalence and antibiotic resistance of commensal Streptococcus pneumoniae in nine European countries. *Future Microbiol* **11**, 737-44 (2016).
18. Mills, R.O. et al. Post-Vaccination Streptococcus pneumoniae Carriage and Virulence Gene Distribution among Children Less Than Five Years of Age, Cape Coast, Ghana. *Microorganisms* **8** (2020).
19. O'Brien, K.L. et al. Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. *Lancet* **374**, 893-902 (2009).
20. Wijayasri, S. et al. The shifting epidemiology and serotype distribution of invasive pneumococcal disease in Ontario, Canada, 2007-2017. *PLoS One* **14**, e0226353 (2019).
21. Prevention, C.f.D.C.a. (2017).
22. Muley, V.A., Ghadage, D.P., Yadav, G.E. & Bhore, A.V. Study of Invasive Pneumococcal Infection in Adults with Reference to Penicillin Resistance. *J Lab Physicians* **9**, 31-35 (2017).
23. Organization, W.H. (2017).
24. Tacconelli, E. et al. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis* **18**, 318-327 (2018).
25. WHO. 9–12 (Geneva, February, Geneva, February, 1999).
26. WHO. 110–19 (Wkly Epidemiol Rec, 2003).
27. Bergenfelz, C. & Hakansson, A.P. Streptococcus pneumoniae Otitis Media Pathogenesis and How It Informs Our Understanding of Vaccine Strategies. *Curr Otorhinolaryngol Rep* **5**, 115-124 (2017).
28. Wahl, B. et al. Burden of Streptococcus pneumoniae and Haemophilus influenzae type b disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000-15. *Lancet Glob Health* **6**, e744-e757 (2018).
29. UNICEF. (United Nations Children's Emergency Fund: New York, NY, USA, 2006).
30. Organization, W.H. 93–104 (Wkly.Epidemiol, 2007).
31. Althouse, B.M. et al. Identifying transmission routes of Streptococcus pneumoniae and sources of acquisitions in high transmission communities. *Epidemiol Infect* **145**, 2750-2758 (2017).
32. Adegbola, R.A. et al. Carriage of Streptococcus pneumoniae and other respiratory bacterial pathogens in low and lower-middle income countries: a systematic review and meta-analysis. *PLoS One* **9**, e103293 (2014).
33. Wang, L., Fu, J., Liang, Z. & Chen, J. Prevalence and serotype distribution of nasopharyngeal carriage of Streptococcus pneumoniae in China: a meta-analysis. *BMC Infect Dis* **17**, 765 (2017).
34. Navne, J.E. et al. Nasopharyngeal bacterial carriage in young children in Greenland: a population at high risk of respiratory infections. *Epidemiol Infect* **144**, 3226-3236 (2016).
35. van Hoek, A.J. et al. Pneumococcal carriage in children and adults two years after introduction of the thirteen valent pneumococcal conjugate vaccine in England. *Vaccine* **32**, 4349-55 (2014).

36. Milucky, J. et al. Streptococcus pneumoniae colonization after introduction of 13-valent pneumococcal conjugate vaccine for US adults 65 years of age and older, 2015-2016. *Vaccine* **37**, 1094-1100 (2019).

37. Bosch, A. et al. Nasopharyngeal carriage of Streptococcus pneumoniae and other bacteria in the 7th year after implementation of the pneumococcal conjugate vaccine in the Netherlands. *Vaccine* **34**, 531-539 (2016).

38. Blasi, F., Mantero, M., Santus, P. & Tarsia, P. Understanding the burden of pneumococcal disease in adults. *Clin Microbiol Infect* **18 Suppl 5**, 7-14 (2012).

39. Zhang, D., Petigara, T. & Yang, X. Clinical and economic burden of pneumococcal disease in US adults aged 19-64 years with chronic or immunocompromising diseases: an observational database study. *BMC Infect Dis* **18**, 436 (2018).

40. Organization, W.H. Position paper on Pneumococcal conjugate vaccines in infants and children under 5 years of age. *Weekly Epidemiol Rec*, 1– 12. (2019).

41. Austrian, R. & Gold, J. PNEUMOCOCCAL BACTEREMIA WITH ESPECIAL REFERENCE TO BACTEREMIC PNEUMOCOCCAL PNEUMONIA. *Ann Intern Med* **60**, 759-76 (1964).

42. Regev-Yochay, G. et al. Nasopharyngeal carriage of Streptococcus pneumoniae by adults and children in community and family settings. *Clin Infect Dis* **38**, 632-9 (2004).

43. Mwenya, D.M. et al. Impact of cotrimoxazole on carriage and antibiotic resistance of Streptococcus pneumoniae and Haemophilus influenzae in HIV-infected children in Zambia. *Antimicrob Agents Chemother* **54**, 3756-62 (2010).

44. Buerano, C.C. et al. Isolation of acanthamoeba genotype t4 from a non-contact lens wearer from the Philippines. *Trop Med Health* **42**, 145-7 (2014).

45. Whitney, C.G. et al. Increasing prevalence of multidrug-resistant Streptococcus pneumoniae in the United States. *N Engl J Med* **343**, 1917-24 (2000).

46. World Health, O. (World Health Organization, 2014).

47. (ECDC), E.C.f.D.P.a.C. (Stockholm: ECDC, 2017).

48. Global, regional, and national under-5 mortality, adult mortality, age-specific mortality, and life expectancy, 1970-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* **390**, 1084-1150 (2017).

49. Hammitt, L.L. et al. Indirect effect of conjugate vaccine on adult carriage of Streptococcus pneumoniae: an explanation of trends in invasive pneumococcal disease. *J Infect Dis* **193**, 1487-94 (2006).

50. Richter, S.S. et al. Pneumococcal serotypes before and after introduction of conjugate vaccines, United States, 1999-2011(1.). *Emerg Infect Dis* **19**, 1074-83 (2013).

51. Izurieta, P., Bahety, P., Adegbola, R., Clarke, C. & Hoet, B. Public health impact of pneumococcal conjugate vaccine infant immunization programs: assessment of invasive pneumococcal disease burden and serotype distribution. *Expert Rev Vaccines* **17**, 479-493 (2018).

52. van der Linden, M. et al. Four years of universal pneumococcal conjugate infant vaccination in Germany: impact on incidence of invasive pneumococcal disease and serotype distribution in children. *Vaccine* **30**, 5880-5 (2012).

53. Deloria Knoll, M. & Bennett, J.C. Global Landscape Review of Serotype-Specific Invasive Pneumococcal Disease Surveillance among Countries Using PCV10/13: The Pneumococcal Serotype Replacement and Distribution Estimation (PSERENADE) Project. **9** (2021).

54. Tettelin, H. et al. Complete genome sequence of a virulent isolate of Streptococcus pneumoniae. *Science* **293**, 498-506 (2001).

55. Hoskins, J. et al. Genome of the bacterium Streptococcus pneumoniae strain R6. *J Bacteriol* **183**, 5709-17 (2001).

56. Johnson, H.L. et al. Systematic evaluation of serotypes causing invasive pneumococcal disease among children under five: the pneumococcal global serotype project. *PLoS Med* **7** (2010).

57. Griffith, F. The Significance of Pneumococcal Types. *J Hyg (Lond)* **27**, 113-59 (1928).

58. Avery, O.T., Macleod, C.M. & McCarty, M. STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III. *J Exp Med* **79**, 137-58 (1944).

59. Blomberg, C. et al. Pattern of accessory regions and invasive disease potential in Streptococcus pneumoniae. *J Infect Dis* **199**, 1032-42 (2009).

60. Mejia, A., G.; Hammerschmidt, S.; Veening,. in University of Greifswald (University of Greifswald,, 2018).

61. Obert, C. et al. Identification of a Candidate Streptococcus pneumoniae core genome and regions of diversity correlated with invasive pneumococcal disease. *Infect Immun* **74**, 4766-77 (2006).

62. Hiller, N.L. et al. Generation of genic diversity among Streptococcus pneumoniae strains via horizontal gene transfer during a chronic polyclonal pediatric infection. *PLoS Pathog* **6**, e1001108 (2010).

63. Andam, C.P. & Hanage, W.P. Mechanisms of genome evolution of Streptococcus. *Infect Genet Evol* **33**, 334-42 (2015).

64. Schulz, C. & Hammerschmidt, S. Exploitation of physiology and metabolomics to identify pneumococcal vaccine candidates. *Expert Rev Vaccines* **12**, 1061-75 (2013).

65. Kadioglu, A., Weiser, J.N., Paton, J.C. & Andrew, P.W. The role of Streptococcus pneumoniae virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol* **6**, 288-301 (2008).

66. Heidelberger, M. & Avery, O.T. THE SOLUBLE SPECIFIC SUBSTANCE OF PNEUMOCOCCUS : SECOND PAPER. *J Exp Med* **40**, 301-17 (1924).

67. Skov Sørensen, U.B., Blom, J., Birch-Andersen, A. & Henrichsen, J. Ultrastructural localization of capsules, cell wall polysaccharide, cell wall proteins, and F antigen in pneumococci. *Infect Immun* **56**, 1890-6 (1988).

68. Brooks, L.R.K. & Mias, G.I. Streptococcus pneumoniae's Virulence and Host Immunity: Aging, Diagnostics, and Prevention. *Front Immunol* **9**, 1366 (2018).

69. Abeyta, M., Hardy, G.G. & Yother, J. Genetic alteration of capsule type but not PspA type affects accessibility of surface-bound complement and surface antigens of Streptococcus pneumoniae. *Infect Immun* **71**, 218-25 (2003).

