Proteome-wide analysis of severe bacterial infections. The battle between host and pathogen.

Bakochi, Anahita

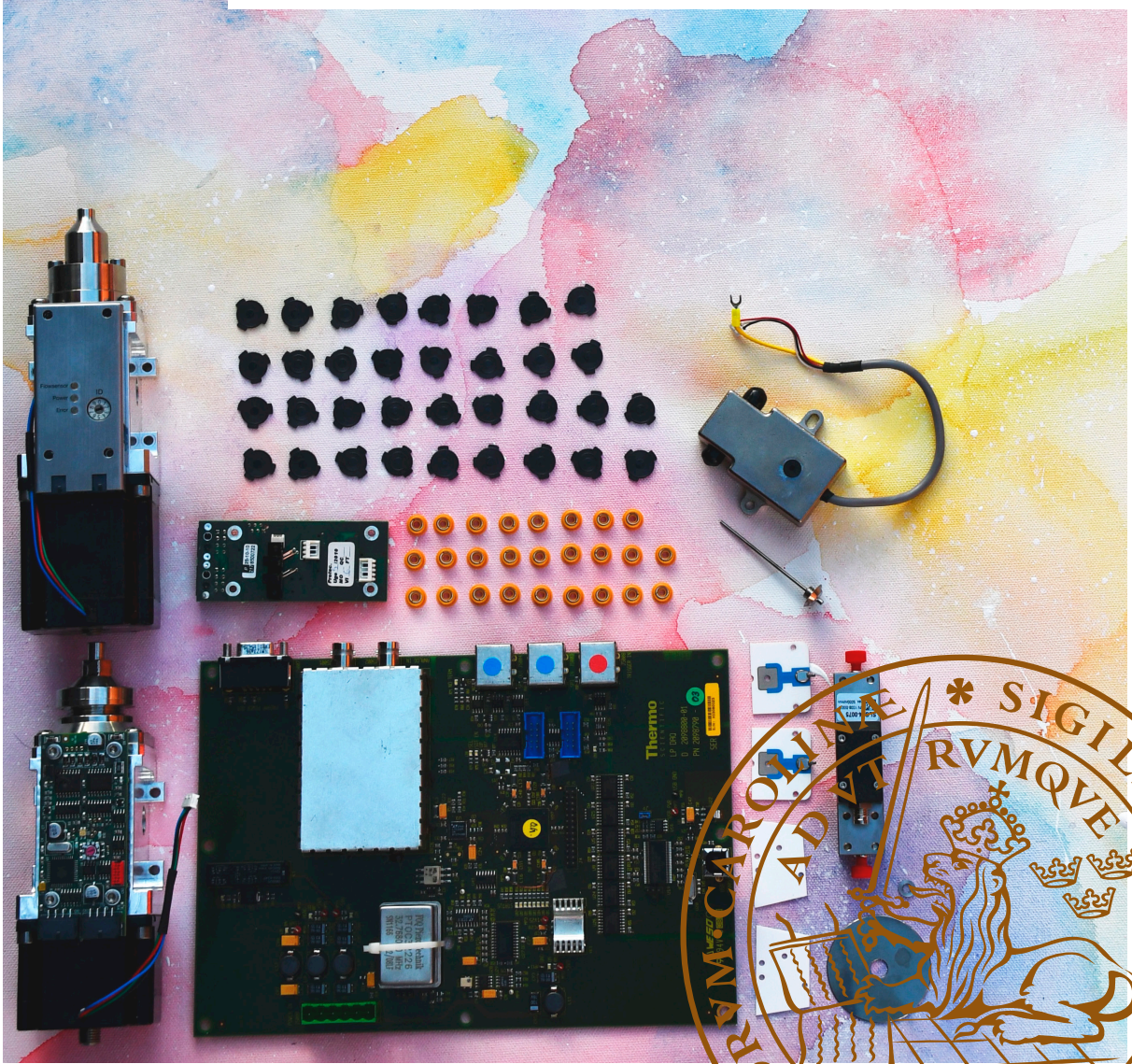
2022

Document Version:
Publisher's PDF, also known as Version of record

Proteome-wide analysis of severe bacterial infections

The battle between host and pathogen

ANAHITA BAKOCHI
DIVISION OF INFECTION MEDICINE | FACULTY OF MEDICINE | LUND UNIVERSITY



Printed by Media-Tryck, Lund 2022 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

FACULTY OF MEDICINE

Department of Clinical Sciences
Division of Infection Medicine

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2022:4
ISBN 978-91-8021-165-9
ISSN 1652-8220

Proteome-wide analysis of severe bacterial infections

The battle between host and pathogen

Proteome-wide analysis of severe bacterial infections

The battle between host and pathogen

Anahita Bakochi

LUND
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at Belfrage lecture hall, Klinikgatan 32, Lund,
Friday January 14th 2022, at 13:00.

Faculty opponent
Dr. Dörte Becher
Institute for Microbiology, University of Greifswald,
Greifswald, Germany

Organization	Document name
LUND UNIVERSITY Division of Infection Medicine Department of Clinical Sciences BMC D13 221 84, Lund, Sweden Author(s) Anahita Bakochi	Doctoral dissertation **Date of issue:** 2021-12-02 Sponsoring organization

Title and subtitle: Proteome-wide analysis of severe bacterial infections: The battle between host and pathogen

Abstract:

The pathogenesis of severe infectious diseases is a complex interplay between the host and the pathogen. The development and progression of a disease encompasses a multitude of processes, which balance between host-damage and host-protection. Proteomic analysis provides the necessary tools to interpret the extensive protein networks of host-pathogen interactions underlying the pathogenesis of a particular disease. This thesis focuses on proteome analysis of two blocks; the host-response during disease, as well as the pathogen during colonization and disease. Data-independent acquisition (DIA) mass spectrometry (MS) was utilized to acquire near-to-complete proteome maps of processes involved in the pathogenesis of one severe infectious disease; meningitis, as well as the host-pathogen interaction counterparts of its leading bacterial cause; *Streptococcus pneumoniae*.

High quantitative ability of DIA-MS was used to construct compendiums of digital cerebrospinal fluid (CSF) proteome maps to define the pathogen-specific host response patterns in meningitis. We generated a predictive multiprotein panel of eighteen human proteins with a high sensitivity and specificity, for discrimination of the meningitis-causing pathogens in the CSF during meningitis. The results also showed a large number of neutrophil-associated proteins in the CSF during bacterial meningitis, and these were found to be due to the presence of neutrophil extracellular traps (NETs). The presence of NETs was further confirmed in the CSF in a rat model of pneumococcal meningitis. Treating the animals with DNase resulted in the abolishment of NETs, and led to increased bacterial killing. We further continued to explore the transcriptional landscape and adaptation of *S. pneumoniae* in human blood plasma by generating a large number of perturbations. A comprehensive pneumococcal proteome repository was constructed to unravel complex protein-protein networks of the bacteria. The results revealed specific regulatory patterns in response to human blood plasma, and pneumococcal transcriptional reorganization regulated by important virulence factors. Furthermore, to describe processes involved in bacterial dissemination in the human nasopharynx, we investigated differences between pneumococcal populations associated with colonization (biofilm bacteria), disease (biofilm-dispersed bacteria) and the conventional broth-grown, planktonic bacteria. The investigated populations showed distinct proteome patterns, especially in regards to metabolic pathways. The virulence of these models was investigated in a murine pneumococcal infection model, where it was showed that virulence of the populations is largely mediated on the investigated pneumococcal serotype.

In conclusion, large-scale proteome analyses produced in this thesis generate fundamental knowledge in understanding host-pathogen interactions as a whole. Furthermore, the constructed repositories can be repetitively queried by the scientific community to deepen the understanding in host-pathogen interactions in bacterial infections.

Key words: Proteomics, mass spectrometry, systems biology, host adaptation, meningitis, bacteria, Streptococcus pneumoniae, pneumococcus

Classification system and/or index terms (if any)

Supplementary bibliographical information		**Language:** English
ISSN and key title: 1652-8220		**ISBN:** 978-91-8021-165-9
Recipient's notes	**Number of pages: 63**	Price
	Security classification	

Signature Anahita Bakochi Date 2021-12-02

Proteome-wide analysis of severe bacterial infections

The battle between host and pathogen

Anahita Bakochi

LUND
UNIVERSITY

Cover photo: "*The number of MS spare parts that an average proteomics PhD student needs in order to successfully complete a PhD*" by Anahita Bakochi

Faculty of Medicine
Department of Clinical Sciences

Doctoral Dissertation Series 2022:4
ISBN 978-91-8021-165-9
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2021

Science knows no country,

because ***knowledge belongs***

to humanity*, and is the torch*

which illuminates the world.

-Louis Pasteur

Table of Contents

Original papers

Paper I

Bakochi, A, Mohanty, T, Pyl, PT, Gueto-Tettay, CA, Malmstrom, L, Linder, A & Malmstrom, J.

Cerebrospinal fluid proteome maps detect pathogen-specific host response patterns in meningitis

Elife, vol. 10. (2021)

Paper II

Mohanty, T, Fisher, J, **Bakochi, A**, Neumann, A, Cardoso, JFP, Karlsson, CAQ, Pavan, C, Lundgaard, I, Nilson, B, Reinstrup, P, Bonnevier, J, Cederberg, D, Malmstrom, J, Bentzer, P & Linder, A.

Neutrophil extracellular traps in the central nervous system hinder bacterial clearance during pneumococcal meningitis

Nat Commun, vol. 10, no. 1, p. 1667. (2019)

Paper III

Bakochi, A, Chao, Y, Scott, A, Gueto-Tettay, CA, Håkansson, AP, Malmström, L, Malmström, J.

Locus driven analysis of the *Streptococcus pneumoniae* proteome organization

Manuscript

Paper IV

Chao, Y, **Bakochi, A**, Bergenfelz, C, Malmström, J, Håkansson, AP.

Proteome profiles of pneumococcal populations associated with colonization and disease

Manuscript

Excluded publications

Malmstrom, L, **Bakochi, A**, Svensson, G, Kilsgard, O, Lantz, H, Petersson, AC, Hauri, S, Karlsson, C & Malmström, J.

Quantitative proteogenomics of human pathogens using DIA-MS

J Proteomics. (2015)

Abbreviations

ABC	ATP-binding cassette
ABM	Acute bacterial meningitis
AUROC	Area under the receiver operating characteristic curve
BBB	Blood-brain barrier
BM	Neuroborreliosis
CBP	Choline-binding protein
COVID-19	Severe coronavirus disease 2019
CSF	Cerebrospinal fluid
DDA	Data-dependent acquisition
DIA	Data-independent acquisition
ESI	Electrospray ionization
FAB	Fatty acid biosynthesis
FDR	False discovery rate
LASSO	Least absolute shrinkage and selection operator (regression model)
LYTA	Autolysin
MPO	Myeloperoxidase
MS	Mass spectrometry
NETs	Neutrophil extracellular traps
PNC	Pneumococcus; *Streptococcus pneumoniae*
PSPA	Pneumococcal surface protein A
PTS	Phosphotransferase system
RP-LC	Reverse phase liquid chromatography
SAH	Subarachnoidal hemorrhage
STKP	Serine/threonine kinase protein
SWATH	Sequential window acquisition of all theoretical mass spectra
TBE	Tick-borne encephalitis
TCS	Two-component regulatory system
VM	Viral meningitis

To my children,

as you are the best experiment that I have ever conducted.

Introduction

Bacteria – in sickness and in health

The concept of the last universal common ancestor of all cells (LUCA)[1] - the evolutionary intermediate of which all life-form on Earth descends from - is widely accepted in the field of early evolution. It gave rise to prokaryotes, unicellular organisms that lack a membrane-enclosed nucleus, which are categorized into bacteria and archaea. This two-domain tree of life later branched out to the evolution of eukaryotes, including animals, plants and fungi.

The first recorded form of life on Earth dates back to 3.5 billion years ago, from when the first fossil records of bacteria were discovered. Introduction of oxygen into the Earth's atmosphere was a major milestone for the evolution of oxygen-breathing life forms, from which the first human ancestors ascend from merely 7 million years ago. Considering that bacteria have existed 500 times longer than humans have, they are now widespread in all of crooks and corners of our planet. One bacterial cell is $1/10^{th}$ of the size of a human cell, typically ranging between 0.5 to 50 μm in size. One way to measure amount of life on earth in quantitative biology is by measuring the amount of carbon; the primary component of all life on Earth (**Figure 1**). Of the ≈550 gigatons of carbon (Gt C) present in our biosphere, bacteria contribute to 13 % of total Gt C in our biosphere.[2] In striking contrast, humans only contribute by a negligible 0,01 % of total Gt C.

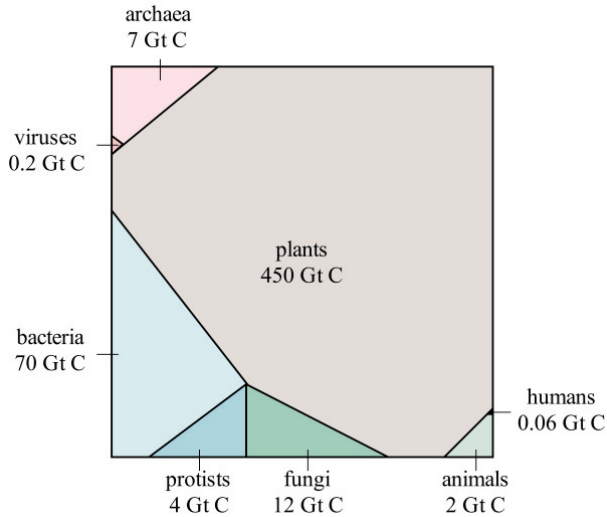


Figure 1. The biomass of the Earth.
An approximate graphical visualization of the global biomass distribution reproduced from Bar-On, et. al, (2018)[2].

