NANOSCOPIC ADVENTURES

Unraveling Macromolecular Complexes in Infectious Diseases via Integrative Structural Biology

Kai Patric Ehrenbolger

Department of Molecular Biology
Department of Medical Biochemistry and Biophysics
Umeå 2024
Mercie, fürd Unterstützig und Verständniss über all die Johr
# Table of content

I. Abstract .................................................................................................................. iii

II. Abbreviations ........................................................................................................ v

III. Publication List ..................................................................................................... vi

1. Introduction to Infectious Diseases ................................................................. 1
   1.1. A Short History of Infectious Diseases .................................................... 1
   1.2. The Burden of Infectious Diseases in the Modern Era ......................... 2
   1.3. Environmental and Socioeconomic Factors: ........................................ 2

2. Co-Evolutionary Host–Pathogen Arms Race .............................................. 5
   2.1. Adaptive and Reductive Co-Evolution .................................................. 5
   2.2. A Unique Fungal Eukaryotic Parasite – Microsporidia ......................... 6
       2.2.1. Discovery .......................................................................................... 6
       2.2.2. General Introduction ...................................................................... 6
       2.2.3. A Specialized Infection Apparatus – The Polar Tube ................. 8
   2.3. From Microbial Adaptation to Antibiotic Resistance: ......................... 9
       2.3.1. Origins and Functions of Antibiotic Resistance in Microbes: 9
       2.3.2. Multidrug Resistance in Enterococcus faecalis ............................. 9
           2.3.2.1. Enterococcus faecalis ................................................................. 9
           2.3.2.2. ABC Transporters as Multidrug Transporters: ................... 10
           2.3.2.3. EfrCD .................................................................................... 11

3. Protein Life Cycle .................................................................................................. 13
   3.1. The Molecular Machinery ........................................................................ 13
   3.2. Protein Synthesis – The Ribosome .......................................................... 13
       3.2.1. Composition, Function, and Regulation ........................................ 13
       3.2.2. The Eukaryotic Ribosome ............................................................... 16
   3.3. Protein Degradation – The Proteasome .................................................... 17

4. The Tools to Investigate Protein Function and Mechanism - Structural Biology ................................................................................................................. 19
   4.1. General Introduction Structural Biology ............................................... 19
   4.2. The Central Methods Employed in This Thesis ....................................... 21
       4.2.1. Protein Production and Purification ................................................ 21
       4.2.2. Cryo-EM: ...................................................................................... 22
           4.2.2.1. Single Particle Cryo-Electron Microscopy - SPA .................. 23
   4.3. Experimental vs Computational Structural Biology: ........................... 27