70. Nelson, A.L. et al. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect Immun* **75**, 83-90 (2007).
71. García, E. & López, R. Molecular biology of the capsular genes of Streptococcus pneumoniae. *FEMS Microbiol Lett* **149**, 1-10 (1997).
72. Geno, K.A., Saad, J.S. & Nahm, M.H. Discovery of Novel Pneumococcal Serotype 35D, a Natural WciG-Deficient Variant of Serotype 35B. *J Clin Microbiol* **55**, 1416-1425 (2017).
73. Wyres, K.L. et al. Pneumococcal capsular switching: a historical perspective. *J Infect Dis* **207**, 439-49 (2013).
74. Morona, J.K., Paton, J.C., Miller, D.C. & Morona, R. Tyrosine phosphorylation of CpsD negatively regulates capsular polysaccharide biosynthesis in streptococcus pneumoniae. *Mol Microbiol* **35**, 1431-42 (2000).
75. Manso, A.S. et al. A random six-phase switch regulates pneumococcal virulence via global epigenetic changes. *Nat Commun* **5**, 5055 (2014).
76. Kim, J.O. & Weiser, J.N. Association of intrastrain phase variation in quantity of capsular polysaccharide and teichoic acid with the virulence of Streptococcus pneumoniae. *J Infect Dis* **177**, 368-77 (1998).
77. Hammerschmidt, S. et al. Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells. *Infect Immun* **73**, 4653-67 (2005).
78. Hall-Stoodley, L. et al. Characterization of biofilm matrix, degradation by DNase treatment and evidence of capsule downregulation in Streptococcus pneumoniae clinical isolates. *BMC Microbiol* **8**, 173 (2008).
79. Sanchez, C.J. et al. Streptococcus pneumoniae in biofilms are unable to cause invasive disease due to altered virulence determinant production. *PLoS One* **6**, e28738 (2011).
80. Vollmer, W., Massidda, O. & Tomasz, A. The Cell Wall of Streptococcus pneumoniae. *Microbiol Spectr* **7** (2019).
81. Jennings, H.J., Lugowski, C. & Young, N.M. Structure of the complex polysaccharide C-substance from Streptococcus pneumoniae type 1. *Biochemistry* **19**, 4712-9 (1980).
82. Heß, N. & Waldow, F. Lipoteichoic acid deficiency permits normal growth but impairs virulence of Streptococcus pneumoniae. **8**, 2093 (2017).
83. Navarre, W.W. & Schneewind, O. Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev* **63**, 174-229 (1999).
84. Bui, H.H., Leohr, J.K. & Kuo, M.S. Analysis of sphingolipids in extracted human plasma using liquid chromatography electrospray ionization tandem mass spectrometry. *Anal Biochem* **423**, 187-94 (2012).
85. Pérez-Dorado, I., Galan-Bartual, S. & Hermoso, J.A. Pneumococcal surface proteins: when the whole is greater than the sum of its parts. *Mol Oral Microbiol* **27**, 221-45 (2012).
86. Pérez-Dorado, I. et al. Insights into pneumococcal fratricide from the crystal structures of the modular killing factor LytC. *Nat Struct Mol Biol* **17**, 576-81 (2010).
87. Hammerschmidt, S. Adherence molecules of pathogenic pneumococci. *Curr Opin Microbiol* **9**, 12-20 (2006).
88. Bergmann, S. & Hammerschmidt, S. Versatility of pneumococcal surface proteins. *Microbiology (Reading)* **152**, 295-303 (2006).
89. Voss, S., Gamez, G. & Hammerschmidt, S. Impact of pneumococcal microbial surface components recognizing adhesive matrix molecules on colonization. *Mol Oral Microbiol* **27**, 246-56 (2012).
90. Fischer, W. Phosphocholine of pneumococcal teichoic acids: role in bacterial physiology and pneumococcal infection. *Res Microbiol* **151**, 421-7 (2000).
91. Ware, D., Watt, J. & Swiatlo, E. Utilization of putrescine by Streptococcus pneumoniae during growth in choline-limited medium. *J Microbiol* **43**, 398-405 (2005).
92. Maestro, B. & Sanz, J.M. Choline Binding Proteins from Streptococcus pneumoniae: A Dual Role as Enzybiotics and Targets for the Design of New Antimicrobials. *Antibiotics (Basel)* **5** (2016).
93. García, E. et al. Molecular evolution of lytic enzymes of Streptococcus pneumoniae and its bacteriophages. *Proc Natl Acad Sci U S A* **85**, 914-8 (1988).
94. Sanz, J.M., Díaz, E. & García, J.L. Studies on the structure and function of the N-terminal domain of the pneumococcal murein hydrolases. *Mol Microbiol* **6**, 921-31 (1992).
95. Albrich, W.C., Monnet, D.L. & Harbarth, S. Antibiotic selection pressure and resistance in Streptococcus pneumoniae and Streptococcus pyogenes. *Emerg Infect Dis* **10**, 514-7 (2004).
96. Hilleringmann, M., Kohler, S., Gàmez, G., and Hammerschmidt, S. Pneumococcal pili and adhesins. In: Streptococcus pneumoniae: Molecular mechanisms of host-pathogen interactions (ed. Jeremy Brown, S.H., Carlos Orihuela) (Academic Press is an imprint of Elsevier, Amsterdam, NL. , 2015).
97. Galán-Bartual, S., Pérez-Dorado, I., García, P. & Hermoso, J. . In Streptococcus pneumoniae, molecular mechanisms of host-pathogen interactions (Elsevier, 2015).
98. Gutiérrez-Fernández, J. et al. Modular Architecture and Unique Teichoic Acid Recognition Features of Choline-Binding Protein L (CbpL) Contributing to Pneumococcal Pathogenesis. *Sci Rep* **6**, 38094 (2016).
99. Hakenbeck, R., Madhour, A., Denapaite, D. & Brückner, R. Versatility of choline metabolism and choline-binding proteins in Streptococcus pneumoniae and commensal streptococci. *FEMS Microbiol Rev* **33**, 572-86 (2009).
100. Gosink, K.K., Mann, E.R., Guglielmo, C., Tuomanen, E.I. & Masure, H.R. Role of novel choline binding proteins in virulence of Streptococcus pneumoniae. *Infect Immun* **68**, 5690-5 (2000).
101. Swiatlo, E., Champlin, F.R., Holman, S.C., Wilson, W.W. & Watt, J.M. Contribution of choline-binding proteins to cell surface properties of Streptococcus pneumoniae. *Infect Immun* **70**, 412-5 (2002).
102. Hammerschmidt, S. (ed. Bioscience, H.) 141-203 (Norfolk, U.K, 2007).
103. Pribyl, T. et al. Influence of impaired lipoprotein biogenesis on surface and exoproteome of Streptococcus pneumoniae. *J Proteome Res* **13**, 650-67 (2014).
104. Molina, R. et al. Crystal structure of CbpF, a bifunctional choline-binding protein and autolysis regulator from Streptococcus pneumoniae. *EMBO Rep* **10**, 246-51 (2009).

105. McDaniel, L.S., Ralph, B.A., McDaniel, D.O. & Briles, D.E. Localization of protection-eliciting epitopes on PspA of Streptococcus pneumoniae between amino acid residues 192 and 260. *Microb Pathog* **17**, 323-37 (1994).

106. Mann, B. et al. Multifunctional role of choline binding protein G in pneumococcal pathogenesis. *Infect Immun* **74**, 821-9 (2006).

107. Kovacs-Simon, A., Titball, R.W. & Michell, S.L. Lipoproteins of bacterial pathogens. *Infect Immun* **79**, 548-61 (2011).

108. Kohler, S., Voß, F., Gómez Mejia, A., Brown, J.S. & Hammerschmidt, S. Pneumococcal lipoproteins involved in bacterial fitness, virulence, and immune evasion. *FEBS Lett* **590**, 3820-3839 (2016).

109. Zückert, W.R. Secretion of bacterial lipoproteins: through the cytoplasmic membrane, the periplasm and beyond. *Biochim Biophys Acta* **1843**, 1509-16 (2014).

110. Chimalapati, S. et al. Effects of deletion of the Streptococcus pneumoniae lipoprotein diacylglyceryl transferase gene lgt on ABC transporter function and on growth in vivo. *PLoS One* **7**, e41393 (2012).

111. Petit, C.M., Brown, J.R., Ingraham, K., Bryant, A.P. & Holmes, D.J. Lipid modification of prelipoproteins is dispensable for growth in vitro but essential for virulence in Streptococcus pneumoniae. *FEMS Microbiol Lett* **200**, 229-33 (2001).

112. Dramsi, S., Magnet, S., Davison, S. & Arthur, M. Covalent attachment of proteins to peptidoglycan. *FEMS Microbiol Rev* **32**, 307-20 (2008).

113. Mitchell, A.M. & Mitchell, T.J. Streptococcus pneumoniae: virulence factors and variation. *Clin Microbiol Infect* **16**, 411-8 (2010).

114. Pluvinage, B. et al. Inhibition of the pneumococcal virulence factor StrH and molecular insights into N-glycan recognition and hydrolysis. *Structure* **19**, 1603-14 (2011).

115. Willis, L.M., Zhang, R., Reid, A., Withers, S.G. & Wakarchuk, W.W. Mechanistic investigation of the endo-alpha-N-acetylgalactosaminidase from Streptococcus pneumoniae R6. *Biochemistry* **48**, 10334-41 (2009).

116. Abbott, D.W., Macauley, M.S., Vocadlo, D.J. & Boraston, A.B. Streptococcus pneumoniae endohexosaminidase D, structural and mechanistic insight into substrate-assisted catalysis in family 85 glycoside hydrolases. *J Biol Chem* **284**, 11676-89 (2009).

117. Yamaguchi, M., Terao, Y., Mori, Y., Hamada, S. & Kawabata, S. PfbA, a novel plasmin- and fibronectin-binding protein of Streptococcus pneumoniae, contributes to fibronectin-dependent adhesion and antiphagocytosis. *J Biol Chem* **283**, 36272-9 (2008).

118. Jensch, I. et al. PavB is a surface-exposed adhesin of Streptococcus pneumoniae contributing to nasopharyngeal colonization and airways infections. *Mol Microbiol* **77**, 22-43 (2010).

119. van Bueren, A.L., Higgins, M., Wang, D., Burke, R.D. & Boraston, A.B. Identification and structural basis of binding to host lung glycogen by streptococcal virulence factors. *Nat Struct Mol Biol* **14**, 76-84 (2007).

120. Garbe, J. & Collin, M. Bacterial hydrolysis of host glycoproteins - powerful protein modification and efficient nutrient acquisition. *J Innate Immun* **4**, 121-31 (2012).

121. Hammerschmidt, S. Surface-exposed adherence molecules of Streptococcus pneumoniae. *Methods Mol Biol* **470**, 29-45 (2009).

122. Kharat, A.S. & Tomasz, A. Inactivation of the srtA gene affects localization of surface proteins and decreases adhesion of Streptococcus pneumoniae to human pharyngeal cells in vitro. *Infect Immun* **71**, 2758-65 (2003).

123. Park, I.H. et al. Nontypeable pneumococci can be divided into multiple cps types, including one type expressing the novel gene pspK. *mBio* **3** (2012).

124. Daniely, D. et al. Pneumococcal 6-phosphogluconate-dehydrogenase, a putative adhesin, induces protective immune response in mice. *Clin Exp Immunol* **144**, 254-63 (2006).

125. Noske, N., Kämmerer, U., Rohde, M. & Hammerschmidt, S. Pneumococcal interaction with human dendritic cells: phagocytosis, survival, and induced adaptive immune response are manipulated by PavA. *J Immunol* **183**, 1952-63 (2009).

126. Chhatwal, G.S. Anchorless adhesins and invasins of Gram-positive bacteria: a new class of virulence factors. *Trends Microbiol* **10**, 205-8 (2002).

127. Bergmann, S., Rohde, M. & Hammerschmidt, S. Glyceraldehyde-3-phosphate dehydrogenase of Streptococcus pneumoniae is a surface-displayed plasminogen-binding protein. *Infect Immun* **72**, 2416-9 (2004).

128. Ljungh, A., Moran, A.P. & Wadström, T. Interactions of bacterial adhesins with extracellular matrix and plasma proteins: pathogenic implications and therapeutic possibilities. *FEMS Immunol Med Microbiol* **16**, 117-26 (1996).

129. Weiser, J.N. The pneumococcus: why a commensal misbehaves. *J Mol Med (Berl)* **88**, 97-102 (2010).

130. Faden, H. et al. Relationship between nasopharyngeal colonization and the development of otitis media in children. Tonawanda/Williamsville Pediatrics. *J Infect Dis* **175**, 1440-5 (1997).

131. Faden, H., Duffy, L., Williams, A., Krystofik, D.A. & Wolf, J. Epidemiology of nasopharyngeal colonization with nontypeable Haemophilus influenzae in the first 2 years of life. *J Infect Dis* **172**, 132-5 (1995).

132. Kyaw, M.H., Christie, P., Jones, I.G. & Campbell, H. The changing epidemiology of bacterial meningitis and invasive non-meningitic bacterial disease in scotland during the period 1983-99. *Scand J Infect Dis* **34**, 289-98 (2002).

133. Nouwen, J.L., van Belkum, A. & Verbrugh, H.A. Determinants of Staphylococcus aureus nasal carriage. *Neth J Med* **59**, 126-33 (2001).