A Dutch man named Antonie Philips van Leeuwenhoek was first to discover bacteria nearly four centuries ago. He inspected the scrapings between his teeth under his self-made microscope, and described his discoveries as series of letters to the Royal Society as follows:

> *"... But examining a kind of gritty Matter from between my Teeth, and mixing it sometimes with Rain-water, and sometimes with Spittle (sputum), both which before had no Animalcules (animals), I discovered therein with admiration a great number of very small ones moving."*[3]

What he discovered in fact was bacteria, and little did he know that he would today be recognized as the "Father of Microbiology". Today it is known that the human body harbours an estimated 500 – 1000 different bacterial species[4], and this number could be considerably higher including the different sub-species or genotypes. Previously it has been a general consensus that the number bacteria in the human body outnumber human cells by a ratio of 10:1. However, a recently published study estimates that there are approximately 39 trillion bacterial (B) cells and 30 trillion human (H) cells in the average human body, making the B/H cells ratio in fact closer to 1.3:1[5]. Our bodies harbour in fact nearly equal numbers of human cells as we do bacterial cells.

The composite term for the microbial domains inhabiting the human body is microbiome, or the normal flora, which in addition to bacteria can also encompass archaea, virus and fungi[6]. The formation of each human being's microbiome

begins at birth, and the mode of delivery (vaginal or caesarean) has been consistently reported as one of the most important means by which an infant microbiome takes form[7]. Bacterial species can be found in the skin, mouth, nose, ears, vagina and the gastrointestinal tract in the human body [8], while the majority of the bacteria are found in the GI.

> *"Within one linear centimetre of your lower colon there lives and works more bacteria (about 100 billion) than all humans who have ever been born. Yet many people continue to assert that it is we who are in charge of the world."*
>
> *-Neil deGrasse Tyson*

While the microbiome is necessary for healthy human function, a beneficial species located in the wrong niche can be harmful and cause disease. Bacteria can also colonize the human body, and remain latent without causing disease. These can be activated by external stimuli, such as increased body temperature to febrile range or presence of viral particles, causing dissemination of the bacteria, and leading to a pathogenic outcome to the host.

Streptococcus pneumoniae

Streptococcus pneumoniae, or the pneumococcus, was first discovered in the late 1800's, and it's genetic material led to the major discovery of DNA by Avery, Macleod and McCarty in 1944[9]. Since its discovery, the pneumococcus has remained as a major cause of morbidity and mortality world-wide.

Colonization and infection

The pneumococcus is a Gram-positive, opportunistic, pathogen that colonizes the mucosal surfaces of the nasopharynx in humans. The carriage of pneumococcus is highest in children, with colonization rates up to 65 %, and is decreased to <10 % in the adult population[10]. The pneumococcus behaves largely as a commensal pathogen and the asymptomatic colonization does not cause detrimental effects to the host. Dissemination from the nasopharyngeal tract to other sites leads to various pneumococcal infections (**Figure 2**), and thus stable colonization is considered as a requirement for development of disease.

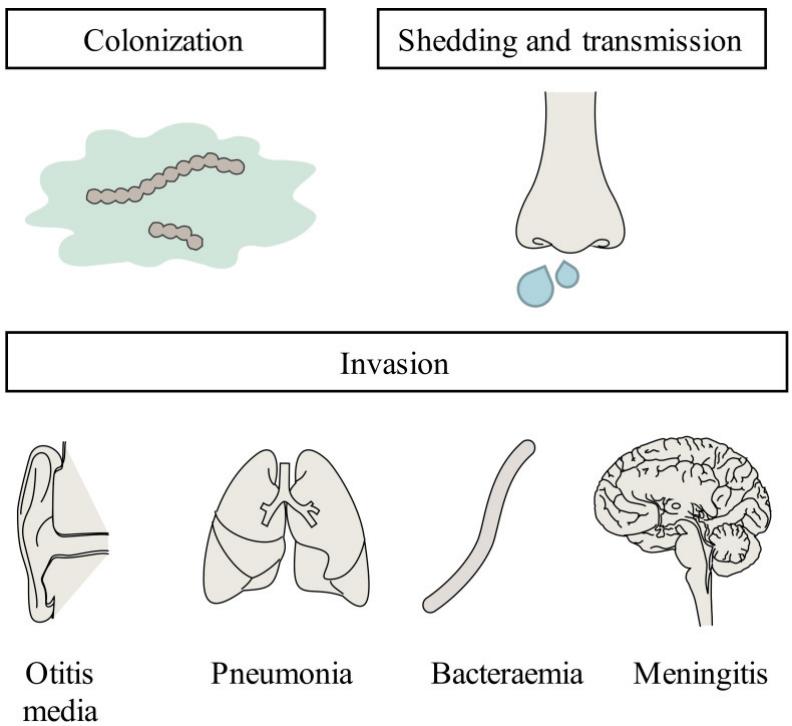


Figure 2. Pneumococcal infections
An overview of pneumococcal localizations in the human body during colonization and infection.

From the mucosal surface of the nasopharynx, the pneumococcus can penetrate into host tissues and progress to cause localized infections at other sites, for example such as sinusitis (sinuses), otitis media (middle ear) and pneumonia (lungs). Invasive infections occur when the pneumococcus enters a sterile site of the body, such as the bloodstream, and can lead to bacteraemia or sepsis. The pneumococcus also harbours molecular mechanisms to transverse across the blood-brain barrier (BBB) where it can cause meningitis.

In 2016, lower respiratory infections caused over 2.3 deaths worldwide, where the pneumococcus contributed to more deaths than other respiratory pathogens such as Haemophilus influenzae type b, influenza and RS (respiratory syncytial virus), put together[11]. In fact, the WHO has included the pneumococcus as a priority pathogen worldwide[12], due to high morbidity and mortality, as well as increasing penicillin and other antibiotic resistance of the bacteria.

Vaccines and serotypes

The development of vaccines has been a major breakthrough in medical history. For the pneumococcus, the vaccines have been targeted against the immunogenic proteins and carbohydrates found on the bacterial surface[13]. These surface carbohydrates make up the pneumococcal surface polysaccharide with unique chemical and immunological properties[14], and is considered as one of the most important pneumococcal virulence factors as it shields the bacteria against the host immune system. Over 90 distinctive capsular polysaccharides have been reported to date[15]. The capsular polysaccharide is used as means to differentiate the pneumococcus into serotypes, and protection by the adaptive immunity is highly serotype-specific.

One of the earliest pneumococcal vaccines, the 23-valent pneumococcal polysaccharide vaccine (PPSV23), was developed in 1980's, and it contained purified polysaccharide antigens protecting against 23 capsular serotypes[16]. Due to the insufficient protective immune response observed in children under the age of two, a new 7-valent pneumococcal polysaccharide conjugate vaccine (PCV7) was developed. Here, the polysaccharide antigens were conjugated to a diphtheria toxin in order to elicit an effective T-cell mediated antibody response, and included seven serotypes common in invasive disease in children. Only a few years after the introduction of the PCV7 vaccine, serotype replacement was observed where serotypes not included in the vaccines increased in the population. More concerningly, the level of antibiotic resistance was also reported to be increased in the non-vaccine serotypes[17]. Several new pneumococcal vaccines have been developed since, such as the 10-valent and 13-valent conjugate vaccines, in order to provide protection for new circulating serotypes not yet covered in previous vaccines.

Virulence

The virulence of a pathogen or a microorganism is defined as the ability to cause damage to the host, and is mediated by virulence factors. The actual disease development is, however, more complex than the sum of all bacterial virulence factors. A successful disease development (**Figure 3**) requires involvement from a) the microbe and its composite virulence factors, b) the host defences and immune system, as well as c) the pre-existing microbial community of the host. In order for a bacteria to cause disease, it has to adapt its virulence in order to enter the host, evade the immune defences and the pre-existing microbial community, and express a number of virulence factors in order to prevail in the human host without being cleared. The successful establishment of a pathogen in its host is a dynamic event, and evolutionary pressure forms the pathogen to adapt to its specific host.

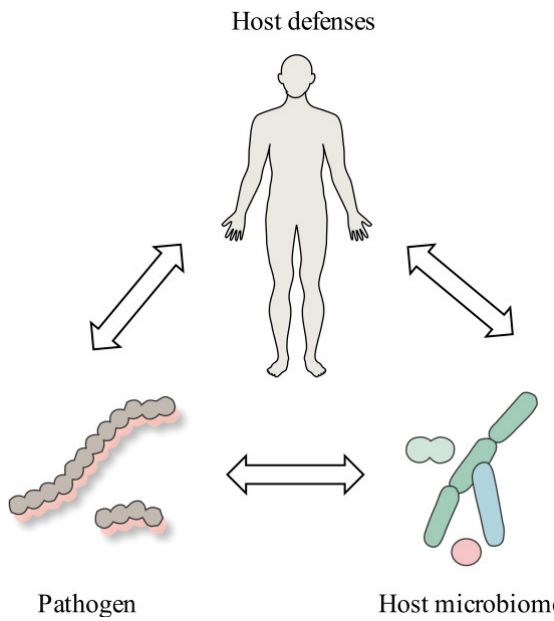


Figure 3. Disease development
Interactions between the invading pathogen, host defenses and the host microbe during disease development.

Capsule

As previously mentioned, the polysaccharide capsule is one of the most important virulence factors of the pneumococcus. In the context of colonization, the capsule inhibits mucus-mediated mechanical entrapment in the nasopharynx, allowing the pneumococcus to access the underlying epithelial surfaces[18]. The thick capsule also has immunological properties by preventing the activation of the components

of both the classical and alternative complement pathways, and therefore prevents phagocytosis[19]. Due to the highly diverse composition of the capsule between serotypes, some serotypes, such as serotype 4 (TIGR4 strain), have been shown to be more effective in preventing complement activation[19]. Capsule is required for the development of invasive disease, as non-encapsulated pneumococci have been shown to be nearly avirulent in murine models, emphasizing the importance of capsule as a major virulence factor in pneumococci.

Autolysis

A unique property of the pneumococcal cell wall is the presence of choline[20], an amino alcohol rarely found in bacteria. This results in the presence of unique group of virulence factors, named as choline-binding proteins (CBP), on the pneumococcal surface. One of them, named autolysin (LytA), is an amidase that cleaves the lactyl-amide bond between cell wall peptides and the peptidoglycan resulting in cellular lysis. The pneumococci are resistant to LytA-mediated lysis during exponential growth phase, but not during stationary growth[21], leading to excessive cell lysis *in vitro.* The function of LytA has not been clarified. However, the importance of LytA as a virulence factor has been shown in murine infection models, where pneumococci deficient in LytA have shown attenuated virulence compared to wildtype strains. It has been suggested that this cell lysis is of importance by releasing other virulence factors, such as pneumolysin, or proteins involved in immune evasion.

Biofilms

In 1970's the link between persistent infection and thick bacterial aggregates was discovered[22]. These aggregates, now known as biofilms, have been since considered as an important aspect of both persistent bacterial infections as well as asymptomatic colonization amongst various bacteria[23]. In biofilms, self-embedded bacteria are attached to a surface in a self-produced matrix consisting of proteins, polysaccharide and DNA. They serve multiple beneficial purposes for the survival of a bacterial species. Bacteria are able to stay latent and hidden from the host immune system, and are even able incorporate host structures to prevent discovery[24]. Biofilm bacteria also exhibit altered metabolic state, gene and protein expression, promoting survival in challenging environments.

Pneumococci use biofilm formation to reside latently and asymptomatically in the nasopharynx during colonization, and the high rate of colonization can largely be contributed by pneumococcal biofilm formation. Clinical observations have shown that eradication of pneumococcal colonization is much more challenging than clearing an infection[25], consistent with the fact that biofilms contribute to resistance to antibiotic treatments. Biofilms have therefore been proposed as an important virulence factor in pneumococci[26].

Meningitis

Pathophysiology

The central-nervous system (CNS) and the brain are considered sterile environments, and the influx and efflux of substances to this body part is carefully controlled. The brain is covered by three meningeal membranes; the pia mater, arachnoid mater and dura mater (**Figure 4**), and the space between the pia and arachnoid matter, the subarachnoid mater, is filled with a clear liquid known as the cerebrospinal fluid (CSF).