5. Objectives of the Thesis ....................................................................................... 29

6. Publications and Supplementary Data ............................................................. 30
   6.1. Project 1: A Molecular Basis of Multi-Drug Export in Enterococcus faecalis (partially published) ................................................................. 30
       6.1.1. Background and aim ...................................................................... 30
6.1.2. Results and Discussion of Supplementary Data .................. 30
   6.1.2.1. Summary of Unpublished Results ................................ 30
6.1.3. Analyzed Datasets and Obtained Conformational States of EfrCD31
   6.1.3.1. Overall Conformational Differences .............................. 31
   6.1.3.2. Nanobody NB193 Supplemented Samples 3 and 6 .................. 32
   6.1.3.3. Cryo-EM Data Collection and Processing of Sample 6 (EfrCD\textsuperscript{wt}:ATP\textgamma{}S: NB193) ............................................. 32
6.1.4. The Structure of EfrCD:ATP\textgamma{}S: NB193 ......................... 32
   6.1.4.1. Overall Structure ................................................. 32
6.1.5. Conclusion of Supplementary Data .................................... 33
6.1.6. Summary of Publication ................................................ 34
6.1.7. Additional Methods for Publication ................................... 34
   6.1.7.1. Sample Preparation ................................................ 34
6.2. Project 2: Structural and Functional Characterization of the Microsporidian Polar Tube Architecture ......................... 35
   6.2.1. Background and Aim .................................................. 35
6.2.2. Results and Discussion of Supplementary Data ...................... 35
   6.2.2.1. Purification of PTP3 Constructs ................................... 35
   6.2.2.2. Initial Insight Into Endogenous Purified PTP3 .................... 37
   6.2.2.3. Structural Characterization of PTP3 ............................... 38
6.2.3. Conclusion of Supplementary Data .................................... 39
6.2.4. Summary of Publication ................................................ 40
6.2.5. Methods not Described in Publication ................................ 41
   6.2.5.1. Cloning of Overexpression Plasmids .............................. 41
   6.2.5.2. Production and Purification of PTP3 .............................. 41
   6.2.5.3. Further Purification of VnPTP3\_201-1'175 ......................... 42
   6.2.5.4. Further Purification of VnPTP3\_201-1'175 ......................... 42
   6.2.5.5. Negative Stain ...................................................... 42
   6.2.5.6. Grid Freezing and Data Collection ............................... 42
6.3. Project 3: Structural Characterization of the Microsporidian Proteasome – Reduction and Regulation ............................ 43
   6.3.1. Background and Aim .................................................. 43
   6.3.2. Summary of Publication ................................................ 44
6.4. Project 4: Structural Characterization of a new Microsporidian Ribosome Hibernation Mechanism ................................. 45
   6.4.1. Background and Aim .................................................. 45
   6.4.2. Summary of Publication ................................................ 45
7. Summary of Research Findings & Discussion .............................. 47
8. Acknowledgment ............................................................... 48
9. References ........................................................................... 49
I. Abstract

This thesis focuses on understanding the underlying molecular mechanisms of infectious diseases, which claim nearly 9 million lives annually. The research centers on critical analysis of pathogen mechanisms and drug resistance. I have mainly focused on two clades of pathogens: Enterococcus faecalis and microsporidia. E. faecalis is a key nosocomial opportunistic pathogen, and microsporidia are a group of emerging fungal pathogens that considerably impact the environment and economy, causing, among other things, the decline of honeybee populations. In this thesis, I have combined biochemistry and cryo-electron microscopy to perform an in-depth molecular analysis of crucial protein complexes that drive the infectivity of these organisms. In E. faecalis, the primary drug efflux pump, EfrCD, is examined to gain insight into its role in antibiotic resistance.

Microsporidia often have a drastically reduced genome and display altered macromolecular structures due to their parasitic lifestyle. The research aims to provide insights into the regulation of translational processes in microsporidia by comparing the dormant spore stage to the active intracellular stage and looking closely into the infection mechanism.

In the publication “Deep mutational scan of a drug efflux pump reveals its structure–function landscape,” I determined the structure of EfrCD and several of its conformations to understand better how this protein complex contributes to E. faecalis' multidrug resistance. In further research, our focus moved to Microsporidia.

During our work on the “Structure of the reduced microsporidian proteasome bound by PI31-like peptides in dormant spores” and “Differences in structure and hibernation mechanism highlight diversification of the microsporidian ribosome,” we solved the structure of the endogenous microsporidian ribosome as well as multiple versions of the proteasome; the dormant form of 20S proteasome and the active form of the 20S and 26S proteasome. This gave a deeper understanding of how microsporidia could highly reduce those conserved macromolecular complexes. By discovering novel inhibitors, we were also able to understand how those energy-demanding molecular machines can efficiently regulate themselves.
Furthermore, we investigated the specialized infection organ of microsporidia, known as the polar tube. As part of the paper titled "Ribosome clustering and surface layer reorganization in the microsporidian host-invasion apparatus," I contributed by performing proteomic analysis of the endogenously affinity-purified polar tubes using a native affinity tag I discovered. Additionally, I identified potential protein-protein interactions of the polar tube proteins. This complemented the work performed on the dynamics and ultrastructure remodeling of the polar tube during germination through light microscopy and cryo-electron tomography. We observed a cargo-filled state with organized arrays of ribosomes clustered along the thin tube wall and an empty post-translocation state with a thicker wall.