134. Davis, K.M., Akinbi, H.T., Standish, A.J. & Weiser, J.N. Resistance to mucosal lysozyme compensates for the fitness deficit of peptidoglycan modifications by Streptococcus pneumoniae. *PLoS Pathog* **4**, e1000241 (2008).

135. Hauck, C.R. Cell adhesion receptors - signaling capacity and exploitation by bacterial pathogens. *Med Microbiol Immunol* **191**, 55-62 (2002).

136. van der Windt, D. et al. Nonencapsulated Streptococcus pneumoniae resists extracellular human neutrophil elastase- and cathepsin G-mediated killing. *FEMS Immunol Med Microbiol* **66**, 445-8 (2012).

137. Elm, C. et al. Ectodomains 3 and 4 of human polymeric Immunoglobulin receptor (hpIgR) mediate invasion of Streptococcus pneumoniae into the epithelium. *J Biol Chem* **279**, 6296-304 (2004).

138. McCullers, J.A. & Rehg, J.E. Lethal synergism between influenza virus and Streptococcus pneumoniae: characterization of a mouse model and the role of platelet-activating factor receptor. *J Infect Dis* **186**, 341-50 (2002).

139. Holmes, A.R. et al. The pavA gene of Streptococcus pneumoniae encodes a fibronectin-binding protein that is essential for virulence. *Mol Microbiol* **41**, 1395-408 (2001).

140. Bergmann, S. et al. Integrin-linked kinase is required for vitronectin-mediated internalization of Streptococcus pneumoniae by host cells. *J Cell Sci* **122**, 256-67 (2009).

141. Binsker, U. et al. Pneumococcal Adhesins PavB and PspC Are Important for the Interplay with Human Thrombospondin-1. *J Biol Chem* **290**, 14542-55 (2015).

142. Orihuela, C.J. et al. Laminin receptor initiates bacterial contact with the blood brain barrier in experimental meningitis models. *J Clin Invest* **119**, 1638-46 (2009).

143. Kanwal, S. et al. Mapping the recognition domains of pneumococcal fibronectin-binding proteins PavA and PavB demonstrates a common pattern of molecular interactions with fibronectin type III repeats. **105**, 839-859 (2017).

144. Pracht, D. et al. PavA of Streptococcus pneumoniae modulates adherence, invasion, and meningeal inflammation. *Infect Immun* **73**, 2680-9 (2005).

145. Cundell, D.R., Weiser, J.N., Shen, J., Young, A. & Tuomanen, E.I. Relationship between colonial morphology and adherence of Streptococcus pneumoniae. *Infect Immun* **63**, 757-61 (1995).

146. Balachandran, P., Brooks-Walter, A., Virolainen-Julkunen, A., Hollingshead, S.K. & Briles, D.E. Role of pneumococcal surface protein C in nasopharyngeal carriage and pneumonia and its ability to elicit protection against carriage of Streptococcus pneumoniae. *Infect Immun* **70**, 2526-34 (2002).

147. Gessner, B.D., Mueller, J.E. & Yaro, S. African meningitis belt pneumococcal disease epidemiology indicates a need for an effective serotype 1 containing vaccine, including for older children and adults. *BMC Infect Dis* **10**, 22 (2010).

148. Clarke, T.B., Francella, N., Huegel, A. & Weiser, J.N. Invasive bacterial pathogens exploit TLR-mediated downregulation of tight junction components to facilitate translocation across the epithelium. *Cell Host Microbe* **9**, 404-14 (2011).

149. Peter, A. et al. Localization and pneumococcal alteration of junction proteins in the human alveolar-capillary compartment. **147**, 707-719 (2017).

150. Bergmann, S., Schoenen, H. & Hammerschmidt, S. The interaction between bacterial enolase and plasminogen promotes adherence of Streptococcus pneumoniae to epithelial and endothelial cells. *Int J Med Microbiol* **303**, 452-62 (2013).

151. Bergmann, S., Rohde, M., Preissner, K.T. & Hammerschmidt, S. The nine residue plasminogen-binding motif of the pneumococcal enolase is the major cofactor of plasmin-mediated degradation of extracellular matrix, dissolution of fibrin and transmigration. *Thromb Haemost* **94**, 304-11 (2005).

152. Weiser, J.N. et al. Antibody-enhanced pneumococcal adherence requires IgA1 protease. *Proc Natl Acad Sci U S A* **100**, 4215-20 (2003).

153. Ring, A., Weiser, J.N. & Tuomanen, E.I. Pneumococcal trafficking across the blood-brain barrier. Molecular analysis of a novel bidirectional pathway. *J Clin Invest* **102**, 347-60 (1998).

154. Wang, L. et al. Streptococcus pneumoniae aminopeptidase N contributes to bacterial virulence and elicits a strong innate immune response through MAPK and PI3K/AKT signaling. *J Microbiol* **58**, 330-339 (2020).

155. Sebert, M.E., Palmer, L.M., Rosenberg, M. & Weiser, J.N. Microarray-based identification of htrA, a Streptococcus pneumoniae gene that is regulated by the CiaRH two-component system and contributes to nasopharyngeal colonization. *Infect Immun* **70**, 4059-67 (2002).

156. Muhammad, M.H. et al. Beyond Risk: Bacterial Biofilms and Their Regulating Approaches. *Front Microbiol* **11**, 928 (2020).

157. Flemming, H.C. & Wingender, J. Relevance of microbial extracellular polymeric substances (EPSs)--Part I: Structural and ecological aspects. *Water Sci Technol* **43**, 1-8 (2001).

158. Flemming, H.C. & Wuertz, S. Bacteria and archaea on Earth and their abundance in biofilms. *Nat Rev Microbiol* **17**, 247-260 (2019).

159. Lewis, K. Multidrug tolerance of biofilms and persister cells. *Curr Top Microbiol Immunol* **322**, 107-31 (2008).

160. Rutherford, S.T. & Bassler, B.L. Bacterial quorum sensing: its role in virulence and possibilities for its control. *Cold Spring Harb Perspect Med* **2** (2012).

161. Dongari-Bagtzoglou, A. Pathogenesis of mucosal biofilm infections: challenges and progress. *Expert Rev Anti Infect Ther* **6**, 201-8 (2008).

162. Poddighe, D. & Vangelista, L. Staphylococcus aureus Infection and Persistence in Chronic Rhinosinusitis: Focus on Leukocidin ED. **12** (2020).

163. López, D., Vlamakis, H. & Kolter, R. Biofilms. *Cold Spring Harb Perspect Biol* **2**, a000398 (2010).

164. Yadav, P. et al. Deciphering Streptococcal Biofilms. *Microorganisms* **8** (2020).

165. Donlan, R.M. Biofilms: microbial life on surfaces. *Emerg Infect Dis* **8**, 881-90 (2002).

166. Berlanga, M. & Guerrero, R. Living together in biofilms: the microbial cell factory and its biotechnological implications. *Microb Cell Fact* **15**, 165 (2016).

167. Stewart, P.S. & Franklin, M.J. Physiological heterogeneity in biofilms. *Nat Rev Microbiol* **6**, 199-210 (2008).

168. Marks, L.R., Reddinger, R.M. & Hakansson, A.P. Biofilm formation enhances fomite survival of Streptococcus pneumoniae and Streptococcus pyogenes. *Infect Immun* **82**, 1141-6 (2014).

169. Walsh, R.L. & Camilli, A. Streptococcus pneumoniae is desiccation tolerant and infectious upon rehydration. *mBio* **2**, e00092-11 (2011).

170. Stoodley, P., Sauer, K., Davies, D.G. & Costerton, J.W. Biofilms as complex differentiated communities. *Annu Rev Microbiol* **56**, 187-209 (2002).

171. Donlan, R.M. & Costerton, J.W. Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev* **15**, 167-93 (2002).

172. Chole, R.A. & Faddis, B.T. Anatomical evidence of microbial biofilms in tonsillar tissues: a possible mechanism to explain chronicity. *Arch Otolaryngol Head Neck Surg* **129**, 634-6 (2003).

173. Hoa, M., Syamal, M., Sachdeva, L., Berk, R. & Coticchia, J. Demonstration of nasopharyngeal and middle ear mucosal biofilms in an animal model of acute otitis media. *Ann Otol Rhinol Laryngol* **118**, 292-8 (2009).

174. Hall-Stoodley, L. & Stoodley, P. Evolving concepts in biofilm infections. *Cell Microbiol* **11**, 1034-43 (2009).

175. Hall-Stoodley, L. et al. Direct detection of bacterial biofilms on the middle-ear mucosa of children with chronic otitis media. *Jama* **296**, 202-11 (2006).

176. Sanderson, A.R., Leid, J.G. & Hunsaker, D. Bacterial biofilms on the sinus mucosa of human subjects with chronic rhinosinusitis. *Laryngoscope* **116**, 1121-6 (2006).

177. Oggioni, M.R. et al. Switch from planktonic to sessile life: a major event in pneumococcal pathogenesis. *Mol Microbiol* **61**, 1196-210 (2006).

178. Trappetti, C. et al. The impact of the competence quorum sensing system on Streptococcus pneumoniae biofilms varies depending on the experimental model. *BMC Microbiol* **11**, 75 (2011).

179. Muñoz-Elías, E.J., Marcano, J. & Camilli, A. Isolation of Streptococcus pneumoniae biofilm mutants and their characterization during nasopharyngeal colonization. *Infect Immun* **76**, 5049-61 (2008).

180. Trappetti, C. et al. Sialic acid: a preventable signal for pneumococcal biofilm formation, colonization, and invasion of the host. *J Infect Dis* **199**, 1497-505 (2009).

181. Gieseke, A., Tarre, S., Green, M. & de Beer, D. Nitrification in a biofilm at low pH values: role of in situ microenvironments and acid tolerance. *Appl Environ Microbiol* **72**, 4283-92 (2006).

182. Weimer, K.E. et al. Coinfection with Haemophilus influenzae promotes pneumococcal biofilm formation during experimental otitis media and impedes the progression of pneumococcal disease. *J Infect Dis* **202**, 1068-75 (2010).

183. Chao, Y., Marks, L.R., Pettigrew, M.M. & Hakansson, A.P. Streptococcus pneumoniae biofilm formation and dispersion during colonization and disease. *Front Cell Infect Microbiol* **4**, 194 (2014).

184. Di Martino, P. Extracellular polymeric substances, a key element in understanding biofilm phenotype. *AIMS Microbiol* **4**, 274-288 (2018).

185. Lefebvre, E., Vighetto, C., Di Martino, P., Larreta Garde, V. & Seyer, D. Synergistic antibiofilm efficacy of various commercial antiseptics, enzymes and EDTA: a study of Pseudomonas aeruginosa and Staphylococcus aureus biofilms. *Int J Antimicrob Agents* **48**, 181-8 (2016).

186. Domenech, M., García, E. & Moscoso, M. Biofilm formation in Streptococcus pneumoniae. *Microb Biotechnol* **5**, 455-65 (2012).

187. Flemming, H.C. & Wingender, J. The biofilm matrix. *Nat Rev Microbiol* **8**, 623-33 (2010).

188. Johnsborg, O. & Håvarstein, L.S. Regulation of natural genetic transformation and acquisition of transforming DNA in Streptococcus pneumoniae. *FEMS Microbiol Rev* **33**, 627-42 (2009).

189. Domenech, M., García, E. & Moscoso, M. Versatility of the capsular genes during biofilm formation by Streptococcus pneumoniae. *Environ Microbiol* **11**, 2542-55 (2009).