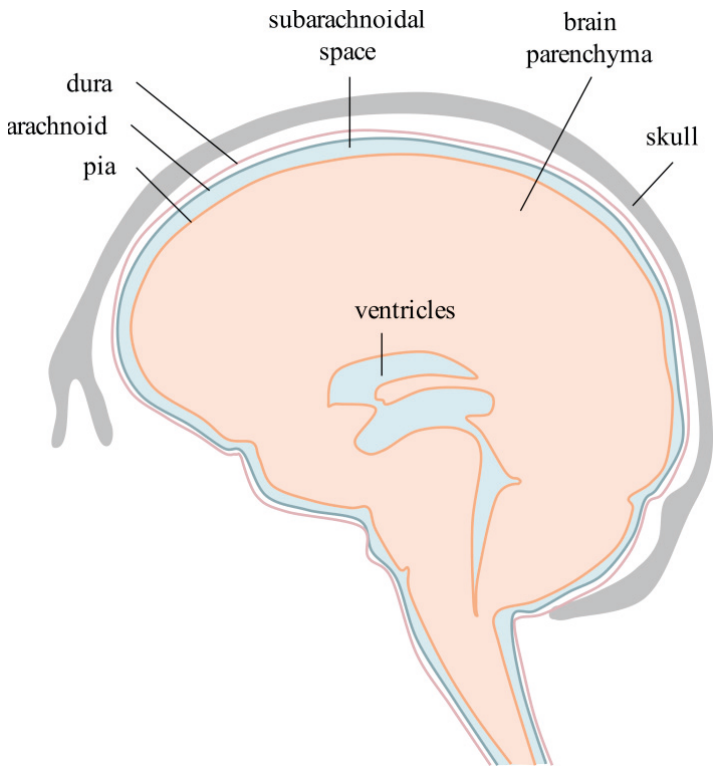


Figure 4. A sagittal cross-section of the brain.
Schematics showing the meningeal membranes (dura, arachnoid and pia) of the brain, as well as the flow of CSF (in light blue) through the ventricles into the subarachnoidal space.

The infection of the meninges of the brain is known as meningitis, and is a severe and life-threatening disease. The characteristic symptoms of meningitis are headache, fever and neck stiffness[27], and consequent neurological and behavioural symptoms can arise with inflammation or complications on the brain parenchyma

or the vascular system. Most commonly, pathogens infect the meninges by spreading from one infection site into the bloodstream, and then transverse through the blood-brain barrier (BBB) to enter the CSF[28]. However, blood-borne pathogens are not the only source of meningitis. Some pathogens can also enter the CSF from a near-site infection, such as otitis media, or by transmigration into the olfactory nerves of the nose and relocate into the CNS[29,30].

Causes

The acute form of meningitis is a heterogenous disease, and can be caused by a wide range of infectious agents[31]. Viruses are the most common cause of meningitis, and are generally associated with low mortality rates[32]. Some subtypes of viral meningitis (VM) are associated with higher mortality rates and the risk of developing long-term neurological sequelae, such as tick-borne encephalitis (TBE)[33]. In contrast, the most severe form of meningitis is caused by bacteria (acute bacterial meningitis; ABM). ABM is one of the leading causes of death worldwide due to infectious diseases, and is associated with an increased risk of development of long-term neurological sequelae as well as a rapid disease progression[27,31].

Bacterial meningitis can separated into two categories; hospital-associated (or post-operative) and community-acquired. In the former case, the infection is a post-operative or a secondary event following head trauma, and in the latter case the infection is acquired spontaneously in the community. The most common causes in community-acquired meningitis are *Haemophilus influenzae* type b, *Streptococcus pneumoniae* and *Neisseria meningitidis*. Mortality rates vary greatly even between the bacterial species, where it is highest in pneumococcal meningitis (between 20 and 30% in high-income countries, and as high as 50% in low-income countries)[34], in contrast to 3 – 10% in meningococcal meningitis[35,36].

Due to the large number of causes of meningitis, it can be in fact considered as an "umbrella term", rather than a singular disease. The disease consists of considerable heterogeneity, which influences both the disease progression and the subsequent clinical diagnosis and treatment.

Therapeutics and challenges in bacterial meningitis

While the symptoms of meningitis are the same, the optimal treatment and care is highly specific to the causative agent. But the diagnosis of meningitis and the exact underlying microbial cause has been challenging. Typically, CSF is collected by lumbar puncture, and various clinical parameters are evaluated. The identification of bacterial pathogens is determined by CSF cultures, and the results can take days to generate. Delays in the administration of antibiotic treatment has

been correlated to poor outcome in patients suffering from ABM, reflecting the crucial need of fast diagnostic methods. In fact, delays as short as six hours between arrival to the emergency room and the administration of antibiotics has been, in some cases, associated with increased mortality in ABM[37]. Timing is not the only challenge: the choice of antibiotic is limited by the ability to cross the BBB into the CSF in sufficient concentrations, as well as increasing antibiotic resistance in bacterial species. The early administration of antibiotic treatment is based on the patient's age and clinical manifestations, and often combines vancomycin and a cephalosporins[31]. The exact antibiotics regime is later adjusted as the results from culturing and susceptibilities have completed.

Even with adequate early administration of antibiotics, the mortality rates of ABM remain high, which can be correlated to irreversible damage to the brain tissue. For example, the administration of "lytic antibiotics", such as cephalosporins, may result in release of bacterial inflammatory and cytotoxic factors in the surrounding environment. In addition to clearing the bacterial infection, the antibiotic treatment may therefore result in harmful side effects. This may be one explanation as to why mortality and the development of neurological sequelae remain high in pneumococcal meningitis, even with sufficient antibiotic treatment.

NETs and DNase

In the onset of bacterial infections, neutrophils are one of the first immune cells that are actively recruited at the site of infection. During normal conditions, the BBB remains impermeable to neutrophils and neutrophils are kept at bay from the central nervous system. However, massive neutrophil infiltrations have been shown during ABM[38]. One mechanism of extracellular killing mediated by neutrophils is the process of NETosis, or NETs[39]. Here, neutrophil granule proteins together with DNA are expelled from the cell, forming extensive and sticky structures for capturing and immobilizing bacteria[39]. The NETosis is considered as a "double-edged sword", as they are both beneficial and harmful to the host[40].

The first reported correlation between NETs and bacterial meningitis dates back to 1959, where administration of DNase together with penicillin was shown to considerably improve mortality rates during ABM[41]. While the researchers at the time did not know the molecular mechanisms of this finding, they speculated that the degradation of DNA may have exposed the bacteria to host defences or antimicrobial therapy. It is now thought to be due to the enzymatic activity of DNase degrading extracellular DNA of NETs.

The safety of DNase as therapeutics is well established, and DNase has been repurposed for various number of diseases. For example, aerosolized recombinant human DNase (under the brand name "Pulmozyme") is approved as a treatment for cystic fibrosis in humans, where extracellular DNA increases viscoelasticity in

airway secretions[42]. More recently, high levels of NETs have been discovered in the sputum of patients suffering from severe COVID-19 infection, and treatment with Pulmozyme improved both oxygen saturation and recovery[43]. Although the use of DNase in meningitis in humans is still largely unexplored, if deemed to be beneficial in treating bacterial meningitis, it would be fairly fast to bring it to a clinical setting.

Proteomics

The proteome

The term "proteome" was first introduced in 1994 by Marc Wilkins[44], where the authors reflected on the challenges of studying the complete set of proteins expressed by a genome as follows:

> *"A great challenge that biologists now face is how the co-expression of thousands of genes can best be examined under physiological and pathophysiological conditions, and how these patterns of expression define an organism."*

Not more than two decades later, the first drafts of the human proteome were published[45,46], and since then the field of proteomics has grown explosively. Today, the proteome is defined as the entire set of proteins expressed in a cell, organ, or organism, and includes all isoform and post-translational variants. Gene products – proteins - exert their functions by complex chains of protein interactions. Thus, the understanding of the proteomic organization on a holistic level is crucial in order to understand processes involved during both homeostasis as well as disease.

As pathogens and humans have evolved alongside each other, both parties have developed extensive mechanisms for adaptation and defence. Extensive research must be conducted in order to understand the mechanisms that lead to clearance of the pathogen, colonization or disease. The fundamental biology in a living organism is driven by proteins, and one way that research often is conducted is to elucidate the role of a single individual protein from all the rest. This results in valuable information of the function of that said protein. In any organism, however, proteins are involved in countless of interactions, both within the organism and in host-pathogen interactions. This type of research is alone not enough to interpret the pathophysiology of a disease. In order to understand the full complexity of host-pathogen interactions, including all of the interacting parts of proteins, quantitative models representing the whole system are required.

Mass spectrometry-based proteomics

One way to measure the proteome is by the means of mass spectrometry (MS), which, in simple, is a technique for measuring the mass-to-charge (m/z) ratio of ions. During the past decades, mass spectrometry has become a powerful tool for the analysis of protein samples, and various techniques have been developed since. MS-based proteomics analyses can be split into two major fields; top-down and bottom-up, where the former one consists of the analysis of intact proteins, and the

latter one of analysis of proteins digested into peptides[47,48]. One would assume that the best approach would be to directly analyse intact proteins from a protein sample, but the successful measurement of intact proteins is limited by the instrument's mass analysers, which perform better at measuring smaller units. Most biological samples consist of a rich mixture of proteins, which has resulted in that bottom-up proteomics has become the most commonly used method for analysis of complex samples. In this thesis, bottom-up proteomics and LC-MS/MS (liquid chromatography tandem mass spectrometry) has been uniformly used across all the papers as a means of proteome analysis.

Prior to entering the mass spectrometer

In a typical LC-MS/MS workflow (**Figure 5**), the protein sample is first cleaved into peptides by enzymatic digestion. The most common enzyme to use is trypsin[49], a specific enzyme generating easily ionized peptides of a feasible length, both which are favourable features for using in MS-analysis. The peptides then undergo separation, e.g., by the means of reverse phase liquid chromatography (RP-LC) – often directly coupled to the mass spectrometer (on-line). Here, the peptides are loaded onto a matrix column, and subsequently eluted based on the peptides hydrophobic properties by washing the column with a pre-specified gradient of hydrophobic/hydrophilic solvent[50]. The separation is important due to the limitations of mass spectrometers in analysing large numbers of analytes, and the resulting retention time information can further aid in identification of peptides in the sample. As the peptides reach the tip of the chromatographic column, they are ionized by electrospray ionization (ESI). In fact, the 2002 Nobel Prize in Chemistry was awarded for the use of ESI with mass spectrometry[51]. The ionized peptides then enter the mass spectrometer.

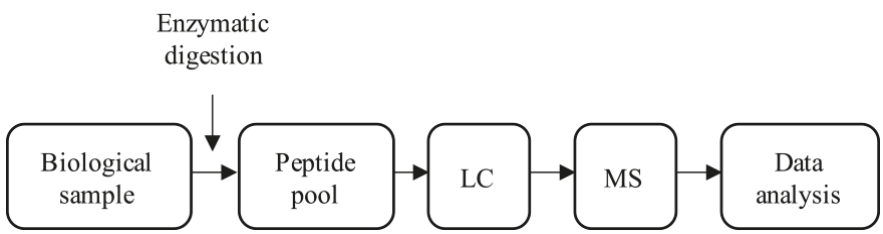


Figure 5. A schematic overview of the main steps of MS-analysis

Data-dependent acquisition

In data-dependent acquisition (DDA; shotgun MS), the aim is in the detection of as many peptides as possible, while concurrently resulting in low sensitivity and poor reproducibility. In DDA (**Figure 6, left panel**), the mass-to-charge ratios and the abundances of the injected peptides are scanned by the mass analyser in full (MS1) scans. The MS1 spectrum does not provide enough information for confident peptide sequence identification, as there are typically a large number of possible amino acid sequence combinations for each peak. Due to this, the top *n*

most intense peptide ions (precursor-ions) are then selected for subsequent fragmentation in the collision cell, where precursor ions fragment by collision with an inert gas. In this thesis, higher-energy collisional dissociation (HCD)[52] was used for fragmentation. These fragment ions are then scanned in a second mass analyser and subsequently detected in order to generate MS2 scans. In order to avoid the repeated measurement of the same precursor ion, a dynamic exclusion[53] list is typically set in the instrument, resulting in constant measurement of different ions. The quantification in DDA is attained at MS1 (peptide) level, which is more rudimentary quantification method than for example DIA (see below).

Data-independent acquisition

The data-independent acquisition (DIA, **Figure 6, right panel**) method known as SWATH-MS[54] (Sequential Window Acquisition of all Theoretical Mass Spectra) was popularized in 2012 and changed the field of mass spectrometry altogether. SWATH-MS generates an unbiased, digital archive of the whole ionizable peptide content of the sample, therefore avoiding the limitations of DDA methods for high-abundant or stochastic selection of peptides.

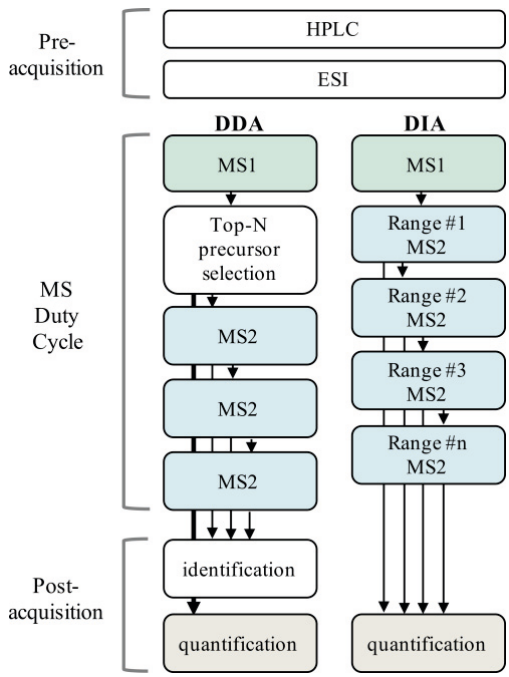


Figure 6. An overview of the DDA and DIA mass spectrometry proteomics workflows.
In pre-acquisition, peptides are separated by high performance liquid chromatography (HPLC), and ionized with electrospray ionization (ESI). In data-dependent acquisition (DDA), or shotgun, workflow the top-N precursor peaks are selected from the MS1 for subsequent fragmentation. In data-independent acquisition (DIA), all precursors are fragmented in predetermined ranges.