The findings of this thesis work expand our understanding of pathogen biology and open up new possibilities for addressing drug development and drug resistance, a significant global health challenge.
II. Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTD</td>
<td>Neglected Tropical Diseases</td>
</tr>
<tr>
<td>AMR</td>
<td>Anti-Microbial Resistance</td>
</tr>
<tr>
<td>PTP</td>
<td>Polar Tube Protein</td>
</tr>
<tr>
<td>MDR</td>
<td>Multi-Drug Resistance</td>
</tr>
<tr>
<td>NBD</td>
<td>Nucleotide Binding Domain</td>
</tr>
<tr>
<td>TMD</td>
<td>Trans Membrane Domain</td>
</tr>
<tr>
<td>P-gp</td>
<td>P-glycoprotein</td>
</tr>
<tr>
<td>BCRP</td>
<td>Breast Cancer Resistance Protein</td>
</tr>
<tr>
<td>MRP1</td>
<td>Multidrug Resistance Protein 1</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transporter RNA</td>
</tr>
<tr>
<td>A-site</td>
<td>Aminoacyl-site</td>
</tr>
<tr>
<td>P-site</td>
<td>Peptidyl-site</td>
</tr>
<tr>
<td>E-site</td>
<td>Exit-sites</td>
</tr>
<tr>
<td>ES</td>
<td>Expension Segment</td>
</tr>
<tr>
<td>UPS</td>
<td>Ubiquitin-Proteasome System</td>
</tr>
<tr>
<td>DUB</td>
<td>Deubiquitinating enzymes</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>cryo-EM</td>
<td>Cryogenic Electron Microscopy</td>
</tr>
<tr>
<td>cryo-ET</td>
<td>Cryo-Electron Tomography</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilized Metal Affinity Chromatography</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione S-Transferase</td>
</tr>
<tr>
<td>MBP</td>
<td>Maltose-Binding Protein</td>
</tr>
<tr>
<td>DMS</td>
<td>Deep Mutational Scanning</td>
</tr>
<tr>
<td>VnPTP3</td>
<td><em>V. necatrix</em> PTP3</td>
</tr>
<tr>
<td>NS</td>
<td>Negative Stain</td>
</tr>
<tr>
<td>PTP</td>
<td>Polar Tube Protein</td>
</tr>
<tr>
<td>PTM</td>
<td>Post Translational Modification</td>
</tr>
<tr>
<td>RBL</td>
<td>Ricin-B Lectin domain-containing proteins</td>
</tr>
<tr>
<td>CV</td>
<td>Column Volume</td>
</tr>
</tbody>
</table>
III. Publication List

I. Deep mutational scan of a drug efflux pump reveals its structure–function landscape.
   Gianmarco Meier, Sujani Thavarasah, Kai Ehrenbolger, Cedric A. J. Hutter, Lea M. Hürlimann, Jonas Barandun & Markus A. Seeger
   https://doi.org/10.1038/s41589-022-01205-1

II. Ribosome clustering and surface layer reorganization in the microsporidian host-invasion apparatus
   Himanshu Sharma, Nathan Jespersen, Kai Ehrenbolger, Lars-Anders Carlson, Jonas Barandun. bioRxiv 2023.05.31.543061;
   https://doi.org/10.1101/2023.05.31.543061
   Accepted for publication in PLOS Biology

III. Differences in structure and hibernation mechanism highlight diversification of the microsporidian ribosome.
   https://doi.org/10.1371/journal.pbio.3000958

IV. Structure of the reduced microsporidian proteasome bound by PI31-like peptides in dormant spores.
   https://doi.org/10.1038/s41467-022-34691-x
1. Introduction to Infectious Diseases

1.1. A Short History of Infectious Diseases.

Infectious diseases have significantly impacted human civilizations, economies, and cultures. These diseases are caused by various microorganisms, including viruses, bacteria, fungi, and other eukaryotic parasites. Evidence of their impact can be discovered throughout history, starting with the remains of prehistoric humans, documentation, and samples from ancient Greece till the Black Death in Europe.