190. Moscoso, M., García, E. & López, R. Biofilm formation by Streptococcus pneumoniae: role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion. *J Bacteriol* **188**, 7785-95 (2006).

191. Marks, L.R., Reddinger, R.M. & Hakansson, A.P. High levels of genetic recombination during nasopharyngeal carriage and biofilm formation in Streptococcus pneumoniae. *mBio* **3** (2012).

192. Allegrucci, M. et al. Phenotypic characterization of Streptococcus pneumoniae biofilm development. *J Bacteriol* **188**, 2325-35 (2006).

193. Chua, S.L. et al. Dispersed cells represent a distinct stage in the transition from bacterial biofilm to planktonic lifestyles. *Nat Commun* **5**, 4462 (2014).

194. Chai, M.H. et al. Proteomic comparisons of opaque and transparent variants of Streptococcus pneumoniae by two dimensional-differential gel electrophoresis. *Sci Rep* **7**, 2453 (2017).

195. Bittaye, M. & Cash, P. Proteomic variation and diversity in clinical Streptococcus pneumoniae isolates from invasive and non-invasive sites. **12**, e0179075 (2017).

196. Dagan, R. et al. Dynamics of pneumococcal nasopharyngeal colonization during the first days of antibiotic treatment in pediatric patients. *Pediatr Infect Dis J* **17**, 880-5 (1998).

197. Dabernat, H. et al. Effects of cefixime or co-amoxiclav treatment on nasopharyngeal carriage of Streptococcus pneumoniae and Haemophilus influenzae in children with acute otitis media. *J Antimicrob Chemother* **41**, 253-8 (1998).

198. Chao, Y., Bergenfelz, C. & Håkansson, A.P. In Vitro and In Vivo Biofilm Formation by Pathogenic Streptococci. *Methods Mol Biol* **1535**, 285-299 (2017).

199. Chao, Y. et al. The serine protease HtrA plays a key role in heat-induced dispersal of pneumococcal biofilms. **10**, 22455 (2020).

200. Bond, J.S. Proteases: History, discovery, and roles in health and disease. *J Biol Chem* **294**, 1643-1651 (2019).

201. Martínez-García, S., Rodríguez-Martínez, S., Cancino-Diaz, M.E. & Cancino-Diaz, J.C. Extracellular proteases of Staphylococcus epidermidis: roles as virulence factors and their participation in biofilm. *Apmis* **126**, 177-185 (2018).

202. Stentzel, S. et al. Staphylococcal serine protease-like proteins are pacemakers of allergic airway reactions to Staphylococcus aureus. *J Allergy Clin Immunol* **139**, 492-500.e8 (2017).

203. Krysko, O. & Teufelberger, A. Protease/antiprotease network in allergy: The role of Staphylococcus aureus protease-like proteins. **74**, 2077-2086 (2019).

204. Burchacka, E. & Witkowska, D. The role of serine proteases in the pathogenesis of bacterial infections. *Postepy Hig Med Dosw (Online)* **70**, 678-94 (2016).

205. Thomas, V.C., Thurlow, L.R., Boyle, D. & Hancock, L.E. Regulation of autolysis-dependent extracellular DNA release by Enterococcus faecalis extracellular proteases influences biofilm development. *J Bacteriol* **190**, 5690-8 (2008).

206. Iwase, T. et al. Staphylococcus epidermidis Esp inhibits Staphylococcus aureus biofilm formation and nasal colonization. *Nature* **465**, 346-9 (2010).

207. Sugimoto, S. et al. Cloning, expression and purification of extracellular serine protease Esp, a biofilm-degrading enzyme, from Staphylococcus epidermidis. *J Appl Microbiol* **111**, 1406-15 (2011).

208. Zarzecka, U. et al. Properties of the HtrA Protease From Bacterium Helicobacter pylori Whose Activity Is Indispensable for Growth Under Stress Conditions. *Front Microbiol* **10**, 961 (2019).

209. Patel, S. A critical review on serine protease: Key immune manipulator and pathology mediator. *Allergol Immunopathol (Madr)* **45**, 579-591 (2017).

210. Barrett, A.J. & Rawlings, N.D. Families and clans of serine peptidases. *Arch Biochem Biophys* **318**, 247-50 (1995).

211. Marquart, M.E. Pathogenicity and virulence of Streptococcus pneumoniae: Cutting to the chase on proteases. *Virulence* **12**, 766-787 (2021).

212. Potempa, M. & Potempa, J. Protease-dependent mechanisms of complement evasion by bacterial pathogens. *Biol Chem* **393**, 873-88 (2012).

213. Banbula, A. et al. Amino-acid sequence and three-dimensional structure of the Staphylococcus aureus metalloproteinase at 1.72 A resolution. *Structure* **6**, 1185-93 (1998).

214. Macfarlane, G.T., Allison, C., Gibson, S.A. & Cummings, J.H. Contribution of the microflora to proteolysis in the human large intestine. *J Appl Bacteriol* **64**, 37-46 (1988).

215. Thibodeaux, B.A., Caballero, A.R., Marquart, M.E., Tommassen, J. & O'Callaghan, R.J. Corneal virulence of Pseudomonas aeruginosa elastase B and alkaline protease produced by Pseudomonas putida. *Curr Eye Res* **32**, 373-86 (2007).

216. Male, C.J. Immunoglobulin A1 protease production by Haemophilus influenzae and Streptococcus pneumoniae. *Infect Immun* **26**, 254-61 (1979).

217. Upadhye, V. et al. Inhibition of Mycobacterium tuberculosis secretory serine protease blocks bacterial multiplication both in axenic culture and in human macrophages. *Scand J Infect Dis* **41**, 569-76 (2009).

218. Wani, J.H., Gilbert, J.V., Plaut, A.G. & Weiser, J.N. Identification, cloning, and sequencing of the immunoglobulin A1 protease gene of Streptococcus pneumoniae. *Infect Immun* **64**, 3967-74 (1996).

219. Ishii, S., Yano, T. & Hayashi, H. Expression and characterization of the peptidase domain of Streptococcus pneumoniae ComA, a bifunctional ATP-binding cassette transporter involved in quorum sensing pathway. *J Biol Chem* **281**, 4726-31 (2006).

220. Kwon, K. et al. Recombinant expression and functional analysis of proteases from Streptococcus pneumoniae, Bacillus anthracis, and Yersinia pestis. *BMC Biochem* **12**, 17 (2011).

221. Proctor, M. & Manning, P.J. Production of immunoglobulin A protease by Streptococcus pneumoniae from animals. *Infect Immun* **58**, 2733-7 (1990).

222. Collin, M. & Olsén, A. Extracellular enzymes with immunomodulating activities: variations on a theme in Streptococcus pyogenes. *Infect Immun* **71**, 2983-92 (2003).

223. Poulsen, K. et al. A comprehensive genetic study of streptococcal immunoglobulin A1 proteases: evidence for recombination within and between species. *Infect Immun* **66**, 181-90 (1998).

224. Kilian, M., Mestecky, J., Kulhavy, R., Tomana, M. & Butler, W.T. IgA1 proteases from Haemophilus influenzae, Streptococcus pneumoniae, Neisseria meningitidis, and Streptococcus sanguis: comparative immunochemical studies. *J Immunol* **124**, 2596-600 (1980).

225. Frolet, C. et al. New adhesin functions of surface-exposed pneumococcal proteins. *BMC Microbiol* **10**, 190 (2010).

226. Gong, Y. et al. Immunization with a ZmpB-based protein vaccine could protect against pneumococcal diseases in mice. *Infect Immun* **79**, 867-78 (2011).

227. Andre, G.O. et al. Role of Streptococcus pneumoniae Proteins in Evasion of Complement-Mediated Immunity. *Front Microbiol* **8**, 224 (2017).

228. Angel, C.S., Ruzek, M. & Hostetter, M.K. Degradation of C3 by Streptococcus pneumoniae. *J Infect Dis* **170**, 600-8 (1994).

229. Zhang, Y. et al. Recombinant PhpA protein, a unique histidine motif-containing protein from Streptococcus pneumoniae, protects mice against intranasal pneumococcal challenge. *Infect Immun* **69**, 3827-36 (2001).

230. Agarwal, V. et al. Binding of Streptococcus pneumoniae endopeptidase O (PepO) to complement component C1q modulates the complement attack and promotes host cell adherence. *J Biol Chem* **289**, 15833-44 (2014).

231. Agarwal, V. et al. Streptococcus pneumoniae endopeptidase O (PepO) is a multifunctional plasminogen- and fibronectin-binding protein, facilitating evasion of innate immunity and invasion of host cells. *J Biol Chem* **288**, 6849-63 (2013).

232. Löfling, J., Vimberg, V., Battig, P. & Henriques-Normark, B. Cellular interactions by LPxTG-anchored pneumococcal adhesins and their streptococcal homologues. *Cell Microbiol* **13**, 186-97 (2011).

233. Nobbs, A.H., Lamont, R.J. & Jenkinson, H.F. Streptococcus adherence and colonization. *Microbiol Mol Biol Rev* **73**, 407-50, Table of Contents (2009).

234. de Stoppelaar, S.F. et al. Streptococcus pneumoniae serine protease HtrA, but not SFP or PrtA, is a major virulence factor in pneumonia. *PLoS One* **8**, e80062 (2013).

235. Courtney, H.S. Degradation of connective tissue proteins by serine proteases from Streptococcus pneumoniae. *Biochem Biophys Res Commun* **175**, 1023-8 (1991).

236. Bethe, G. et al. The cell wall-associated serine protease PrtA: a highly conserved virulence factor of Streptococcus pneumoniae. *FEMS Microbiol Lett* **205**, 99-104 (2001).

237. Desa, M.N., Sekaran, S.D., Vadivelu, J. & Parasakthi, N. Distribution of CBP genes in Streptococcus pneumoniae isolates in relation to vaccine types, penicillin susceptibility and clinical site. *Epidemiol Infect* **136**, 940-2 (2008).

238. Mirza, S. et al. Serine Protease PrtA from Streptococcus pneumoniae Plays a Role in the Killing of S. pneumoniae by Apolactoferrin. *Infection and Immunity* **79**, 2440-2450 (2011).

239. Ali, M.Q. et al. Extracellular Pneumococcal Serine Proteases Affect Nasopharyngeal Colonization. *Frontiers in Cellular and Infection Microbiology* **10** (2021).

240. Mirza, S. et al. Serine protease PrtA from Streptococcus pneumoniae plays a role in the killing of S. pneumoniae by apolactoferrin. *Infect Immun* **79**, 2440-50 (2011).

241. Ibrahim, Y.M., Kerr, A.R., McCluskey, J. & Mitchell, T.J. Control of virulence by the two-component system CiaR/H is mediated via HtrA, a major virulence factor of Streptococcus pneumoniae. *J Bacteriol* **186**, 5258-66 (2004).

242. Ibrahim, Y.M., Kerr, A.R., McCluskey, J. & Mitchell, T.J. Role of HtrA in the virulence and competence of Streptococcus pneumoniae. *Infect Immun* **72**, 3584-91 (2004).

243. Cassone, M., Gagne, A.L., Spruce, L.A., Seeholzer, S.H. & Sebert, M.E. The HtrA protease from Streptococcus pneumoniae digests both denatured proteins and the competence-stimulating peptide. *J Biol Chem* **287**, 38449-59 (2012).

244. Kochan, T.J. & Dawid, S. The HtrA protease of Streptococcus pneumoniae controls density-dependent stimulation of the bacteriocin blp locus via disruption of pheromone secretion. *J Bacteriol* **195**, 1561-72 (2013).