Here, the mass spectrometer repeatedly cycles of performing MS1 scans at predefined m/z isolation windows, typically 25 Da, covering the range of 400 – 1,200 m/z throughout the chromatographic retention time (**Figure 7**). During each isolation window, all precursors are subsequently isolated, fragmentated and scanned, resulting in the generation of a large number of MS2. At each cycle time, the whole range is scanned anew, producing a time-resolved and complete recording of fragment ions within the biological sample.

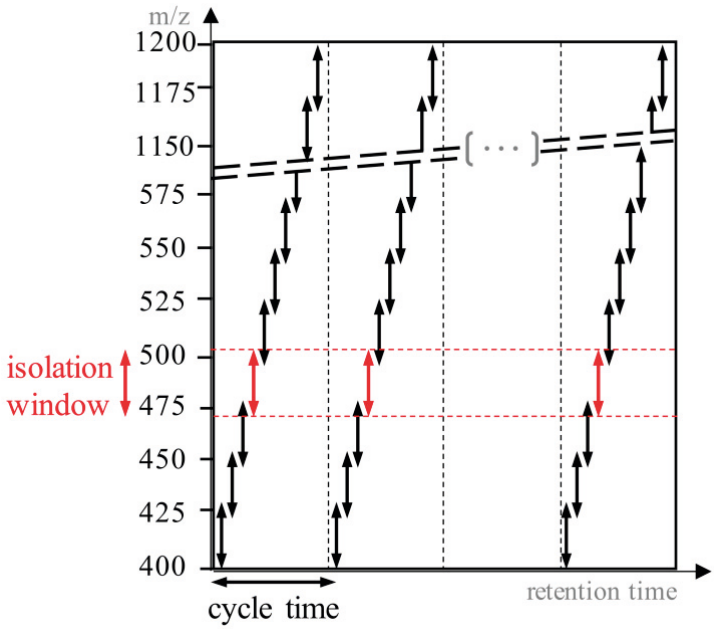


Figure 7. Schematic overview of the SWATH-MS method.
Reproduced from Gillet et al., (2012) 'Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis', Mol Cell Proteomics, vol. 11, no. 6, p. O111 016717.

Data analysis and challenges

The annotation of generated mass spectra into peptides is done computationally, and in the case of DDA, this is achieved by matching the recorded spectra to the *in silico* digested theoretical proteome. For this, the genome of the analysed organism naturally requires to have been sequenced[55]. By using the given genome sequence and the digestion enzyme used, it can be predicted which peptides could technically be present in the sample. From this, a theoretical database of predicted MS2 fragmentation spectra is generated, which is then used to compare to the experimental MS2 spectra for annotation of peptide sequences.

As this method relies on comparing predicted spectra to experimental spectra, statistical approaches are used for validation. This is typically done by determination of the false discovery rate (FDR)[56]. A new database is generated, where the amino acid sequences of all predicted proteins are either shuffled or reversed, resulting in decoy proteins[56]. The generated MS-data is then searched against both the theoretical proteome database, as well as the decoy database. Experimental spectra should match to a decoy much worse, and only by random. A score is then calculated based on whether the experimental spectra is "a true match" or "a decoy match". Commonly, a cut-off score is placed at 1% FDR[56], meaning that 1% false identifications are accepted.

In DIA, the generated mass spectra are much more complex, as it contains spectra of all precursors within each isolation window. For identification and quantification of DIA-data, the deconvolution of these spectra is achieved by assays (assay libraries)[57]. Assay libraries are typically generated from previous DDA analyses, and provide information regarding the protein content in the analysed sample[57]. Often, a DIA analysis is preceded by multiple DDA analyses, in order to generate an experiment-specific assay library. Due to the stochastic nature of precursor selection in DDA, only peptides detected in DDA will be included in the assay library. Consequently, the identification and the subsequent quantification of peptides from a DIA analysis is limited by the assay library. Also, different or pre-existing assay libraries can be used to repeatedly query DIA-generated data. This is specifically beneficial when working with clinical samples[58] (which are often scarce and of limited volume), as only one DIA analysis is sufficient for generation of full MS spectra of the sample. An additional advantage to DDA is that the quantification in DIA is attained at MS2 (fragment) level, yielding more precise quantification of peptides than in DDA.

Certain considerations need to be taken into account as the mass spectrometer measures peptides, and not proteins. For example, mapping a peptide to a protein is straight-forward if all peptides would be unique to their corresponding proteins. This is however not always the case. Certain peptides, and even proteins to a certain degree, are evolutionarily conserved, even across species. Peptides matching to multiple proteins, non-proteotypic peptides, give little information regarding which protein was in fact in the biological sample. It is therefore common to consider proteotypic peptides only in a MS-analysis.

Aim of this thesis

The red thread

The aim of this thesis was to use **advanced quantitative capabilities of mass spectrometry** to study the **host-pathogen interactions** of a severe infectious disease, meningitis. The pathogenesis of any disease is a complex interplay between the host and the pathogen, a constant battle between host-damage and host-protection. Therefore, this thesis was divided into two blocks (**Figure 8**).

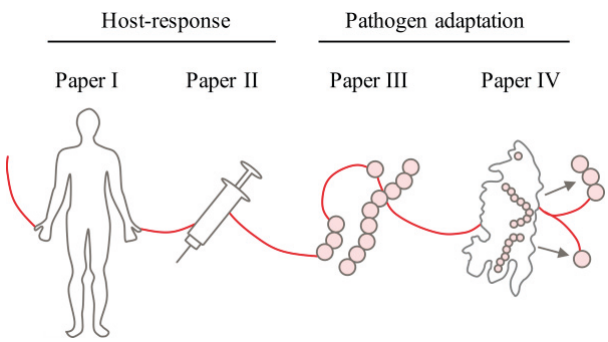


Figure 8. The red thread of this thesis.
The papers in this thesis are divided into two blocks: host response during disease (Papers I and II) and pathogen adaptation (Papers III and IV).

In the **first block**, the aim was to investigate how the host proteins in CSF are regulated in response to different meningitis-causing pathogens. Furthermore, the role of a specific host-response mediated by neutrophils, namely NETs, was studied in the most severe form of meningitis, acute bacterial meningitis (ABM).

The **second block** concentrates on investigating the *Streptococcus pneumoniae* – the most common cause of ABM. Here, the aim was to map out the transcriptional landscape of pneumococci, and discover how the bacteria regulates its proteome during various perturbations. Also, the differences in the proteomes of pneumococcal populations associated with colonization and invasive disease was examined.

Specific aims

Paper I:

- Develop a compendium of DIA-MS cerebrospinal fluid proteome maps
- Provide novel insights into central nervous system functioning and host-response, as well as discover pathogen-specific host-response patterns in a cohort of patients with meningitis.

Paper II:

- Investigate the presence of NETs in CSF during acute bacterial meningitis
- Determine the effect of NET removal by DNase on bacterial killing in the CNS

Paper III:

- Develop a large repository of pneumococcal proteome maps by using DIA-MS
- Investigate the proteome composition and regulation patterns of pneumococci under different conditions
- Define the composition of stable, fluctuating and highly variable fraction of pneumococcal proteome

Paper IV:

- Investigate the differences in the proteome compositions and virulence of pneumococcal populations associated with
    - *colonization* (pneumococci in biofilms)
    - *disease* (dispersed pneumococci)
    - *conventionally grown pneumococci* (broth-grown; planktonic)

Results

In this section, I present a brief summary of the main results of each paper. The papers in original format can be found at the end of this thesis, where the reader can find an in-depth presentation of the results.

Paper I

Changes in the proteome pattern in CSF during meningitis

In this paper, clinical CSF samples were obtained from a cohort of patients admitted to the hospital with the suspicion of meningitis. This large cohort (**Figure 9**) included a total of 107 patients suffering from ABM, neuroborreliosis (BM), VM (including tick-borne encephalitis; TBE), suspected ABM or VM, inflammation without infection, as well as healthy (headache) controls.

	Pathogen	N	Gender – n (%) male	Age (years) – Mean
ACUTE BACTERIAL MENINGITIS	–	35	16 (47.1)	58
Community–acquired	Listeria Monocytogenes	2	1 (50)	79
	Neisseria Meningitidis	5	3 (60)	25
	Streptococcus Pneumoniae	17	7 (43.8)	63
	Streptococcus Pyogenes	2	1 (50)	72
	Pseudomonas Aeruginosa	1	0 (0)	59
Post–operative	Bacteroides Fragilis	1	1 (100)	97
	Enterococcus Faecalis	1	0 (0)	33
	Escherichia Coli	1	1 (100)	73
	Staphylococcus Aureus	3	1 (33.3)	68
	Streptococcus Agalactiae	1	1 (100)	1
	Streptococcus Salivarius	1	0 (0)	63
VIRAL MENINGITIS	–	21	10 (50)	45
	Cytomegalovirus	1	0 (0)	22
	Enterovirus	5	2 (50)	26
	Herpes Simplex Virus 1	2	1 (50)	61
	Herpes Simplex Virus 2	4	1 (25)	38
	Herpes Zoster Ophthalmicus	1	1 (100)	83
	Tick–borne Encephalitis	5	3 (60)	53
	Varicella Zoster Virus	3	2 (66.7)	51
NEUROBORRELIOSIS	–	7	3 (42.9)	63
HEADACHE CONTROLS	–	49	20 (40.8)	48

Figure 9. CSF samples included in this paper.
A summary, showing the number of clinical CSF samples, as well as the pathogen strains that caused meningitis.

We could show that the largest effect on the human CSF proteome was mediated by ABM, where both the number of uniquely detected proteins, as well as the total protein intensities showed highest increase compared to the other groups. The majority of the significantly regulated proteins were associated with neutrophils (32% in ABM and 14% in VM), the brain (44% in ABM and 32% in VM), as well as plasma-derived proteins, including several acute-phase proteins.

❖

In fact, deep analysis of the neutrophil-derived proteins in ABM revealed numerous proteins involved NET formation. This implied the presence of NETs within the CSF during bacterial meningitis. This finding lead to a spin-off project, where NETs were investigated in the CSF during meningitis. This project was pursued in line with Paper I, and published in Nature Communications (in this thesis as Paper II).

❖

Only three proteins were similarly, and statistically, regulated in all three disease groups (ABM, BM and VM); chitinase-3-like protein 2 (CHI3L2), immunoglobulin mu heavy chain disease protein (MUCB) and profilin-1 (PROF1). This demonstrates the heterogeneity of meningitis as a singular disease, signifying the vast differences in CSF proteome caused by the plethora of infectious agents capable of causing meningitis.

LASSO regression modelling for detecting predictive proteome patterns

We used LASSO (least absolute shrinkage and selection operator) regression model as a means of machine learning to generate predictive models to discriminate between the samples (a summary of the LASSO model is shown in **Figure 10**). Five sample groups were used here (ABM, BM, VM, TBE and control), and each model consisted of predictive scores for proteins used for discriminating between these sample groups.

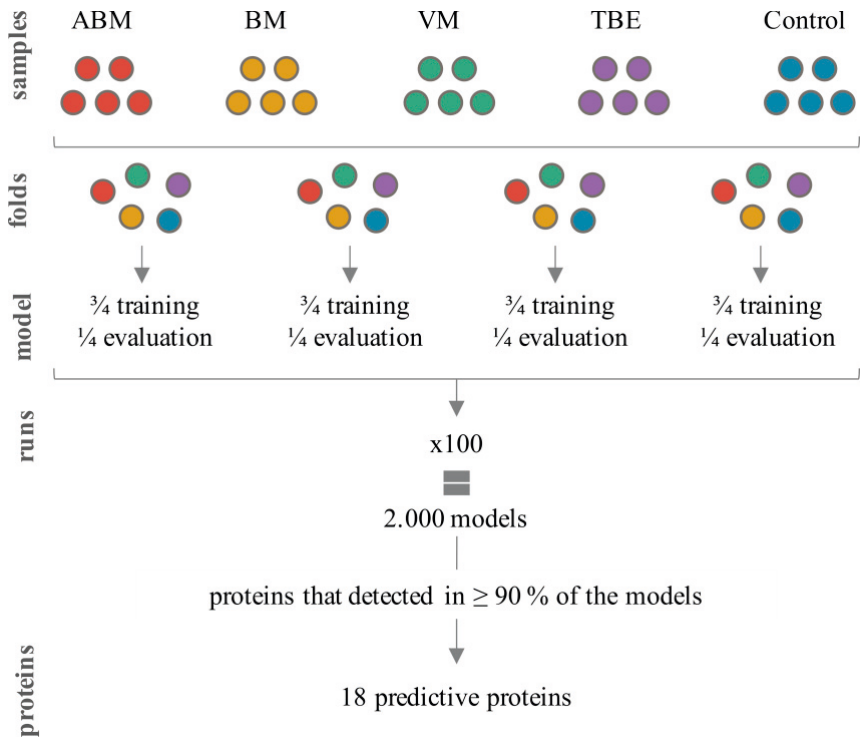


Figure 10. Overview of the LASSO regression model.
SAMPLES: The sample groups consisted of five groups (ABM, BM, VM, TBE and control), each containing differing number of clinical samples. **FOLDS**: The samples were split into four folds by random, where each fold contained the same proportion of pathogen-specific samples as the whole cohort. **MODEL**: ¾ of each fold was used to train the regression modeling on. The remaining ¼ was used to evaluate the performance of the model. **RUNS**: Each model was performed 100 times, resulting in a total of 2,000 models (5 sample groups x 4 folds x 100 runs). **PROTEINS**: All proteins included in this study did not display a predictive ability in the generated models. However, those that did, and which were detected in ≥ 90 % of the 2000 models, were selected as predictive proteins.