The Black Death was one of the most devastating events of the medieval period. During those times, microorganisms were first proposed as the cause of infectious diseases. However, this theory of scholars from al-Andalus was not widely accepted in Europe until the Renaissance. Nevertheless, the foundation for understanding disease transmission was already laid in the early 11th century by the famous scholar Avicenna, who recognized the contagious nature of infectious diseases.

The 17th to 19th century was a time of global exploration—unfortunately, this led to the spread of smallpox and other diseases to new continents. Furthermore, growing populations and the lack of proper sanitation led to a higher incidence of cholera and typhoid outbreaks in urban populations. It was also a time of groundbreaking advancements in microbiology. Anton van Leeuwenhoek enhanced our understanding of the microbial world through light microscopy.

In the 19th century, Robert Koch and Louis Pasteur revolutionized microbiology with their germ theory. At the same time, pioneers like Edward Jenner and Jonas Salk introduced vaccines against various diseases. Alexander Fleming's discovery of penicillin and Gerhard Domagk's development of sulphonamides accelerated the antibiotic movement. This period between 1857 and 1914 is often called the "golden age" of microbiology due to the significant advancements made during this time. Unfortunately, this ended with the shift of resources due to the start of World War I.

In the 20th century, the occurrence of multidrug resistance along with new threats like HIV/AIDS and diseases such as SARS, MERS, and COVID-19 in the 21st century has presented challenges that are yet to be solved. Moreover, over half of the world's population currently lives in urban areas. By 2050, this number is expected to increase by another two-
thirds, which increases the threat of diseases even more. The population density worsens sanitation, leading to the spread of various diseases from mosquitoes (dengue, Zika virus, chikungunya, etc.), water sources (cholera, typhoid, giardiasis, cryptosporidiosis), and airborne respiratory infections (SARS, avian flu, and tuberculosis).

The relationship between humans and pathogens has been complex throughout history. We must be constantly aware and able to adapt to an ever-changing infectious threat. The progress in understanding and combating diseases from our prehistoric ancestors to today, as well as the recent pandemic, reminds us that continued research and preparedness are needed.

1.2. The Burden of Infectious Diseases in the Modern Era.

Infectious diseases have far-reaching economic and health implications for people worldwide. Even before the COVID-19 pandemic, the burden of infectious diseases was colossal. Research revealed that diseases like HIV/AIDS, malaria, and tuberculosis were responsible for approximately 680,000 (2020), 619,000 (2021), and 1.6 million (2021) deaths, respectively. The global burden of “years of life lost,” as well as the financial burden of those diseases, is already massive even without taking measles, hepatitis, dengue fever, rabies, yellow fever, and neglected tropical diseases (NTDs) into account. In 2016 alone, infectious diseases caused a loss of over 156 million life years and incurred up to $8 trillion in costs, including direct healthcare expenses and broader socio-economic impacts.

The numbers above don’t include one of the most severe threats we face today—antimicrobial resistance (AMR). Shockingly, in the year 2019 alone, 1.25 million people died due to AMR. If left unchecked, AMR could cause up to 10 million deaths annually and lead to potential economic costs of up to $100 trillion by 2050. Thus, we must undertake more research and launch a global effort to combat this issue.

1.3. Environmental and Socioeconomic Factors:

Infectious diseases are deeply entwined with environmental and socioeconomic contexts. Climate change, for instance, promotes the spread of vector-borne diseases like malaria, dengue, and Zika by expanding the habitats of carrier mosquitoes. Deforestation brings
humans closer to wildlife, elevating the risk of zoonotic diseases while disrupting ecological balances that could keep potential infections in check. Water scarcity and poor sanitation are fertile grounds for water-borne diseases, such as those caused by the parasite Schistosoma or bacteria like Vibrio cholerae. Human migration introduces diseases into new regions, as evidenced by the spread of Chikungunya to America.

Furthermore, high population densities in urban areas can accelerate the transmission of respiratory diseases such as tuberculosis and influenza and facilitate the spread of other diseases, such as HIV/AIDS, due to diverse social interactions. Other socioeconomic factors, including access to healthcare, economic status, and general and public health education, play critical roles in disease dynamics. Poverty, malnutrition, lack of clean water, and inadequate housing increase vulnerability to infections.