245. Sender, V. et al. Capillary leakage provides nutrients and antioxidants for rapid pneumococcal proliferation in influenza-infected lower airways. *Proceedings of the National Academy of Sciences* **117**, 31386-31397 (2020).

246. Håvarstein, L.S., Martin, B., Johnsborg, O., Granadel, C. & Claverys, J.P. New insights into the pneumococcal fratricide: relationship to clumping and identification of a novel immunity factor. *Mol Microbiol* **59**, 1297-307 (2006).

247. Cao, J. et al. CD4(+) T lymphocytes mediated protection against invasive pneumococcal infection induced by mucosal immunization with ClpP and CbpA. *Vaccine* **27**, 2838-44 (2009).

248. Cao, J. et al. Pneumococcal ClpP modulates the maturation and activation of human dendritic cells: implications for pneumococcal infections. *J Leukoc Biol* **93**, 737-49 (2013).

249. Chastanet, A., Prudhomme, M., Claverys, J.P. & Msadek, T. Regulation of Streptococcus pneumoniae clp genes and their role in competence development and stress survival. *J Bacteriol* **183**, 7295-307 (2001).

250. Robertson, G.T., Ng, W.L., Foley, J., Gilmour, R. & Winkler, M.E. Global transcriptional analysis of clpP mutations of type 2 Streptococcus pneumoniae and their effects on physiology and virulence. *J Bacteriol* **184**, 3508-20 (2002).

251. Liu, X. et al. High-throughput CRISPRi phenotyping identifies new essential genes in Streptococcus pneumoniae. *Mol Syst Biol* **13**, 931 (2017).

252. Sung, C.K. & Morrison, D.A. Two distinct functions of ComW in stabilization and activation of the alternative sigma factor ComX in Streptococcus pneumoniae. *J Bacteriol* **187**, 3052-61 (2005).

253. Kwon, H.Y. et al. Effect of heat shock and mutations in ClpL and ClpP on virulence gene expression in Streptococcus pneumoniae. *Infect Immun* **71**, 3757-65 (2003).

254. Ritchie, N.D. & Evans, T.J. Dual RNA-seq in Streptococcus pneumoniae Infection Reveals Compartmentalized Neutrophil Responses in Lung and Pleural Space. *mSystems* **4** (2019).

255. Armstrong, R.N. Mechanistic diversity in a metalloenzyme superfamily. *Biochemistry* **39**, 13625-32 (2000).

256. Carter, R. et al. Genomic analyses of pneumococci from children with sickle cell disease expose host-specific bacterial adaptations and deficits in current interventions. *Cell Host Microbe* **15**, 587-599 (2014).

257. Rowe, H.M. et al. Bacterial Factors Required for Transmission of Streptococcus pneumoniae in Mammalian Hosts. *Cell Host Microbe* **25**, 884-891.e6 (2019).

258. Hostetter, M.K. Opsonic and nonopsonic interactions of C3 with Streptococcus pneumoniae. *Microb Drug Resist* **5**, 85-9 (1999).

259. Nganje, C.N. et al. PepN is a non-essential, cell wall-localized protein that contributes to neutrophil elastase-mediated killing of Streptococcus pneumoniae. *PLoS One* **14**, e0211632 (2019).

260. Blevins, L.K. et al. A Novel Function for the Streptococcus pneumoniae Aminopeptidase N: Inhibition of T Cell Effector Function through Regulation of TCR Signaling. *Front Immunol* **8**, 1610 (2017).

261. Shu, Z. et al. Streptococcus pneumoniae PepO promotes host anti-infection defense via autophagy in a Toll-like receptor 2/4 dependent manner. *Virulence* **11**, 270-282 (2020).

262. Yao, H. et al. Purified Streptococcus pneumoniae Endopeptidase O (PepO) Enhances Particle Uptake by Macrophages in a Toll-Like Receptor 2- and miR-155-Dependent Manner. *Infect Immun* **85** (2017).

263. Zhang, H. et al. Streptococcus pneumoniae Endopeptidase O (PepO) Elicits a Strong Innate Immune Response in Mice via TLR2 and TLR4 Signaling Pathways. *Front Cell Infect Microbiol* **6**, 23 (2016).

264. Abdullah, M.R. et al. Structure of the pneumococcal l,d-carboxypeptidase DacB and pathophysiological effects of disabled cell wall hydrolases DacA and DacB. *Mol Microbiol* **93**, 1183-206 (2014).

265. Barendt, S.M., Sham, L.T. & Winkler, M.E. Characterization of mutants deficient in the L,D-carboxypeptidase (DacB) and WalRK (VicRK) regulon, involved in peptidoglycan maturation of Streptococcus pneumoniae serotype 2 strain D39. *J Bacteriol* **193**, 2290-300 (2011).

266. Hoyland, C.N. et al. Structure of the LdcB LD-carboxypeptidase reveals the molecular basis of peptidoglycan recognition. *Structure* **22**, 949-60 (2014).

267. Kilian, M., Mestecky, J. & Schrohenloher, R.E. Pathogenic species of the genus Haemophilus and Streptococcus pneumoniae produce immunoglobulin A1 protease. *Infect Immun* **26**, 143-9 (1979).

268. Wikström, M.B., Dahlén, G., Kaijser, B. & Nygren, H. Degradation of human immunoglobulins by proteases from Streptococcus pneumoniae obtained from various human sources. *Infect Immun* **44**, 33-7 (1984).

269. Reinholdt, J. & Kilian, M. Comparative analysis of immunoglobulin A1 protease activity among bacteria representing different genera, species, and strains. *Infect Immun* **65**, 4452-9 (1997).
270. Poulsen, K., Reinholdt, J. & Kilian, M. Characterization of the Streptococcus pneumoniae immunoglobulin A1 protease gene (iga) and its translation product. *Infect Immun* **64**, 3957-66 (1996).
271. Batten, M.R., Senior, B.W., Kilian, M. & Woof, J.M. Amino acid sequence requirements in the hinge of human immunoglobulin A1 (IgA1) for cleavage by streptococcal IgA1 proteases. *Infect Immun* **71**, 1462-9 (2003).
272. Bergé, M. et al. The puzzle of zmpB and extensive chain formation, autolysis defect and non-translocation of choline-binding proteins in Streptococcus pneumoniae. *Mol Microbiol* **39**, 1651-60 (2001).
273. Chiavolini, D. et al. The three extra-cellular zinc metalloproteinases of Streptococcus pneumoniae have a different impact on virulence in mice. *BMC Microbiol* **3**, 14 (2003).
274. Blue, C.E. et al. ZmpB, a novel virulence factor of Streptococcus pneumoniae that induces tumor necrosis factor alpha production in the respiratory tract. *Infect Immun* **71**, 4925-35 (2003).
275. Hsieh, Y.C. et al. Establishment of a young mouse model and identification of an allelic variation of zmpB in complicated pneumonia caused by Streptococcus pneumoniae. *Crit Care Med* **36**, 1248-55 (2008).
276. Novak, R. et al. Extracellular targeting of choline-binding proteins in Streptococcus pneumoniae by a zinc metalloprotease. *Mol Microbiol* **36**, 366-76 (2000).
277. Camilli, R. et al. Zinc metalloproteinase genes in clinical isolates of Streptococcus pneumoniae: association of the full array with a clonal cluster comprising serotypes 8 and 11A. *Microbiology (Reading)* **152**, 313-321 (2006).
278. Cremers, A.J., Kokmeijer, I., Groh, L., de Jonge, M.I. & Ferwerda, G. The role of ZmpC in the clinical manifestation of invasive pneumococcal disease. *Int J Med Microbiol* **304**, 984-9 (2014).
279. Surewaard, B.G. et al. Pneumococcal immune evasion: ZmpC inhibits neutrophil influx. *Cell Microbiol* **15**, 1753-65 (2013).
280. Oggioni, M.R. et al. Pneumococcal zinc metalloproteinase ZmpC cleaves human matrix metalloproteinase 9 and is a virulence factor in experimental pneumonia. *Mol Microbiol* **49**, 795-805 (2003).
281. Chen, Y., Hayashida, A., Bennett, A.E., Hollingshead, S.K. & Park, P.W. Streptococcus pneumoniae sheds syndecan-1 ectodomains through ZmpC, a metalloproteinase virulence factor. *J Biol Chem* **282**, 159-67 (2007).
282. Yamaguchi, M. et al. Zinc metalloproteinase ZmpC suppresses experimental pneumococcal meningitis by inhibiting bacterial invasion of central nervous systems. *Virulence* **8**, 1516-1524 (2017).
283. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **44**, D7-19 (2016).
284. Seol, J.H. et al. Protease Do is essential for survival of Escherichia coli at high temperatures: its identity with the htrA gene product. *Biochem Biophys Res Commun* **176**, 730-6 (1991).
285. Lipinska, B., Zylicz, M. & Georgopoulos, C. The HtrA (DegP) protein, essential for Escherichia coli survival at high temperatures, is an endopeptidase. *J Bacteriol* **172**, 1791-7 (1990).
286. Spiess, C., Beil, A. & Ehrmann, M. A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell* **97**, 339-47 (1999).
287. Backert, S., Bernegger, S., Skorko-Glonek, J. & Wessler, S. Extracellular HtrA serine proteases: An emerging new strategy in bacterial pathogenesis. *Cell Microbiol* **20**, e12845 (2018).
288. Clausen, T., Southan, C. & Ehrmann, M. The HtrA family of proteases: implications for protein composition and cell fate. *Mol Cell* **10**, 443-55 (2002).
289. Kim, D.Y. et al. Crystal structure of the protease domain of a heat-shock protein HtrA from Thermotoga maritima. *J Biol Chem* **278**, 6543-51 (2003).
290. Zarzecka, U., Harrer, A., Zawilak-Pawlik, A., Skorko-Glonek, J. & Backert, S. Chaperone activity of serine protease HtrA of Helicobacter pylori as a crucial survival factor under stress conditions. **17**, 161 (2019).
291. Hansen, G. & Hilgenfeld, R. Architecture and regulation of HtrA-family proteins involved in protein quality control and stress response. *Cell Mol Life Sci* **70**, 761-75 (2013).
292. Singh, K.H., Yadav, S., Kumar, D. & Biswal, B.K. The crystal structure of an essential high-temperature requirement protein HtrA1 (Rv1223) from Mycobacterium tuberculosis reveals its unique features. *Acta Crystallogr D Struct Biol* **74**, 906-921 (2018).
293. Poquet, I. et al. HtrA is the unique surface housekeeping protease in Lactococcus lactis and is required for natural protein processing. *Mol Microbiol* **35**, 1042-51 (2000).
294. Clausen, T., Kaiser, M., Huber, R. & Ehrmann, M. HTRA proteases: regulated proteolysis in protein quality control. *Nat Rev Mol Cell Biol* **12**, 152-62 (2011).
295. Malet, H. et al. Newly folded substrates inside the molecular cage of the HtrA chaperone DegQ. *Nat Struct Mol Biol* **19**, 152-7 (2012).
296. Boehm, M. et al. Function of Serine Protease HtrA in the Lifecycle of the Foodborne Pathogen Campylobacter jejuni. *Eur J Microbiol Immunol (Bp)* **8**, 70-77 (2018).
297. Bai, X.C. et al. Characterization of the structure and function of Escherichia coli DegQ as a representative of the DegQ-like proteases of bacterial HtrA family proteins. *Structure* **19**, 1328-37 (2011).
298. Chang, Z. The function of the DegP (HtrA) protein: Protease versus chaperone. *IUBMB Life* **68**, 904-907 (2016).
299. Cortes, T. et al. Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in Mycobacterium tuberculosis. *Cell Rep* **5**, 1121-31 (2013).
300. Murwantoko et al. Binding of proteins to the PDZ domain regulates proteolytic activity of HtrA1 serine protease. *Biochem J* **381**, 895-904 (2004).
301. Pestova, E.V., Håvarstein, L.S. & Morrison, D.A. Regulation of competence for genetic transformation in Streptococcus pneumoniae by an auto-induced peptide pheromone and a two-component regulatory system. *Mol Microbiol* **21**, 853-62 (1996).