Our predictive models generated discriminative protein profiles for all sample groups, with an average of > 80 % area under the receiver operating characteristic curve (AUROC). The model performed best for prediction of ABM and controls (mean AUROC 96 % and 95 %, respectively). TBE displayed a lower discriminatory AUROC at 87 %, but this was, however, the most stable across the modelling.

Multi-protein panel for discrimination of causative-agent subgroups

From the LASSO regression model, we were able to select 18 proteins with high predictive ability to differentiate between the five sample groups. Out of these proteins, five were elevated in ABM and BM, predicting ABM or BM from all other classes. Nine predictive proteins were elevated to a high degree in ABM, and to a lesser extent in controls, suggesting that a high concentration of these proteins

is necessary for indication of ABM. For prediction of VM or TBE, one discriminatory protein was detected, respectively.

Neurological sequelae: the longitudinal effect

As the development of neurological sequelae is a serious and debilitating possible long-term effect of meningitis, we analysed longitudinally collected CSF samples from patients suffering from ABM or a non-infectious brain trauma, subarachnoidal hemorrhage (SAH). For ABM, our analyses show that certain proteins remain elevated (LASSO-predicted proteins for ABM), whilst certain down-regulated proteins consistently increase, or even peak (brain-associated proteins), during the investigated 10-day time period. The levels of neutrophil-associated proteins that were statistically and specifically up-regulated in ABM show a tendency for decrease, but are also accompanied by a high degree of variability amongst the patients.

Paper II

In **Paper I**, The CSF samples during ABM displayed significant increase in neutrophil-associated proteins. In addition to this, NET-associated proteins were detected, suggesting for the first time the presence of NETs in CSF during ABM. The presence, and role of these NETs during CNS infection was further investigated in this paper.

Presence of NETs in human CSF during pneumococcal meningitis

Examination of the CSF samples from patients afflicted with pneumococcal ABM (which will hereafter be referred to as simply ABM in this section), BM, VM, and SAH showed that specifically ABM samples contain considerably higher neutrophil counts and neutrophil-associated proteins than other sample groups. Furthermore, for the first time, we showed that NETs are present in the CSF of ABM patients, and the levels of NETs were low or undetectable in the other conditions investigated. In vitro analysis of various clinical isolates of meningitis-causing bacterial species challenged with purified human neutrophils exhibited a significant and rapid NET formation, indicating that ABM-causing bacteria are potent NET-inducers.

DNase treatment in a rat model of pneumococcal meningitis

Additionally, NETs were also formed in a rat model of pneumococcal meningitis, where suspensions of a potent NET-inducing pneumococcal isolate were subarachnoidally injected. In this model, the NET formation was successfully prevented with a treatment of a recombinant human DNase I (DNase). The treatment with DNase also reduced the number of viable bacteria in the brain, as well as in other organs, and thus resolved the infection and prevented dissemination to distant organs. Both intrathecal as well as intravenous administration of DNase led to the reduced bacterial counts in the investigated organs, meaning that DNase is able to cross the BBB.

NETs promote bacterial survival

Treatment with DNase and the consequential removal of NETs led to increased intracellular uptake of bacteria in neutrophils, when compared to control-treated samples. Additionally, the activity of myeloperoxidase (MPO) was significantly increased in the supernatant following DNase treatment. Combined, these findings showed that DNase-mediated NETs removal results in increased neutrophil-mediated phagocytosis of bacteria, as well as increased extracellular killing mechanisms, leading to enhanced bacterial clearance in pneumococcal meningitis.

Paper III

A total of 176 pneumococcal samples and a total of 30 unique conditions, including various isogenic mutants and growth conditions were generated (**Figure 11**), in order to investigate how the specific modifications alter the transcriptional landscape of the pneumococcus. The selected mutants included virulence factors of different classes, including choline-binding proteins, lipoproteins, secreted proteins, the regulatory protein serine/threonine kinase, as well as capsular alterations (capsule mutants and hypercapsulated strains).

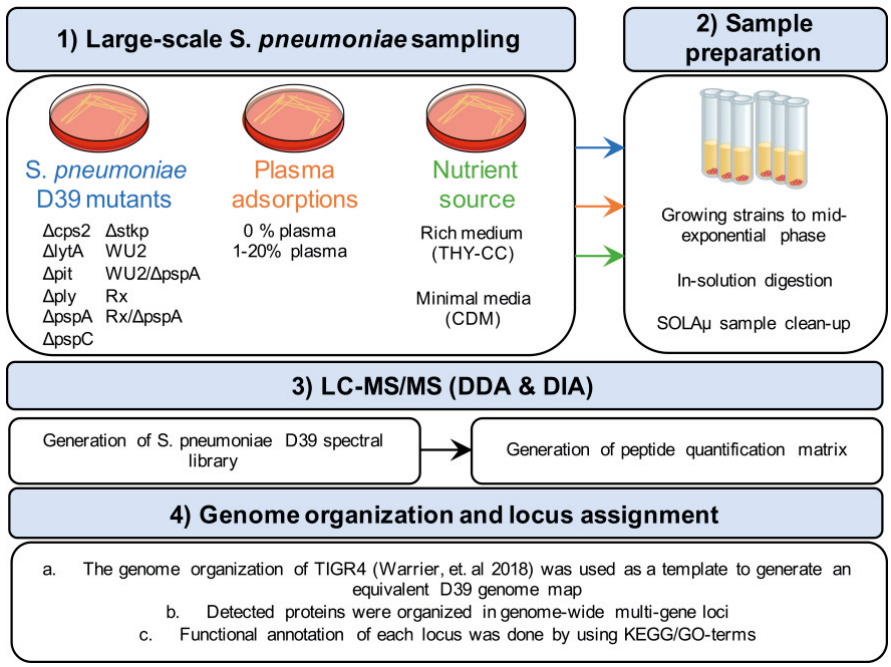


Figure 11. Graphical abstract of the project workflow.
Wild-type *Streptococcus pneumoniae* D39 strain and 11 mutants were grown in various conditions and prepared for analysis by mass spectrometry. A large pneumococcal spectral library was generated, and the samples were ran on DIA for generation of peptide quantification matrix. Based on the genome organization of the TIGR4 strain as published in Warrier, et. al., 2018, an equivalent D39 genome map was generated. The detected proteins in this paper were organized in line with the generated D39 genome map, and grouped as multi-gene (and consequently protein) loci. Each locus, containing multiple proteins, was annotated into functional groups by using KEGG and GO-terms. csp2; capsule biosynthesis protein, lytA; autolysin, pit; pneumococcal iron transporter, ply; pneumolysin, pspA and pspC; pneumococcal surface protein A and C, stkp; serine/threonine kinase protein, WU2; hypercapsulated pneumococcal strain, Rx; spontaneous capsule mutant strain.

Genetic locus-based functional annotation of the proteome

Bacterial genomes are organized in multi-gene operons – here termed as loci (sing. locus). Commonly, the genes within a locus are involved and share the same

biological function. In this paper, we organized the detected pneumococcal D39 strain proteins into loci consisting of multiple proteins, based on a genome-map generated from Warrier, et. al., (2018)[59]. Here, we measured a total of 14,239 peptides, matching to a total of 1,336 unique proteins – resulting in a high theoretical D39 strain proteome coverage of approximately 70 %. We assigned these proteins into 297 loci, which were biologically annotated into 30 unique functional groups.

Pneumococcal proteome composition in homeostasis

By analysing the proteome composition, we showed that in the wild-type pneumococcus – reflecting the natural state of pneumococci – the highly expressed (high abundant) functional groups consist of ribosomal proteins, ATP-binding cassette (ABC) transporters and carbohydrate metabolism proteins. In contrast, proteins involved in peptidoglycan/capsule biosynthesis, quorum sensing and protein transport reflected the total protein abundances the least.

Plasma -induced proteome reorganization

Human blood plasma is rich in nutrients, as well as a reservoir of large number of immune system factors, and represents to the environment that the bacteria would encounter in the human host. Pathogens encountering human plasma must adapt to the environment, in order to evade clearance. The interactions of the pathogen with plasma proteins provides valuable information regarding the adaptation of pathogens in humans during infection.

Pneumococci subjected to human blood plasma significantly down-regulate five proteins involved in fatty acid biosynthesis, for which only one locus was identified in the pneumococcal genome. The down-regulation fatty acid biosynthesis proteins suggests that the bacteria are able to utilize plasma as a resource for fatty acids, and are therefore able to shut down the internal fatty acid biosynthesis machinery.

Effect of the isogenic mutants on the pneumococcal transcriptional landscape

Of all the mutants investigated in this paper, we saw the largest effect on the proteome in ΔlytA (autolysin), Δcps2 (capsule biosynthesis protein), and Δstkp (serine/threonine kinase protein) samples. The Δstkp samples regulated 17 % of the total pneumococcal theoretical proteome, up-regulating proteins associated with chromosomal processes and ion binding/transport, and down-regulating peptidoglycan biosynthesis, phosphotransferase systems (PTS), quorum sensing and translation. Furthermore, two two-component regulatory systems displayed significant Δstkp mediated regulation, where saeRS (tcs08) and vncRS (tcs10) were up- and down-regulated, respectively.

Another isogenic mutant, ΔpspA (pneumococcal surface protein A), while displaying generally more modest regulation patterns in the pneumococcal proteome, showed a distinct and specific regulation of one carbohydrate metabolism locus involved in galactose metabolism. In this locus, the ΔpspA samples up-regulated members of the family of tagatose-6-phosphate pathway proteins lacA-C, suggesting a novel role of pspA in the regulation of galactose metabolism. We further showed that this pspA-mediated regulation is dependent on a wild-type capsule, as both hyper-capsulated and non-capsulated ΔpspA samples were void of lacA-C regulation.

Paper IV

Pneumococcal populations

In this paper, we analysed proteomes of three pneumococcal populations, namely planktonic (broth-grown), biofilm-derived bacteria and biofilm-released bacteria (collected from the supernatant of the biofilm samples) (**Figure 12**). Exposure to heat (by mimicking fever) significantly increased the release of bacteria from biofilms, compared to bacterial release as a part of normal biofilm turnover. The three populations displayed distinct proteomic patterns, as described below.

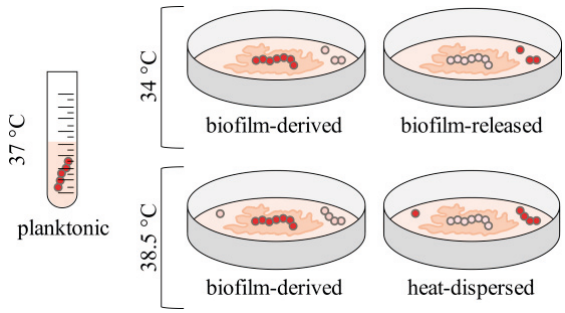


Figure 12. Description of the pneumococcal populations.
Planktonic, broth-grown, bacteria were grown at 37 °C (as conventionally used in laboratory experiments). Biofilms were grown on an epithelial substrate at 34 °C, mimicing both nasopharyngeal environment and temperature. Biofilm-derived isolates were collected from the pellet (e.g., biofilm) of the sample, whereas the biofilm-released isolates are collected from the supernatant (as they are released from the biofilm). Increase in temperature to febrile-range is a factor associated with viral infection and active release of pneumococci from the biofilms. The isolates released from the biofilms in response to heat (38.5 °C) were referred to as heat-dispersed.

Differences between biofilm-derived and heat-dispersed bacteria

Proteomic analysis showed that the heat-dispersed bacteria were metabolically more active than biofilm bacteria, and up-regulated metabolic pathways involved in various carbohydrates, amino acids, fatty acids and nucleotides. Instead, biofilm-derived bacteria up-regulated processes associated with energy metabolism (oxidative phosphorylation and glycan metabolism) and environmental information processing (ABC transporters and signal transduction proteins).

Differences between planktonic and biofilm bacteria

Comparison between planktonic, broth-grown, and biofilm bacteria showed an up-regulation of proteins associated with translation, as well as metabolic pathways involved in fatty acid and nucleotide metabolism to be specific to the planktonic bacterial population. In contrast, proteins expressed by biofilm bacteria were associated with a more sessile lifestyle, exemplified by the up-regulated proteins involved in processes including competence, replication/repair, bacteriocin

production, secretion systems, protein export, and certain ABC transporters. Additionally, the virulence factor pneumolysin was upregulated in planktonic bacteria.

Virulence of the pneumococcal populations in vivo

The virulence of the pneumococcal populations associated with colonization, infection and planktonic bacteria was investigated in four serotypes, by intraperitoneal injection of these populations in a murine pneumococcal infection model. In all the strains, the heat-dispersed isolates showed a significantly increased virulence, however the highest bacterial loads were observed in strains D39 and EF10175 (serotypes 2 and 19F). For these strains, even planktonic bacteria grown in febrile-range temperature showed increased bacterial loads compared to biofilm bacteria, suggesting that temperature itself is an important determinant factor of virulence for certain pneumococcal strains. The combined results indicate large inter-strain differences in pneumococci.

Discussion

The common theme in the papers in this thesis is the use of proteomics to measure the protein content in the included samples. While it is important to remember that proteomics alone cannot deliver answers to every scientific question, it grants the investigating researcher a certain objectivity. Instead of starting with a highly specific hypothesis, we allow the measured protein content of the specifically selected samples speak for itself. As proteins are the mediators of nearly all physiological aspects involved in cellular activities, the abundances of proteins in a sample are naturally of high scientific significance.

The goal of this thesis was to use the strength and power of quantitative proteomics to study the mechanisms involved in a debilitating, and not so well studied, disease. The disease that was chosen, meningitis, is a complex disease caused by a plethora of infectious agents, and is additionally of clinical importance as well. In order to pursue this broad aim, this thesis was split into two blocks, where **Paper I** and **Paper II** are aimed at understanding the molecular mechanisms of the host-response during disease, and **Paper III** and **Paper IV** study the adaptation and specific virulence mechanisms of a disease-causing pathogen.

One major strength of **Paper I** was the patient cohorts from which the CSF samples were collected[60], which included samples from 19 unique pathogenic species, headache controls as well as SAH. Acquiring such a versatile sample collection can be time-consuming, as the prevalence of each specific pathogen responsible for causing meningitis varies largely[61]. For example, the viral and clinical illness involving the CNS known as tick-borne encephalitis (TBE), often presents with the symptoms of meningitis, is reported to have as few as 10,000 yearly cases in Europe[62] – although the number is likely an underestimate. The incidence of *Haemophilus influenzae* type B (Hib) -mediated meningitis has decreased by approximately 49 % between 1990 and 2016[61,63], largely due to the incorporation of Hib in the vaccination programs. While vaccines also exist against certain pneumococcal serotypes, and a decrease in pneumococcal meningitis has also been reported, this form of meningitis still presents with the highest probability of long-term impairment in surviving patients[61].

Several studies have previously attempted to establish the proteome of the human CSF[64,65], but are often limited by the level of quantification or the number of

samples included. More often, the CSF proteome is studied in relation to neurodegenerative diseases[66-70], such as Alzheimer's or Huntington's disease. The proteomic analysis of meningitis CSF has been only done at a more limited scale, by either including a CSF samples from pneumococcal meningitis only[71,72], or include limited number of samples[73].

As the pathogenic species during meningitis either are, or at some point have been, present in the CNS, it would naturally be very interesting to analyse the pathogen-secreted proteins present in the CSF. This was, in fact, one of the original aims for **Paper I**, as we aimed to detect both pathogen-induced host patterns as well as pathogen-secreted peptides in the patient CSF. Assay libraries were generated for eight different bacterial species (*Escherichia coli, Enterococcus faecalis, Streptococcus pyogenes, Streptococcus agalactiae, Streptococcus pneumoniae, Listeria monocytogenes, Pseudomonas aeruginosa* and *Staphylococcus aureus*), which were used to interrogate the DIA-analysed CSF samples. We were unable to detect any bacterial peptides in the CSF samples, which was most likely due to sample processing. The bacterial assay libraries were made publicly available, and still remain as a valuable resource. For future CSF studies, it would be vastly interesting to investigate the pathogen proteome in addition to the host-response in the CSF, which would further our understanding of pathogen adaptation in the sterile environment of CNS.

The correct diagnosis of the underlying meningitis-causing pathogen is of vital importance in a clinical setting, as fast diagnostic tools are crucial for preventing death and reducing the development of long-term effects. Currently, the diagnostic methods are time-consuming, and require bacterial cell culturing. The molecular mechanisms between the interactions of a pathogen and the host dictate how the disease outcome will become. In a disease at heterogeneous as meningitis, a single biomarker approach will not suffice to cover the range of variables observed between meningitis patients. A multiprotein panel approach, such as generated with LASSO in **Paper I**, can therefore provide better coverage in complex diseases. The LASSO regression including the 18 predictive proteins together generated high AUROC for all pathogen groups (80 – 96 %), suggesting that this approach is highly feasible for capturing host proteome patterns between pathogen groups. For many of the predictive proteins, we found supporting evidence from the literature suggesting known biological functions associated with the respective pathogens, while for some of the predictive proteins we found no correlation from the literature. It would be interesting to see if future studies would verify these currently unconfirmed associations between those predictive proteins and the pathogen groups.

Using an assay library containing a layer of information regarding the tissue- and cell-enrichment of human proteins (unpublished data), we could infer the cellular or tissue origin of CSF detected proteins. In addition of detecting large number of neutrophil-associated proteins in the CSF of ABM, several NET-related proteins

were also discovered. The infiltration of neutrophils to CSF during ABM has been previously shown[74,75], but the presence of NETs has not. These findings lead to the initiation of a separate study, where the presence and role of NETs during pneumococcal meningitis was investigated, leading to the preparation of **Paper II**.

In **Paper II**, NETs were for the first time detected in the human CSF during ABM – a finding that has previously only been reported in a porcine bacterial meningitis model[76]. The production of NETs in a sensitive environment of CNS, together with the fact that the CSF-observed NETs were large cellular aggregates, could be a contributing factor to the detrimental side-effects during ABM, such as increased intracranial pressure. In a rat model of pneumococcal meningitis, administration of a DNase treatment cleared the NETs, which led to increased bacterial clearance by the means of increased neutrophil killing mechanisms.

The possible use of DNase as a therapeutic for ABM would be advantageous for multiple reasons. First, as the DNase does not directly target the bacteria, they are therefore unlikely to develop resistance against. In addition, successful DNase treatment would be valuable as an additional therapeutic against ABM in times of increasing antibiotics resistance, as DNase treatment was efficient in clearing both antibiotic-resistant *S. pneumoniae* and *S. aureus*. Furthermore, the safety of use in humans has been already achieved, as aerosolized DNase has recently been repurposed for use in COVID-19 infections[43]. As DNase as a treatment for ABM has not yet been done in a clinical setting, future studies are needed to clarify how the treatment would perform together with current therapeutics, and whether this would have a beneficial effect in the development of long-term neurological sequelae in surviving patients.

In order to understand how a pathogen causes disease, it is important to understand the basic proteomic wiring of the pathogen in question. In this thesis, this was achieved by large-scale proteomics analysis of the most common bacterial cause of meningitis, *Streptococcus pneumoniae.* In **Paper III**, this was achieved by a locus-based annotation and quantitative proteomics analysis of a large number of pneumococcal samples grown in various conditions.

The incentive to separate the pneumococcal proteins into loci based on the pneumococcal genome derives from the fact that most bacterial genes are located under the control of a single promoter (hence, in one locus), which encodes proteins involved in similar functions[77]. While there are various strategies to control the production of a specific protein, we hypothesize that genes (and thereby proteins) belonging to a locus most likely share a similar biological function. This approach could serve as a starting point in suggesting novel roles for uncharacterized proteins that often appear in proteomic analyses. For example, if a locus consists of ten proteins, out of which only nine have a known function in carbohydrate metabolism, it could be presumed that the one last uncharacterized protein would share, at least to a certain extent, the same function.

This approached was utilized in **Paper III**, where the transcriptional landscape of *Streptococcus pneumoniae* strain D39 was mapped by using the published transcriptional landscape of the pneumococcal TIGR4 strain[59] as a template. A large number of conditions were generated, including isogenic mutants of major virulence as well as growth conditions. For the selection of mutants, different classes of virulence factors were chosen, such as choline-binding proteins, lipoproteins, secreted proteins, the important regulatory relay protein serine/threonine kinase protein, as well as variations in the capsule (non-encapsulation and hyper-capsulation), in order to maximise the coverage of different regulatory patterns. By combining the data from the generated conditions, it allowed for the analysis of the stable portion of the proteome versus the hypervariable portion of the proteome. The stable portion of the proteome is interesting from the evolutionary perspective. Any given organism will only produce the exact number of copies of a protein that are required, thereby sparing unnecessary use of energy where it is not needed. The proteins that remained at a stable abundance across all of our conditions would represent the basic machinery that is required for bacterial survival at any given time in the studied conditions.

The hypervariable portion of the proteome, on the other hand, is a result of, and specific, to the investigated condition. One such observation was the down-regulation of the fatty acid biosynthesis (FAB) proteins when pneumococci were exposed to human blood plasma. This has been previously described in *Streptococcus pyogenes*[78], where it was shown that *S. pyogenes* shut off their intrinsic FAB machinery in response to exposure to plasma. The authors further showed that *S. pyogenes* utilizes human serum albumin, the transporter of fatty acids in human plasma. In the case of pneumococci, no albumin-binding surface proteins have been previously reported, and whether this observation is due to albumin-binding still remains to be investigated.

Pneumococci, together with other bacterial species, encompass comprehensive signaling systems which rapidly initiate adaptive responses to environmental stimuli, leading to gene expression. These systems, known as the two-component regulatory systems (TCS), are mediated by the phosphorylation/dephosphorylation cascades of a histidine kinase coupled to a response regulator[79,80]. In pneumococci, thirteen TCSs have been described[81]. The regulatory role of each TCS is not well described, but possible functions have been described to be involved in large cellular pathways, such as energy sensing, nutrient perception, competence and quorum sensing[80,81]. In addition to TCSs, the pneumococcal serine/threonine kinase protein (STKP) acts as a regulator of various cellular functions, and some STKP-mediated signaling networks are thought to overlap with those of the TCSs. A pneumococcal stkp isogenic mutant was included in the proteomic investigation in **Paper III**. The pneumococcal samples void of stkp displayed the second largest changes in the pneumococcal proteome (first one being the deletion of capsule biosynthesis protein), representing the complex regulatory role of stkp. While we

also investigated how the different conditions effected the 13 pneumococcal TCSs, an interesting future project would be a quantitative MS analysis of pneumococcal isolates lacking each of the 13 TCSs, thus gaining valuable information regarding the regulatory role of each of these systems.

Paper IV focused on studying the phenotypic characteristics of wild-type pneumococcal populations. Most of conventional pneumococcal research is conducted on planktonic, broth-grown bacteria in 37 °C. It is arguable whether this correctly represents the in vivo phenotype of the niche-adapted, pneumococci. Growing the pneumococci planktonically in a nutrient-rich media void of immune cells or epithelial cells only produces a snapshot of the bacterial state in that environment, while simultaneously lacking numerous processes that are valuable, and present, in the host during colonization or disease. In addition to planktonic organisms, it is clinically significant to discover the protein patterns discriminating biofilm-derived pneumococci (associated with latent colonization) and pneumococci dispersed from biofilms. In the latter case, it is shown that exposure of biofilms to viral infection -associated factors (such as febrile-range temperatures) lead to the release of pneumococci from the biofilms[82], and that these isolates are associated with invasive disease.

This characterization of colonization versus infection phenotypes in bacteria has been not well characterized. For *Pseudomonas aeruginosa* and *Salmonella enterica* it has been shown that biofilm-dispersed bacteria display a distinct gene expression profile as compared to planktonic or biofilm bacteria[83,84]. **Paper IV** represents a comprehensive proteomic analysis of these three pneumococcal populations, which has previously not been done. The paper displayed distinct differences in proteome patterns between planktonic, biofilm-derived and biofilm-dispersed populations. However, we were unsuccessful in detecting significantly regulated proteins between the biofilm-derived and biofilm-dispersed populations as we had originally hypothesized. One explanation could be that the pneumococcal isolates released from the biofilms are of different “activation state”, and therefore display a heterogenous protein expression. Additionally, a low number of biological replicates were included in this paper. Therefore, additional experimental improvements are required for further investigation of these populations.

In summary, the advancements of **Paper III** and **Paper IV** to the field of pneumococcal research are considerable, as the high-level quantitative proteomics analyses at this scale and the inclusion of specifically chosen, large, sample cohorts have not been previously done. In the scope of this thesis, we can only find pieces of the puzzle for understanding the adaptation and wiring of the pneumococcal proteome, and more studies are required to put the pieces of the puzzle together.

Throughout this thesis, we have collected and generated large sample sets, and one important goal has been the continuous deposition of libraries in the form of proteome map for future use. The CSF proteome maps in **Paper I** (consisting of 171 exceedingly valuable clinical CSF samples) bacterial libraries in **Paper I** (containing samples from eight different bacterial species), as well as the pneumococcal proteome maps in **Paper III** (consisting of 176 pneumococcal samples generated under 30 varying conditions) altogether make up valuable resources for researchers for future investigations. These maps can be continuously reanalysed to test new hypotheses, and to gain more understanding in the field of bacterial adaptation and host-response in health and disease.

As we keep producing vaccines to protect ourselves from virulent strains, and with the increased antibiotics resistance of bacteria, the pathogens will continue to diversify for their survival. Bacterial strains are constantly adapting, which was evident in **Paper IV**, resulting in increased or decreased virulence mechanisms – which could have dire consequences on the human body. Further research is warranted in order to understand the domains and sequences of pathogens that have evolutionarily been conserved, and more importantly, which are prone, and viable, for alteration, and the consequences that these can have in disease.

The combined work of this thesis shows the large versatility of both the human's response as well as the pathogen's adaptation in both health and disease. Furthermore, this thesis provides both knowledge as well as proteomics resources that can hopefully serve as a foundation for future studies of severe infectious diseases.