302. Peters, K., Schweizer, I., Hakenbeck, R. & Denapaite, D. New Insights into Beta-Lactam Resistance of Streptococcus pneumoniae: Serine Protease HtrA Degrades Altered Penicillin-Binding Protein 2x. *Microorganisms* **9** (2021).

303. Li, Y., Hill, A., Beitelshees, M., Shao, S. & Lovell, J.F. Directed vaccination against pneumococcal disease. **113**, 6898-903 (2016).

304. Wessler, S., Schneider, G. & Backert, S. Bacterial serine protease HtrA as a promising new target for antimicrobial therapy? *Cell Commun Signal* **15**, 4 (2017).

305. Kanz, C. et al. The EMBL Nucleotide Sequence Database. *Nucleic Acids Res* **33**, D29-33 (2005).

306. Yang, M., Derbyshire, M.K., Yamashita, R.A. & Marchler-Bauer, A. NCBI's Conserved Domain Database and Tools for Protein Domain Analysis. *Curr Protoc Bioinformatics* **69**, e90 (2020).

307. Kazemian, H. et al. CbpM and CbpG of Streptococcus Pneumoniae Elicit a High Protection in Mice Challenged with a Serotype 19F Pneumococcus. *Iran J Allergy Asthma Immunol* **17**, 574-585 (2018).

308. Garcia-Bustos, J.F. & Tomasz, A. Teichoic acid-containing muropeptides from Streptococcus pneumoniae as substrates for the pneumococcal autolysin. *J Bacteriol* **169**, 447-53 (1987).

309. Cheng, Q., Stafslien, D., Purushothaman, S.S. & Cleary, P. The group B streptococcal C5a peptidase is both a specific protease and an invasin. *Infect Immun* **70**, 2408-13 (2002).

310. Beckmann, C., Waggoner, J.D., Harris, T.O., Tamura, G.S. & Rubens, C.E. Identification of novel adhesins from Group B streptococci by use of phage display reveals that C5a peptidase mediates fibronectin binding. *Infect Immun* **70**, 2869-76 (2002).

311. Kukkonen, M. & Korhonen, T.K. The omptin family of enterobacterial surface proteases/adhesins: from housekeeping in Escherichia coli to systemic spread of Yersinia pestis. *Int J Med Microbiol* **294**, 7-14 (2004).

312. Bateman, A. & Rawlings, N.D. The CHAP domain: a large family of amidases including GSP amidase and peptidoglycan hydrolases. *Trends Biochem Sci* **28**, 234-7 (2003).

313. Park, C.Y. et al. Virulence attenuation of Streptococcus pneumoniae clpP mutant by sensitivity to oxidative stress in macrophages via an NO-mediated pathway. *J Microbiol* **48**, 229-35 (2010).

314. Mahdi, L.K., Ogunniyi, A.D., LeMessurier, K.S. & Paton, J.C. Pneumococcal virulence gene expression and host cytokine profiles during pathogenesis of invasive disease. *Infect Immun* **76**, 646-57 (2008).

315. Sempere, J. et al. Pneumococcal Choline-Binding Proteins Involved in Virulence as Vaccine Candidates. *Vaccines (Basel)* **9** (2021).

316. Siezen, R.J. Multi-domain, cell-envelope proteinases of lactic acid bacteria. *Antonie Van Leeuwenhoek* **76**, 139-55 (1999).

317. Zysk, G. et al. Detection of 23 immunogenic pneumococcal proteins using convalescent-phase serum. *Infect Immun* **68**, 3740-3 (2000).

318. Mahdi, L.K., Van der Hoek, M.B., Ebrahimie, E., Paton, J.C. & Ogunniyi, A.D. Characterization of Pneumococcal Genes Involved in Bloodstream Invasion in a Mouse Model. *PLoS One* **10**, e0141816 (2015).

319. Wizemann, T.M. et al. Use of a whole genome approach to identify vaccine molecules affording protection against Streptococcus pneumoniae infection. *Infect Immun* **69**, 1593-8 (2001).

320. Vos, P. et al. A maturation protein is essential for production of active forms of Lactococcus lactis SK11 serine proteinase located in or secreted from the cell envelope. *J Bacteriol* **171**, 2795-802 (1989).

321. Holck, A. & Naes, H. Cloning, sequencing and expression of the gene encoding the cell-envelope-associated proteinase from Lactobacillus paracasei subsp. paracasei NCDO 151. *J Gen Microbiol* **138**, 1353-64 (1992).

322. Bonifait, L. et al. The cell envelope subtilisin-like proteinase is a virulence determinant for Streptococcus suis. *BMC Microbiol* **10**, 42 (2010).

323. Hsu, C.F. et al. PrtA immunization fails to protect against pulmonary and invasive infection by Streptococcus pneumoniae. *Respir Res* **19**, 187 (2018).

324. Bryan, J.D. & Shelver, D.W. Streptococcus agalactiae CspA is a serine protease that inactivates chemokines. *J Bacteriol* **191**, 1847-54 (2009).

325. Harris, T.O., Shelver, D.W., Bohnsack, J.F. & Rubens, C.E. A novel streptococcal surface protease promotes virulence, resistance to opsonophagocytosis, and cleavage of human fibrinogen. *J Clin Invest* **111**, 61-70 (2003).

326. Gamez, G. & Hammerschmidt, S. Combat pneumococcal infections: adhesins as candidates for protein-based vaccine development. *Curr Drug Targets* **13**, 323-37 (2012).

327. Tomasz, A. Choline in the cell wall of a bacterium: novel type of polymer-linked choline in Pneumococcus. *Science* **157**, 694-7 (1967).

328. McCullers, J.A. & Tuomanen, E.I. Molecular pathogenesis of pneumococcal pneumonia. *Front Biosci* **6**, D877-89 (2001).

329. Ali, M.Q., Kohler, T.P., Schulig, L., Burchhardt, G. & Hammerschmidt, S. Pneumococcal Extracellular Serine Proteases: Molecular Analysis and Impact on Colonization and Disease. *Front Cell Infect Microbiol* **11**, 763152 (2021).

330. NCBI. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **44**, D7-19 (2016).

331. Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**, 276-7 (2000).

332. Kanehisa, M. et al. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* **34**, D354-7 (2006).

333. Zdobnov, E.M. & Apweiler, R. InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* **17**, 847-8 (2001).

334. Fan, K., Zhang, J., Zhang, X. & Tu, X. Solution structure of HtrA PDZ domain from Streptococcus pneumoniae and its interaction with YYF-COOH containing peptides. *J Struct Biol* **176**, 16-23 (2011).

335. Wilken, C., Kitzing, K., Kurzbauer, R., Ehrmann, M. & Clausen, T. Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease. *Cell* **117**, 483-94 (2004).

336. Hasselblatt, H. et al. Regulation of the sigmaE stress response by DegS: how the PDZ domain keeps the protease inactive in the resting state and allows integration of different OMP-derived stress signals upon folding stress. *Genes Dev* **21**, 2659-70 (2007).

337. Bohnsack, J.F. et al. Genetic polymorphisms of group B streptococcus scpB alter functional activity of a cell-associated peptidase that inactivates C5a. *Infect Immun* **68**, 5018-25 (2000).

338. Nielsen, H. Predicting Secretory Proteins with SignalP. *Methods Mol Biol* **1611**, 59-73 (2017).

339. Yother, J. & White, J.M. Novel surface attachment mechanism of the Streptococcus pneumoniae protein PspA. *J Bacteriol* **176**, 2976-85 (1994).

340. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. & Barton, G.J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-91 (2009).

341. Andersson, B. et al. Identification of an active disaccharide unit of a glycoconjugate receptor for pneumococci attaching to human pharyngeal epithelial cells. *J Exp Med* **158**, 559-70 (1983).

342. Junges, R., Maienschein-Cline, M., Morrison, D.A. & Petersen, F.C. Complete Genome Sequence of Streptococcus pneumoniae Serotype 19F Strain EF3030. **8** (2019).

343. Wüst, A.W. in Molecular Genetics and Infection Microbiology (Universität Greifswald, 2017).

344. Mehlitz, N. in Department of Molecular Genetics and Infection Biology, Interfaculty Institute of Genetics and Functional Genomics, Center for Functional Genomics of Microbes (University of Greifswald, University of Greifswald 2014).

345. Riley, D.R., Angiuoli, S.V., Crabtree, J., Dunning Hotopp, J.C. & Tettelin, H. Using Sybil for interactive comparative genomics of microbes on the web. *Bioinformatics* **28**, 160-166 (2011).

346. Mickelson, M.N. CHEMICALLY DEFINED MEDIUM FOR GROWTH STREPTOCOCCUS PYOGENES. *J Bacteriol* **88**, 158-64 (1964).

347. Leonard, C.G., Ranhand, J.M. & Cole, R.M. Competence factor production in chemically defined media by noncompetent cells of group H Streptococcus strain Challis. *J Bacteriol* **104**, 674-83 (1970).

348. Härtel, T. et al. Impact of glutamine transporters on pneumococcal fitness under infection-related conditions. *Infect Immun* **79**, 44-58 (2011).

349. Bootsma, H.J., Egmont-Petersen, M. & Hermans, P.W. Analysis of the in vitro transcriptional response of human pharyngeal epithelial cells to adherent Streptococcus pneumoniae: evidence for a distinct response to encapsulated strains. *Infect Immun* **75**, 5489-99 (2007).

350. Bogaert, D., de Groot, R. & Hermans, P.W.M. Streptococcus pneumoniae colonisation: the key to pneumococcal disease. *The Lancet Infectious Diseases* **4**, 144-154 (2004).

351. Short, K.R., Reading, P.C., Wang, N., Diavatopoulos, D.A. & Wijburg, O.L. Increased nasopharyngeal bacterial titers and local inflammation facilitate transmission of Streptococcus pneumoniae. *mBio* **3** (2012).

352. Hoa, M. et al. Identification of adenoid biofilms with middle ear pathogens in otitis-prone children utilizing SEM and FISH. *Int J Pediatr Otorhinolaryngol* **73**, 1242-8 (2009).

353. Cohen, R. et al. One dose ceftriaxone vs. ten days of amoxicillin/clavulanate therapy for acute otitis media: clinical efficacy and change in nasopharyngeal flora. *Pediatr Infect Dis J* **18**, 403-9 (1999).

354. Håkansson, A. et al. Aspects on the interaction of Streptococcus pneumoniae and Haemophilus influenzae with human respiratory tract mucosa. *Am J Respir Crit Care Med* **154**, S187-91 (1996).

355. Marks, L.R., Davidson, B.A., Knight, P.R. & Hakansson, A.P. Interkingdom signaling induces Streptococcus pneumoniae biofilm dispersion and transition from asymptomatic colonization to disease. *mBio* **4** (2013).

356. Bolsmann, R. (Universität Greifswald, 2017).

357. Chung, C.H. & Goldberg, A.L. Purification and characterization of protease So, a cytoplasmic serine protease in Escherichia coli. *J Bacteriol* **154**, 231-8 (1983).