Concluding remarks

Research is conducted mainly for two reasons, and the first one is curiosity. Our curiosity stems from our deep yearning to understand the laws and forces of nature around us. This starts at a young age when children start to play, which is the first experiment that a human will do. If a child builds a tower of three blocks, and then hits the tower, the blocks will fall. The child will repeat this experiment (children do love repetitions), and the blocks will fall every time. Repetitiveness is another aspect of the experiment of playing. Will the blocks fall again? What happens if I hit the tower with another object? Or make the tower taller? This is in fact research at its simplest form, including an aim (playing), experimentation (building a tower), materials and methods (blocks and a hand), results (tower fell) and conclusions (the tower falls every time). As researchers, we are all doing exactly this – but just at a much more complicated form.

This brings me to the second reason why we conduct research, which is need of information. Most research is aimed at improving health, whether it be diagnostic, therapeutic or else. And in order to achieve these things, the systems we are studying need to be understood at a deep level. It may have been enough for Antonie Philips van Leeuwenhoek to discover the existence of bacteria nearly four centuries ago. However, the discoveries of medicine today are more often at a molecular level – requiring the improvement of instrumentation, technology, skills and knowledge.

The reason why I like proteomics, is that it is a technique that gives you all the information in a sample (as much as the technology and instrumentation of today can allow, of course). It allows the scientist to do a read-out of the investigated samples, without being limited by pre-set and highly specific hypotheses. Especially after the development of DIA tools, the analysed samples become permanent reservoirs of information that can be mined over and over again. I believe this has been a valuable improvement in the field of proteomics and medicine combined. Rare clinical samples, or large experiments, only need to be analysed once, and the data can be iteratively queried based on new hypotheses. This has been one goal of the papers in this thesis. I hope that the work included in this book will be a useful resource in the scientific community to build further knowledge on.

Popular science summary

At any given time, the cells in our bodies are constantly active in processes that keep us healthy and alive. Cells mediate their function through certain types of molecules, known as proteins. Proteins are often known as an important source of energy and as the main component of muscles. However, they are very important for basic functioning of living organisms. During a disease, the proteins of the host (human) interact with the proteins of the infecting organism (pathogen, such as bacteria or virus). Measuring the amounts of different proteins in samples can provide information to understand processes involved in disease. Proteins are however very small, and measuring their amounts cannot be done easily. In this thesis, proteomics is used as a tool to identify and measure the proteins in samples. To do this, an instrument known as the mass spectrometer is used. Simply put, a mass spectrometer is a very large scale, which can measure the weight of the proteins in the samples. By using knowledge of the structure of each protein, we can use the mass that was measured to identify which proteins are in the sample.

Because of the COVID-19 pandemic, severe infections have been a very hot topic worldwide. But other severe infections have not disappeared, and are still causing disease and death worldwide. One such severe infection is meningitis, which is caused by the inflammation of the protective layers around the brain and the spinal cord. This inflammation can be caused by basically any pathogen, including different types of bacteria and viruses. It is a fast-progressing disease, that can, in worst case, lead to sepsis or death. The mortality is particularly high in meningitis caused by bacteria, and survivors often suffer from long-term neurological disability. In addition to the damage caused by the pathogen, our immune system can respond so strongly against the pathogen that this also can lead to damage. Neutrophils are a type of immune cell in the body that fight against infections. They are known to produce traps for bacteria, known as “neutrophil extracellular traps”, or NETs. These are sticky material consisting of DNA where the bacteria can get stuck, and then killed by the neutrophils.

In this thesis the liquid surrounding the brain and the spinal cord, known as cerebrospinal fluid (CSF) was collected from patients suffering from different types of meningitis. We produced a combination of 18 human proteins, which can be used to find out which pathogen is causing the meningitis in patients. Patients suffering from bacterial meningitis had high levels of neutrophils in their CSF, and we showed that the neutrophils produced a lot of NETs. The results showed that

the bacteria during meningitis was able to hide in NETs, and were not killed. A drug known as DNase was used in rats suffering from bacterial meningitis in order to break the NETs. This led to the bacteria being released from NETs and improved the killing of bacteria by neutrophils.

Bacterial meningitis is very dangerous compared to other types of meningitis, and the most common cause of bacterial meningitis is one bacteria, known as *Streptococcus pneumoniae* (*S. pneumoniae*). In fact, up to 80 % of the world's population is carrying this bacteria in their nose and throat right now – without getting sick. This is because the bacteria spends most of its time living in a layer known as "biofilms", and lives there as a colonizer that coexists peacefully with our bodies. But sometimes, certain factors (such as fever or a viral infection) can trigger the biofilms to release a number of bacteria outside the biofilm. These bacteria leave the biofilms, and spread to other parts of our body and can cause disease, such as ear infection, lung infection, sepsis or meningitis.

We studied how the bacteria living inside the biofilms, and the bacteria released from the biofilms (that can cause disease) were different from each other. By measuring the amount and types of proteins that they contain, we showed that the bacteria released from biofilms have a high metabolism, and contain proteins involved in high activity. Often, researchers study *S. pneumoniae* by growing them in a broth, which is not representative to how the bacteria lives in our bodies. We studied these broth-grown bacteria and compared them to biofilms. We found that these two populations are very different from each other.

When *S. pneumoniae* spreads from the biofilms to other parts of our bodies, it comes in contact with our blood. Our blood contains a large number of immune cells. A part of our blood is known as plasma, which contains proteins that are meant to attack invading pathogens. We studied *S. pneumoniae* in plasma, and showed that it has developed mechanisms to increase its survival in our plasma.

This thesis has concentrated on understanding how the human body reacts to pathogens during meningitis. The DNase represents a promising future drug to help kill bacteria in meningitis, but more research needs to be done to make sure it works, and is safe to use, in humans. We have also done extensive research in increasing our understanding about how *S. pneumoniae* has learned to live and cause disease in our body. This can hopefully help us in the future to come up with new medications against bacterial infections caused by *S. pneumoniae.*

Acknowledgements

First and foremost, I would like to thank my supervisor **Johan Malmström**. Your endless source of support and guidance throughout the ups and downs of past years has been invaluable. You have a passion for science that is truly contagious – and I am infected! Thank you for providing with an inspiring workplace that has made coming to work feel like a privilege.

Anita Berglund, you are the glue that keeps everything together, thank you for all that you do for us.

Thank you **Christofer Karlsson**, my co-supervisor, for being the core unit of our group. You have always been a source of support and knowledge during my PhD studies.

To my co-supervisor **Adam Linder**, thank you for supporting and giving me valuable input throughout my PhD. I'm very happy you let me scavenge through your freezers for valuable clinical samples!

Lars Malmström, thank you for all your effort in making our work running and constantly improving.

Lotta, on my very first day, I saw that you had a poster of the "cursing hedgehog" (kiroileva siili) by your desk. Right then I knew that you were a fast-speaking fellow-Finn. Thank you for making Lund feel like a faraway home.

Sounak, you joined the group when I was a lone PhD student, and it has been a blast working and chatting with you. Thank you for your physical and emotional help with the autoclave.

Tirthankar, thank you for always being ready for scientific discussions. Also, talking about Indian and Persian cuisine with you while growing germs at the bact lab has been memorable.

Carlos and Aaron, for always ready to help me with my projects and analyses. Thanks for being so easy to work with.

Tommaso, your sense of humor never fails, thank you for all the laughs and the gigantic cookies that you have baked for us throughout the years.

Di, thank you for being a great office mate, although we lost a lot of time due to the pandemic.

Moritz, Thank you for all the help with the MS method optimizations. I hope I can join you on your next fishing trip.

The IMP group, throughout the years people have come and gone – but the group has always maintained an inspiring atmosphere. Thank you to all, present and past group members, for making my PhD journey so memorable.

Thank you to all the wonderful people at **BioMS** and the **D13 floor**, for a great work atmosphere!

To the people at B14, you have always made me feel welcome at the floor. Thank you for offering a helping hand and guidance whenever I've needed it.

My previous office mate, **Emma Åhrman**, we've shared a lot of fun experiences together! Thank you for your company, help with R, mass spectrometers and Marvin.

My collaborators, **Anders** and **Yosh**, we stumbled into each other and started working on some really fascinating projects. Thank you for all your expertise and knowledge, and being fantastic collaborators.

Barbara, we've been through so much together. Thank you for always being a shoulder to lean on, and a constant source of support and laughter. Soon we will be doctors!

Emmi, although we are far away and rarely meet, our friendship has always stayed strong. Thank you for always listening to me, and for the constant support. I don't know what I would have done without you!

Mikko, my oldest friend. Just because we *can* walk to Ystad and back, doesn't mean we *should*! Thank you for exploring the world with me, and always being there for me.

To my parents, without you I wouldn't be where I am. You have worked so hard to get me where I am today, and I am eternally grateful to you. Thank you for everything you have done for me.

My brother, although you are much taller than me, you will always be my little brother. Thank you for always being there when I need you – no matter where you are in the world.

Christofer, my partner and best friend. It has not been easy few years for us, but throughout it all you have always been my rock. Thank you for your everlasting support, and putting up with all the crazy ideas I come up with.

My children, you two are the best thing that has ever happened to me. Thank you for putting everything into a perspective and giving new meaning to life.

References

1 Weiss, M. C. *et al.* The physiology and habitat of the last universal common ancestor. *Nat Microbiol* **1**, 16116, doi:10.1038/nmicrobiol.2016.116 (2016).

2 Bar-On, Y. M., Phillips, R. & Milo, R. The biomass distribution on Earth. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 6506-6511, doi:10.1073/pnas.1711842115 (2018).

3 Van Leuwenhoek, A. An extract of a letter from Mr. Anth. Van Leuwenhoek, concerning animalcules found on the teeth; of the scaleyness of the skin, &c. *Philos Trans R Soc Lond B Biol Sci* **17**, 646–649, doi:https://doi.org/10.1098/rstl.1693.0012 (1693).

4 Gilbert, J. A. *et al.* Current understanding of the human microbiome. *Nature medicine* **24**, 392-400, doi:10.1038/nm.4517 (2018).

5 Sender, R., Fuchs, S. & Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol* **14**, e1002533, doi:10.1371/journal.pbio.1002533 (2016).

6 Cox, M. J., Cookson, W. O. & Moffatt, M. F. Sequencing the human microbiome in health and disease. *Hum Mol Genet* **22**, R88-94, doi:10.1093/hmg/ddt398 (2013).

7 Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A. & Brown, P. O. Development of the human infant intestinal microbiota. *PLoS Biol* **5**, e177, doi:10.1371/journal.pbio.0050177 (2007).

8 Bik, E. M. Composition and function of the human-associated microbiota. *Nutr Rev* **67 Suppl 2**, S164-171, doi:10.1111/j.1753-4887.2009.00237.x (2009).

9 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. *The Journal of experimental medicine* **79**, 137-158, doi:10.1084/jem.79.2.137 (1944).

10 Yahiaoui, R. Y. *et al.* Prevalence and antibiotic resistance of commensal Streptococcus pneumoniae in nine European countries. *Future microbiology* **11**, 737-744, doi:10.2217/fmb-2015-0011 (2016).

11 Collaborators, G. B. D. L. R. I. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet infectious diseases* **18**, 1191-1210, doi:10.1016/S1473-3099(18)30310-4 (2018).

12 Asokan, G. V., Ramadhan, T., Ahmed, E. & Sanad, H. WHO Global Priority Pathogens List: A Bibliometric Analysis of Medline-PubMed for Knowledge Mobilization to Infection Prevention and Control Practices in Bahrain. *Oman Med J* **34**, 184-193, doi:10.5001/omj.2019.37 (2019).

13 Bridy-Pappas, A. E., Margolis, M. B., Center, K. J. & Isaacman, D. J. Streptococcus pneumoniae: description of the pathogen, disease epidemiology, treatment, and prevention. *Pharmacotherapy* **25**, 1193-1212, doi:10.1592/phco.2005.25.9.1193 (2005).

14 Geno, K. A. *et al.* Pneumococcal Capsules and Their Types: Past, Present, and Future. *Clinical microbiology reviews* **28**, 871-899, doi:10.1128/CMR.00024-15 (2015).

15 Henrichsen, J. Six newly recognized types of Streptococcus pneumoniae. *Journal of clinical microbiology* **33**, 2759-2762, doi:10.1128/jcm.33.10.2759-2762.1995 (1995).

16 Daniels, C. C., Rogers, P. D. & Shelton, C. M. A Review of Pneumococcal Vaccines: Current Polysaccharide Vaccine Recommendations and Future Protein Antigens. *J Pediatr Pharmacol Ther* **21**, 27-35, doi:10.5863/1551-6776-21.1.27 (2016).

17 Hsu, H. E. *et al.* Effect of pneumococcal conjugate vaccine on pneumococcal meningitis. *N Engl J Med* **360**, 244-256, doi:10.1056/NEJMoa0800836 (2009).

18 Nelson, A. L. *et al.* Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infection and immunity* **75**, 83-90, doi:10.1128/IAI.01475-06 (2007).

19 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. *Infection and immunity* **78**, 704-715, doi:10.1128/IAI.00881-09 (2010).

20 Garcia, J. L., Sanchez-Beato, A. R., Medrano, F. J. & Lopez, R. Versatility of choline-binding domain. *Microb Drug Resist* **4**, 25-36, doi:10.1089/mdr.1998.4.25 (1998).

21 Tomasz, A., Albino, A. & Zanati, E. Multiple antibiotic resistance in a bacterium with suppressed autolytic system. *Nature* **227**, 138-140, doi:10.1038/227138a0 (1970).