358. Simell, B. et al. The fundamental link between pneumococcal carriage and disease. *Expert Rev Vaccines* **11**, 841-55 (2012).

359. Harrer, A., Boehm, M., Backert, S. & Tegtmeyer, N. Overexpression of serine protease HtrA enhances disruption of adherens junctions, paracellular transmigration and type IV secretion of CagA by Helicobacter pylori. *Gut Pathog* **9**, 40 (2017).

360. Cleary, P.P., Prahbu, U., Dale, J.B., Wexler, D.E. & Handley, J. Streptococcal C5a peptidase is a highly specific endopeptidase. *Infect Immun* **60**, 5219-23 (1992).

361. Supuran, C.T., Scozzafava, A. & Clare, B.W. Bacterial protease inhibitors. *Med Res Rev* **22**, 329-72 (2002).

362. Janoff, E.N. et al. Pneumococcal IgA1 protease subverts specific protection by human IgA1. *Mucosal Immunology* **7**, 249-256 (2014).

363. Govindarajan, B. et al. A metalloproteinase secreted by Streptococcus pneumoniae removes membrane mucin MUC16 from the epithelial glycocalyx barrier. *PLoS One* **7**, e32418 (2012).

364. Joloba, M.L. et al. Pneumococcal conjugate vaccine serotypes of Streptococcus pneumoniae isolates and the antimicrobial susceptibility of such isolates in children with otitis media. *Clin Infect Dis* **33**, 1489-94 (2001).

365. Blevins, L.K. et al. Coinfection with Streptococcus pneumoniae negatively modulates the size and composition of the ongoing influenza-specific CD8[+] T cell response. *J Immunol* **193**, 5076-87 (2014).

366. Brondsted, L., Andersen, M.T., Parker, M., Jorgensen, K. & Ingmer, H. The HtrA protease of Campylobacter jejuni is required for heat and oxygen tolerance and for optimal interaction with human epithelial cells. *Appl Environ Microbiol* **71**, 3205-12 (2005).

367. Frees, D., Brondsted, L. & Ingmer, H. Bacterial proteases and virulence. *Subcell Biochem* **66**, 161-92 (2013).

368. Simson, D., Boehm, M. & Backert, S. HtrA-dependent adherence and invasion of Campylobacter jejuni in human vs avian cells. *Lett Appl Microbiol* **70**, 326-330 (2020).

369. Bæk, K.T., Vegge, C.S. & Brøndsted, L. HtrA chaperone activity contributes to host cell binding in Campylobacter jejuni. *Gut Pathog* **3**, 13 (2011).
370. Frees, D., Brøndsted, L. & Ingmer, H. Bacterial proteases and virulence. *Subcell Biochem* **66**, 161-92 (2013).
371. Hoy, B., Brandstetter, H. & Wessler, S. The stability and activity of recombinant Helicobacter pylori HtrA under stress conditions. *J Basic Microbiol* **53**, 402-9 (2013).
372. Jomaa, A., Iwanczyk, J., Tran, J. & Ortega, J. Characterization of the autocleavage process of the Escherichia coli HtrA protein: implications for its physiological role. *J Bacteriol* **191**, 1924-32 (2009).
373. Desa, M.N., Navaratnam, P., Vadivelu, J. & Sekaran, S.D. Expression analysis of adherence-associated genes in pneumococcal clinical isolates during adherence to human respiratory epithelial cells (in vitro) by real-time PCR. *FEMS Microbiol Lett* **288**, 125-30 (2008).
374. Luo, R. et al. Solution structure of choline binding protein A, the major adhesin of Streptococcus pneumoniae. *Embo j* **24**, 34-43 (2005).
375. Radziwill-Bienkowska, J.M. et al. Contribution of plasmid-encoded peptidase S8 (PrtP) to adhesion and transit in the gut of Lactococcus lactis IBB477 strain. **101**, 5709-5721 (2017).
376. van Ginkel, F.W. et al. Pneumococcal carriage results in ganglioside-mediated olfactory tissue infection. *Proceedings of the National Academy of Sciences* **100**, 14363-14367 (2003).
377. McCool, T.L. & Weiser, J.N. Limited role of antibody in clearance of Streptococcus pneumoniae in a murine model of colonization. *Infect Immun* **72**, 5807-13 (2004).
378. Ji, Y., McLandsborough, L., Kondagunta, A. & Cleary, P.P. C5a peptidase alters clearance and trafficking of group A streptococci by infected mice. *Infect Immun* **64**, 503-10 (1996).
379. Potempa, J. & Pike, R.N. Corruption of innate immunity by bacterial proteases. *J Innate Immun* **1**, 70-87 (2009).
380. Saleh, M. et al. Following in real time the impact of pneumococcal virulence factors in an acute mouse pneumonia model using bioluminescent bacteria. *J Vis Exp*, e51174 (2014).
381. Gingles, N.A. et al. Role of genetic resistance in invasive pneumococcal infection: identification and study of susceptibility and resistance in inbred mouse strains. *Infect Immun* **69**, 426-34 (2001).
382. Schulz, C. et al. Regulation of the arginine deiminase system by ArgR2 interferes with arginine metabolism and fitness of Streptococcus pneumoniae. **5** (2014).
383. Frolet, C. et al. New adhesin functions of surface-exposed pneumococcal proteins. *BMC Microbiology* **10**, 190 (2010).
384. Paterson, G.K. & Mitchell, T.J. The role of Streptococcus pneumoniae sortase A in colonisation and pathogenesis. *Microbes Infect* **8**, 145-53 (2006).
385. Parker, D. et al. The NanA neuraminidase of Streptococcus pneumoniae is involved in biofilm formation. *Infect Immun* **77**, 3722-30 (2009).
386. Biswas, S. & Biswas, I. Role of HtrA in surface protein expression and biofilm formation by Streptococcus mutans. *Infect Immun* **73**, 6923-34 (2005).
387. Donlan, R.M. et al. Model system for growing and quantifying Streptococcus pneumoniae biofilms in situ and in real time. *Appl Environ Microbiol* **70**, 4980-8 (2004).
388. Connolly, K.L., Roberts, A.L., Holder, R.C. & Reid, S.D. Dispersal of Group A streptococcal biofilms by the cysteine protease SpeB leads to increased disease severity in a murine model. *PLoS One* **6**, e18984 (2011).
389. Domenech, M. & García, E. The N-Acetylglucosaminidase LytB of Streptococcus pneumoniae Is Involved in the Structure and Formation of Biofilms. *Appl Environ Microbiol* **86** (2020).
390. Ramos-Sevillano, E., Moscoso, M., García, P., García, E. & Yuste, J. Nasopharyngeal colonization and invasive disease are enhanced by the cell wall hydrolases LytB and LytC of Streptococcus pneumoniae. *PLoS One* **6**, e23626 (2011).
391. Giebink, G.S. Otitis media: the chinchilla model. *Microb Drug Resist* **5**, 57-72 (1999).
392. Xu, Y. et al. Structure of the nisin leader peptidase NisP revealing a C-terminal autocleavage activity. *Acta Crystallogr D Biol Crystallogr* **70**, 1499-505 (2014).
393. Stamenkovic, I. Extracellular matrix remodelling: the role of matrix metalloproteinases. *J Pathol* **200**, 448-64 (2003).
394. Iovino, F., Seinen, J., Henriques-Normark, B. & van Dijl, J.M. How Does Streptococcus pneumoniae Invade the Brain? *Trends Microbiol* **24**, 307-315 (2016).
395. Devaux, C.A., Mezouar, S. & Mege, J.L. The E-Cadherin Cleavage Associated to Pathogenic Bacteria Infections Can Favor Bacterial Invasion and Transmigration, Dysregulation of the Immune Response and Cancer Induction in Humans. *Front Microbiol* **10**, 2598 (2019).
396. Courtney, H.S., Ofek, I., Simpson, W.A., Hasty, D.L. & Beachey, E.H. Binding of Streptococcus pyogenes to soluble and insoluble fibronectin. *Infect Immun* **53**, 454-9 (1986).
397. Kuypers, J.M. & Proctor, R.A. Reduced adherence to traumatized rat heart valves by a low-fibronectin-binding mutant of Staphylococcus aureus. *Infect Immun* **57**, 2306-12 (1989).
398. Schmidt, T.P., Goetz, C., Huemer, M., Schneider, G. & Wessler, S. Calcium binding protects E-cadherin from cleavage by Helicobacter pylori HtrA. *Gut Pathogens* **8**, 29 (2016).
399. Zawilak-Pawlik, A., Zarzecka, U., Żyła-Uklejewicz, D., Lach, J. & Strapagiel, D. Establishment of serine protease htrA mutants in Helicobacter pylori is associated with secA mutations. **9**, 11794 (2019).
400. Hulpiau, P. & van Roy, F. Molecular evolution of the cadherin superfamily. *Int J Biochem Cell Biol* **41**, 349-69 (2009).
401. Anderton, J.M. et al. E-cadherin is a receptor for the common protein pneumococcal surface adhesin A (PsaA) of Streptococcus pneumoniae. *Microb Pathog* **42**, 225-36 (2007).
402. Abfalter, C.M. et al. HtrA-mediated E-cadherin cleavage is limited to DegP and DegQ homologs expressed by gram-negative pathogens. **14**, 30 (2016).
403. Boehm, M. et al. Rapid paracellular transmigration of Campylobacter jejuni across polarized epithelial cells without affecting TER: role of proteolytic-active HtrA cleaving E-cadherin but not fibronectin. *Gut Pathog* **4**, 3 (2012).

404. von Schillde, M.A. et al. Lactocepin secreted by Lactobacillus exerts anti-inflammatory effects by selectively degrading proinflammatory chemokines. *Cell Host Microbe* **11**, 387-96 (2012).

405. Hörmannsperger, G., von Schillde, M.A. & Haller, D. Lactocepin as a protective microbial structure in the context of IBD. *Gut Microbes* **4**, 152-7 (2013).

406. Nikolić, M., Tolinacki, M., Fira, D., Golić, N. & Topisirović, L. Variation in specificity of the PrtP extracellular proteinases in Lactococcus lactis and Lactobacillus paracasei subsp. paracasei. *Folia Microbiol (Praha)* **54**, 188-94 (2009).

407. Inoue, M. et al. Subdural empyema due to Lactococcus lactis cremoris: case report. *Neurol Med Chir (Tokyo)* **54**, 341-7 (2014).

408. Abedi, D., Feizizadeh, S., Akbari, V. & Jafarian-Dehkordi, A. In vitro anti-bacterial and anti-adherence effects of Lactobacillus delbrueckii subsp bulgaricus on Escherichia coli. *Res Pharm Sci* **8**, 260-8 (2013).

409. Kunji, E.R., Mierau, I., Hagting, A., Poolman, B. & Konings, W.N. The proteolytic systems of lactic acid bacteria. *Antonie Van Leeuwenhoek* **70**, 187-221 (1996).

410. Courtin, P., Monnet, V. & Rul, F. Cell-wall proteinases PrtS and PrtB have a different role in Streptococcus thermophilus/Lactobacillus bulgaricus mixed cultures in milk. *Microbiology (Reading)* **148**, 3413-3421 (2002).

411. Chmouryguina, I., Suvorov, A., Ferrieri, P. & Cleary, P.P. Conservation of the C5a peptidase genes in group A and B streptococci. *Infect Immun* **64**, 2387-90 (1996).