22 Hoiby, N. A short history of microbial biofilms and biofilm infections. *APMIS* **125**, 272-275, doi:10.1111/apm.12686 (2017).

23 Potera, C. Forging a link between biofilms and disease. *Science* **283**, 1837, 1839, doi:10.1126/science.283.5409.1837 (1999).

24 Blanchette, K. A. & Orihuela, C. J. Future perspective on host-pathogen interactions during bacterial biofilm formation within the nasopharynx. *Future microbiology* **7**, 227-239, doi:10.2217/fmb.11.160 (2012).

25 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. *The Journal of*

antimicrobial chemotherapy **41**, 253-258, doi:10.1093/jac/41.2.253 (1998).
26 Brooks, L. R. K. & Mias, G. I. Streptococcus pneumoniae's Virulence and Host Immunity: Aging, Diagnostics, and Prevention. *Frontiers in immunology* **9**, 1366, doi:10.3389/fimmu.2018.01366 (2018).
27 van de Beek, D. *et al.* Clinical features and prognostic factors in adults with bacterial meningitis. *N Engl J Med* **351**, 1849-1859, doi:10.1056/NEJMoa040845 (2004).
28 van Sorge, N. M. & Doran, K. S. Defense at the border: the blood-brain barrier versus bacterial foreigners. *Future microbiology* **7**, 383-394, doi:10.2217/fmb.12.1 (2012).
29 Dubey, S. P., Larawin, V. & Molumi, C. P. Intracranial spread of chronic middle ear suppuration. *Am J Otolaryngol* **31**, 73-77, doi:10.1016/j.amjoto.2008.10.001 (2010).
30 van Ginkel, F. W. *et al.* Pneumococcal carriage results in ganglioside-mediated olfactory tissue infection. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 14363-14367, doi:10.1073/pnas.2235844100 (2003).
31 van de Beek, D., de Gans, J., Tunkel, A. R. & Wijdicks, E. F. Community-acquired bacterial meningitis in adults. *N Engl J Med* **354**, 44-53, doi:10.1056/NEJMra052116 (2006).
32 Chadwick, D. R. Viral meningitis. *Br Med Bull* **75-76**, 1-14, doi:10.1093/bmb/ldh057 (2005).
33 Bogovic, P. & Strle, F. Tick-borne encephalitis: A review of epidemiology, clinical characteristics, and management. *World J Clin Cases* **3**, 430-441, doi:10.12998/wjcc.v3.i5.430 (2015).
34 Brouwer, M. C., Tunkel, A. R. & van de Beek, D. Epidemiology, diagnosis, and antimicrobial treatment of acute bacterial meningitis. *Clinical microbiology reviews* **23**, 467-492, doi:10.1128/CMR.00070-09 (2010).
35 Lucas, M. J., Brouwer, M. C. & van de Beek, D. Neurological sequelae of bacterial meningitis. *J Infect* **73**, 18-27, doi:10.1016/j.jinf.2016.04.009 (2016).
36 Heckenberg, S. G. B. *et al.* Clinical features, outcome, and meningococcal genotype in 258 adults with meningococcal meningitis: a prospective cohort study. *Medicine (Baltimore)* **87**, 185-192, doi:10.1097/MD.0b013e318180a6b4 (2008).
37 Proulx, N., Frechette, D., Toye, B., Chan, J. & Kravcik, S. Delays in the administration of antibiotics are associated with mortality from adult acute bacterial meningitis. *QJM* **98**, 291-298, doi:10.1093/qjmed/hci047 (2005).
38 Tunkel, A. R. *et al.* Practice guidelines for the management of bacterial meningitis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **39**, 1267-1284, doi:10.1086/425368 (2004).

39 Brinkmann, V. *et al.* Neutrophil extracellular traps kill bacteria. *Science* **303**, 1532-1535, doi:10.1126/science.1092385 (2004).
40 Kaplan, M. J. & Radic, M. Neutrophil extracellular traps: double-edged swords of innate immunity. *Journal of immunology* **189**, 2689-2695, doi:10.4049/jimmunol.1201719 (2012).
41 Johnson, A. J., Ayvazian, J. H. & Tillett, W. S. Crystalline pancreatic desoxyribonuclease as an adjunct to the treatment of pneumococcal meningitis. *N Engl J Med* **260**, 893-900, doi:10.1056/NEJM195904302601801 (1959).
42 Jones, A. P. & Wallis, C. E. Recombinant human deoxyribonuclease for cystic fibrosis. *Cochrane Database Syst Rev*, CD001127, doi:10.1002/14651858.CD001127 (2003).
43 Fisher, J. *et al.* Proteome Profiling of Recombinant DNase Therapy in Reducing NETs and Aiding Recovery in COVID-19 Patients. *Molecular & cellular proteomics : MCP* **20**, 100113, doi:10.1016/j.mcpro.2021.100113 (2021).
44 Wilkins, M. R. *et al.* Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev* **13**, 19-50, doi:10.1080/02648725.1996.10647923 (1996).
45 Kim, M. S. *et al.* A draft map of the human proteome. *Nature* **509**, 575-581, doi:10.1038/nature13302 (2014).
46 Wilhelm, M. *et al.* Mass-spectrometry-based draft of the human proteome. *Nature* **509**, 582-587, doi:10.1038/nature13319 (2014).
47 Yates, J. R., Ruse, C. I. & Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng* **11**, 49-79, doi:10.1146/annurev-bioeng-061008-124934 (2009).
48 Chait, B. T. Chemistry. Mass spectrometry: bottom-up or top-down? *Science* **314**, 65-66, doi:10.1126/science.1133987 (2006).
49 Vandermarliere, E., Mueller, M. & Martens, L. Getting intimate with trypsin, the leading protease in proteomics. *Mass Spectrom Rev* **32**, 453-465, doi:10.1002/mas.21376 (2013).
50 Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198-207, doi:10.1038/nature01511 (2003).
51 Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64-71, doi:10.1126/science.2675315 (1989).
52 Olsen, J. V. *et al.* Higher-energy C-trap dissociation for peptide modification analysis. *Nature methods* **4**, 709-712, doi:10.1038/nmeth1060 (2007).
53 Brun, V. *et al.* Isotope-labeled protein standards: toward absolute quantitative proteomics. *Molecular & cellular proteomics : MCP* **6**, 2139-2149, doi:10.1074/mcp.M700163-MCP200 (2007).
54 Gillet, L. C. *et al.* Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent

and accurate proteome analysis. *Molecular & cellular proteomics : MCP* **11**, O111 016717, doi:10.1074/mcp.O111.016717 (2012).

55 Marcotte, E. M. How do shotgun proteomics algorithms identify proteins? *Nature biotechnology* **25**, 755-757, doi:10.1038/nbt0707-755 (2007).

56 Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature methods* **4**, 207-214, doi:10.1038/nmeth1019 (2007).

57 Rost, H. L. *et al.* OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nature biotechnology* **32**, 219-223, doi:10.1038/nbt.2841 (2014).

58 Guo, T. *et al.* Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps. *Nature medicine* **21**, 407-413, doi:10.1038/nm.3807 (2015).

59 Warrier, I. *et al.* The Transcriptional landscape of Streptococcus pneumoniae TIGR4 reveals a complex operon architecture and abundant riboregulation critical for growth and virulence. *PLoS pathogens* **14**, e1007461, doi:10.1371/journal.ppat.1007461 (2018).

60 Linder, A. *et al.* Heparin-binding protein: a diagnostic marker of acute bacterial meningitis. *Crit Care Med* **39**, 812-817, doi:10.1097/CCM.0b013e318206c396 (2011).

61 Collaborators, G. B. D. M. Global, regional, and national burden of meningitis, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* **17**, 1061-1082, doi:10.1016/S1474-4422(18)30387-9 (2018).

62 Dobler, G. Zoonotic tick-borne flaviviruses. *Vet Microbiol* **140**, 221-228, doi:10.1016/j.vetmic.2009.08.024 (2010).

63 Wenger, J. D. Epidemiology of Haemophilus influenzae type b disease and impact of Haemophilus influenzae type b conjugate vaccines in the United States and Canada. *Pediatr Infect Dis J* **17**, S132-136, doi:10.1097/00006454-199809001-00008 (1998).

64 Schutzer, S. E. *et al.* Establishing the proteome of normal human cerebrospinal fluid. *PloS one* **5**, e10980, doi:10.1371/journal.pone.0010980 (2010).

65 Zougman, A. *et al.* Integrated analysis of the cerebrospinal fluid peptidome and proteome. *Journal of proteome research* **7**, 386-399, doi:10.1021/pr070501k (2008).

66 Abdi, F. *et al.* Detection of biomarkers with a multiplex quantitative proteomic platform in cerebrospinal fluid of patients with neurodegenerative disorders. *J Alzheimers Dis* **9**, 293-348 (2006).

67 Guldbrandsen, A. *et al.* CSF-PR 2.0: An Interactive Literature Guide to Quantitative Cerebrospinal Fluid Mass Spectrometry Data from Neurodegenerative Disorders. *Molecular & cellular proteomics : MCP* **16**, 300-309, doi:10.1074/mcp.O116.064477 (2017).

68 Higginbotham, L. *et al.* Integrated proteomics reveals brain-based cerebrospinal fluid biomarkers in asymptomatic and symptomatic Alzheimer's disease. *Sci Adv* **6**, doi:10.1126/sciadv.aaz9360 (2020).
69 Kroksveen, A. C., Opsahl, J. A., Aye, T. T., Ulvik, R. J. & Berven, F. S. Proteomics of human cerebrospinal fluid: discovery and verification of biomarker candidates in neurodegenerative diseases using quantitative proteomics. *Journal of proteomics* **74**, 371-388, doi:10.1016/j.jprot.2010.11.010 (2011).
70 Minjarez, B. *et al.* Identification of proteins that are differentially expressed in brains with Alzheimer's disease using iTRAQ labeling and tandem mass spectrometry. *Journal of proteomics* **139**, 103-121, doi:10.1016/j.jprot.2016.03.022 (2016).
71 Gomez-Baena, G. *et al.* Quantitative Proteomics of Cerebrospinal Fluid in Paediatric Pneumococcal Meningitis. *Sci Rep* **7**, 7042, doi:10.1038/s41598-017-07127-6 (2017).
72 Goonetilleke, U. R., Scarborough, M., Ward, S. A. & Gordon, S. B. Proteomic analysis of cerebrospinal fluid in pneumococcal meningitis reveals potential biomarkers associated with survival. *The Journal of infectious diseases* **202**, 542-550, doi:10.1086/654819 (2010).
73 Cordeiro, A. P. *et al.* Comparative proteomics of cerebrospinal fluid reveals a predictive model for differential diagnosis of pneumococcal, meningococcal, and enteroviral meningitis, and novel putative therapeutic targets. *BMC Genomics* **16 Suppl 5**, S11, doi:10.1186/1471-2164-16-S5-S11 (2015).
74 Hoffman, O. & Weber, R. J. Pathophysiology and treatment of bacterial meningitis. *Ther Adv Neurol Disord* **2**, 1-7, doi:10.1177/1756285609337975 (2009).
75 Tunkel, A. R. & Scheld, W. M. Pathogenesis and pathophysiology of bacterial meningitis. *Clinical microbiology reviews* **6**, 118-136 (1993).
76 de Buhr, N. *et al.* Neutrophil extracellular trap formation in the Streptococcus suis-infected cerebrospinal fluid compartment. *Cellular microbiology* **19**, doi:10.1111/cmi.12649 (2017).
77 Homuth, G., Mogk, A. & Schumann, W. Post-transcriptional regulation of the Bacillus subtilis dnaK operon. *Molecular microbiology* **32**, 1183-1197, doi:10.1046/j.1365-2958.1999.01428.x (1999).
78 Malmstrom, J. *et al.* Streptococcus pyogenes in human plasma: adaptive mechanisms analyzed by mass spectrometry-based proteomics. *The Journal of biological chemistry* **287**, 1415-1425, doi:10.1074/jbc.M111.267674 (2012).
79 Jung, K., Fried, L., Behr, S. & Heermann, R. Histidine kinases and response regulators in networks. *Curr Opin Microbiol* **15**, 118-124, doi:10.1016/j.mib.2011.11.009 (2012).
80 Gomez-Mejia, A., Gamez, G. & Hammerschmidt, S. Streptococcus pneumoniae two-component regulatory systems: The interplay of the

pneumococcus with its environment. *Int J Med Microbiol* **308**, 722-737, doi:10.1016/j.ijmm.2017.11.012 (2018).
81 Lange, R. *et al.* Domain organization and molecular characterization of 13 two-component systems identified by genome sequencing of Streptococcus pneumoniae. *Gene* **237**, 223-234, doi:10.1016/s0378-1119(99)00266-8 (1999).
82 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**, doi:10.1128/mBio.00438-13 (2013).
83 Chambers, J. R., Cherny, K. E. & Sauer, K. Susceptibility of Pseudomonas aeruginosa Dispersed Cells to Antimicrobial Agents Is Dependent on the Dispersion Cue and Class of the Antimicrobial Agent Used. *Antimicrobial agents and chemotherapy* **61**, doi:10.1128/AAC.00846-17 (2017).
84 Gal-Mor, O. Persistent Infection and Long-Term Carriage of Typhoidal and Nontyphoidal Salmonellae. *Clinical microbiology reviews* **32**, doi:10.1128/CMR.00088-18 (2019).