412. Thurlow, L.R. et al. Gelatinase contributes to the pathogenesis of endocarditis caused by Enterococcus faecalis. *Infect Immun* **78**, 4936-43 (2010).

413. Pietrocola, G., Nobile, G., Rindi, S. & Speziale, P. Staphylococcus aureus Manipulates Innate Immunity through Own and Host-Expressed Proteases. *Front Cell Infect Microbiol* **7**, 166 (2017).

414. Jajere, S.M. A review of Salmonella enterica with particular focus on the pathogenicity and virulence factors, host specificity and antimicrobial resistance including multidrug resistance. *Vet World* **12**, 504-521 (2019).

415. Meijers, R. et al. The crystal structure of glutamyl endopeptidase from Bacillus intermedius reveals a structural link between zymogen activation and charge compensation. *Biochemistry* **43**, 2784-91 (2004).

416. Montalbán-López, M., Deng, J., van Heel, A.J. & Kuipers, O.P. Specificity and Application of the Lantibiotic Protease NisP. *Front Microbiol* **9**, 160 (2018).

417. Wexler, D.E., Chenoweth, D.E. & Cleary, P.P. Mechanism of action of the group A streptococcal C5a inactivator. *Proceedings of the National Academy of Sciences* **82**, 8144-8148 (1985).

418. Lyon, W.R. & Caparon, M.G. Role for serine protease HtrA (DegP) of Streptococcus pyogenes in the biogenesis of virulence factors SpeB and the hemolysin streptolysin S. *Infect Immun* **72**, 1618-25 (2004).

419. Zarzecka, U. et al. Functional analysis and cryo-electron microscopy of Campylobacter jejuni serine protease HtrA. *Gut Microbes* **12**, 1-16 (2020).

420. Hoy, B. et al. Helicobacter pylori HtrA is a new secreted virulence factor that cleaves E-cadherin to disrupt intercellular adhesion. *EMBO Rep* **11**, 798-804 (2010).

421. Tegtmeyer, N. et al. Helicobacter pylori Employs a Unique Basolateral Type IV Secretion Mechanism for CagA Delivery. *Cell Host Microbe* **22**, 552-560.e5 (2017).

422. Wu, K.J., Boutte, C.C., Ioerger, T.R. & Rubin, E.J. Mycobacterium smegmatis HtrA Blocks the Toxic Activity of a Putative Cell Wall Amidase. *Cell Rep* **27**, 2468-2479.e3 (2019).

423. Håvarstein, L.S., Coomaraswamy, G. & Morrison, D.A. An unmodified heptadecapeptide pheromone induces competence for genetic transformation in Streptococcus pneumoniae. *Proc Natl Acad Sci U S A* **92**, 11140-4 (1995).

424. Hammerschmidt, S., Talay, S.R., Brandtzaeg, P. & Chhatwal, G.S. SpsA, a novel pneumococcal surface protein with specific binding to secretory immunoglobulin A and secretory component. *Mol Microbiol* **25**, 1113-24 (1997).

425. Birnboim, H.C. & Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**, 1513-23 (1979).

426. Saiki, R.K. et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-91 (1988).

427. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-5 (1970).

428. Pettigrew, M.M. et al. Dynamic changes in the Streptococcus pneumoniae transcriptome during transition from biofilm formation to invasive disease upon influenza A virus infection. *Infect Immun* **82**, 4607-19 (2014).

429. Keck, T., Leiacker, R., Riechelmann, H. & Rettinger, G. Temperature profile in the nasal cavity. *Laryngoscope* **110**, 651-4 (2000).

430. Briles, D.E., Crain, M.J., Gray, B.M., Forman, C. & Yother, J. Strong association between capsular type and virulence for mice among human isolates of Streptococcus pneumoniae. *Infect Immun* **60**, 111-6 (1992).

431. Yu, N.Y., Laird, M.R., Spencer, C. & Brinkman, F.S. PSORTdb--an expanded, auto-updated, user-friendly protein subcellular localization database for Bacteria and Archaea. *Nucleic Acids Res* **39**, D241-4 (2011).

432. Petersen, T.N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* **8**, 785-6 (2011).

433. Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* **2**, 953-71 (2007).

434. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E.L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567-80 (2001).

8.10 Eigenständigkeitserklärung

Hiermit erkläre ich, dass diese Arbeit bisher von mir weder an der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Greifswald noch einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde.

Ferner erkläre ich, dass ich diese Arbeit selbständig verfasst und keine anderen als die darin angegebenen Hilfsmittel und Hilfen benutzt und keine Textabschnitte eines Dritten ohne Kennzeichnung übernommen habe.

Greifswald, …………………   …………………………………...

                                              Murtadha Q. Ali, BSc. MSc.

8.11 Curriculum Vitae

Personal Information

Address:
Felix-Hausdorff-Stra. 8   Mobile: +4917661265641
D-17487, Greifswald, Germany   Email: murtadhaqali@gmail.com
                                                          Date of Birth: 13/August/1990

LinkedIn: **https://www.linkedin.com/in/murtadha-q-ali-1a6964137/**

About me

Award-winning DAAD-Research Scientist based in Germany, affiliated with Department of Molecular Genetics and Infection Biology, University of Greifswald, offering 4+ years of blended expertise in molecular biology up-to-date technologies and approaches. These experiences have already resulted in three peer-reviewed international publications.

Education

2017–2021   **Ph.D. in Molecular Genetics and Infection Biology at the University of Greifswald, Germany.**

06-10/2017   **German language course for the PhD. scholars in Carl Duisberg Institute in Berlin. German Academic Exchange Service (DAAD).**

2012–2014   **M.Sc. in Biology, Zoology, University of Baghdad, Iraq.** Passed with **Distinction.**

2008–2012   **B.Sc. in Biology, College of Education for Pure Science, University of Baghdad, Iraq.** Passed with a **Very Good** grade. (Hons. Rank 19/161).

2005–2008   **Al-Rawabe Preparatory School (GCSEs), Baghdad, Iraq.**
                           Passed with **a Good** grade.

Job Experience

01/2022 present   Human Geneticist and Medical Reporter, Arcensus GmbH, Rostock, Germany.

10/2017–12/2021   Ph.D. student, German Academic Exchange Service (DAAD) Doctoral scholar in Prof. Sven Hammerschmidt's group, Molecular Genetics and Infection Biology at the University of Greifswald, Germany.

11/2015–06/2017   Biology teacher, Ministry of Education, Baghdad, Iraq.

04-10/2015   Lecturer, Ministry of High Education and Scientific Research, Al-Isra University, Baghdad, Iraq.

10/2012–10/2017   Laboratory assistant (Part-time work), Medical diagnostic laboratory, Baghdad, Iraq.

10/2009–10/2012   Quality Control (Part-time work), Soft drinks factory, Baghdad, Iraq.

Language skills:

Arabic: Mother tongue
English: Fluent (written and spoken)
German: Good
Spanish: Basic

Professional Training and Memberships
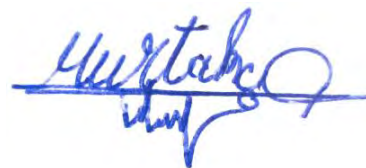
2021-Present   Member of the German Society for Hygiene and Microbiology – DGHM
2020-Present   Member of the German Society for General and Applied Microbiology –VAAM
2018   European Health Law of the Federation of Laboratory Animal Science Associations (FELASA) mice and rat. Greifswald, Germany.
2020- Present   Member of the non-profit organization Humanitär ohne Grenzen (HOG), Germany.
2012-2017   Committee Advisor in Directorial General Education of Baghdad, Iraq.
2012-2017   Iraqi teachers union, Baghdad, Iraq

Awards and Scholarships

Top 10 selected for DAAD-Fellowship

Laboratory Toolbox

DNA/RNA/protein isolation, DNA electrophoresis, PCR, RT-PCR, bioinformatics tools, primer design, molecular cloning, bacterial transformation, bacterial mutation, nucleic acid quantitation and quality control, SDS-PAGE, western blot, ELISA, flow cytometry, affinity purification, FACS, immunofluorescence, cell isolation/culture, confocal microscopy, fluorescent microscopy, Serialconal Snapgene, EndNote, GraphPad Prism, DNAStar, ImageJ, Photoshop and Illustrator.

Greifswald, …………………   …………………………………...

Murtadha Q. Ali, BSc. MSc.

8.12 Publications and conference contributions

8.12.1 Peer-reviewed articles

1. **<u>Murtadha Q. Ali</u>,** Thomas Peter Kohler, Gerhard Burchhardt, Andreas Wüst, Nadin Henck, Robert Bolsmann, Franziska Voß and Sven Hammerschmidt. Extracellular pneumococcal serine proteases affect nasopharyngeal colonization.
Published in Frontiers in Cellular and Infection Microbiology, 2021.
DOI:10.3389/fcimb.2020.613467

2. **<u>Murtadha Q. Ali</u>**, Thomas P. Kohler, Lukas Schulig, Gerhard Burchhardt, and Sven Hammerschmidt. Pneumococcal Extracellular Serine Proteases: Molecular Analysis and Impact on colonization and disease.
Published in Frontiers in Cellular and Infection Microbiology, 2021
DOI: 10.3389/fcimb.2021.763152

3. Lea Tölken, **<u>Murtadha Q. Ali</u>**, Patience Shumba, Sylvio Redauz, Sven Hammerschmidt, Bernd Kreikemeyer, Steinar Skrede, Anna Norrby-Teglund and Nikolai Siemens. The role of Streptokinase in Group G Streptococcal tissue Infection. (***Manuscript in preparation***).
Front. Molecular Microbiology.

4. **<u>Murtadha Q. Ali</u>**, Thomas Peter Kohler, Gerhard Burchhardt, and Sven Hammerschmidt. The role Extracellular Pneumococcal Serine Proteases on Biofilm Formation. (***Manuscript in preparation***).

8.12.2 International conferences and symposiums attended

1. **Murtadha Q. Ali,** Nadin Mehlitz, Andreas Wüst, Robert Bolsmann, Thomas P. Kohler, Gerhard Burchhardt, Sven Hammerschmidt. Impact of extracellular pneumococcal serine proteases onpneumococcal phagocytosis and virulence. 6th Joint Conference of the DGHM & VAAM-72nd Annual Meeting of the DGHM & Annual Meeting of the VAAM, March, 8 -11, 2020 Leipzig, Germany. (**Poster presentation**).

2. **Murtadha Q. Ali,** Thomas Kohler, Gerhard Burchhardt, Nadine Henck, Andreas Wüst, Robert Bolsmann and Sven Hammerschmidt. The Influence of Extracellular Serine proteases on Pneumococcal persistence in the nasopharynx. 73rd Annual Conference of the German Society for Hygiene and Microbiology (DGHM). Sep 12-14, 2021. (virtual) Germany. **(Poster presentation)**.

3. **Murtadha Q. Ali**, Supradipta De, Thomas P. Kohler, Gerhard Burchhardt, Nadine Henck, Andreas Wüst, Robert Bolsmann and Sven Hammerschmidt. The Impact of Pneumococcal Serine Proteases on *in vivo* colonization. 15th European Meeting on the Molecular Biology of the Pneumococcus (Europneumo2022). 31 May - 03 June 2022, Liverpool, UK. **(Poster presentation)**.

4. Participated in the 4th German Pneumococcal Streptococcal Symposium on October 22nd-24th of Oct. 2018, Freie Universistät Berlin, Germany.

5. Participated in 14th European meeting on the molecular biology of the Pneumococcus from 11 -14th of June 2019, University of Greifswald.

8.13 Acknowledgments

. . . . . *Murtadha*