Among the man-made factors we have considered, climate change could be the most consequential. The impact of climate change is substantial and has been identified as a catalyst for 58% of human infectious diseases. Altering ecosystems leads vectors to spread. Additionally, higher global temperatures contribute to the increased heat tolerance of fungal pathogens, which can lead to more widespread infections.

How human-induced environmental and socioeconomic changes lead to the spread of infectious diseases can be exemplified by the Saint Gothard anemia. The construction of the Gotthard Tunnel in Switzerland was a remarkable achievement of the 19th century. However, it also became a case study for spreading infectious diseases and the intersection of human labor migration and disease. Many workers faced the endemic spread of hookworm infections, particularly Ancylostoma duodenale, which led to widespread cases of "Saint Gothard anemia." The construction created a cramped and damp environment due to geothermal heat and explosives. Furthermore, poor sanitation forced workers to relieve themselves inside the tunnel. This provided a fertile environment for the tropical parasite to survive. This phenomenon was not just limited to the Gotthard Tunnel. It mirrored a broader pattern where human migration contributed to the spread of parasites.

This illustrates that managing infectious diseases is a complex task. It involves various environmental and human factors further complicated by globalization, accelerating the spread of diseases across borders. Adopting a comprehensive approach that combines medicine with social, economic, and environmental strategies is needed to overcome this burden. At the
same time, more research in host-pathogen interactions is required to tackle the complex landscape of infectious diseases.
2. Co-Evolutionary Host–Pathogen Arms Race

2.1. Adaptive and Reductive Co-Evolution

Host-pathogen coevolution is a crucial process that shapes the evolution of life, affects genetic diversity, and influences disease dynamics. This process involves a continuous arms race between the host and the pathogen, leading to pathogen specialization and increased virulence while concurrently promoting diversity.

To explain the dynamics between hosts and pathogens, two evolutionary hypotheses have been proposed: the Red Queen Hypothesis and the Black Queen Hypothesis. The Red Queen Hypothesis postulates that species must continually adapt and evolve in response to other species. It highlights the ongoing evolutionary arms race between hosts and pathogens. Furthermore, it emphasizes that sexual reproduction is an efficient and fast mechanism to adapt to pathogenic and parasitic threats. This ongoing host-pathogen arms race is demonstrated by the rapid development of antibiotic resistance in response to increased drug usage in the past century, which poses a significant threat to human health. Bacteria, such as Enterococcus faecalis, emerge as prime examples. The multidrug efflux pump EfrCD from E. faecalis exemplifies an adaptive response to the selective pressure exerted by antibiotic use (Project 1). This mechanism enables the bacterium to resist a wide range of antibiotics, representing a direct challenge to medical treatment and showcasing a pathogen’s ability to adapt efficiently to human interventions.

In contrast, the black Queen Hypothesis explains that natural selection drives gene loss, particularly in microbial communities, leading to reductive evolution. It posits that organisms may lose specific genes or functions if they become redundant due to availability from other community members. This streamlines organisms’ genomes for efficiency in coevolutionary contexts by removing redundant genes.

Microsporidia are an illustrative case of reductive evolution, as many streamlined their genome to an extreme level, demonstrates the interconnectedness in biological systems and the implications of coevolution in host-pathogen relationships.
2.2. A Unique Fungal Eukaryotic Parasite – Microsporidia

2.2.1. Discovery

Microsporidia, now recognized as fungi or a sister group to fungi\(^{43}\), were first identified as a distinct group in the mid-19th century\(^{44}\). These unicellular parasites form spores and have a unique mechanism for infecting host cells. Initially considered primitive eukaryotes, they are now recognized as highly specialized organisms\(^{45}\).

In the mid-1800s, the silk production industry in France was severely impacted by a disease that had a devastating effect on silkworm agriculture. To tackle the issue, the French government enlisted the help of Louis Pasteur, who was already a well-known scientist at that time. Despite having no prior experience with silkworms or animals, Pasteur conducted research on the disease between 1865 and 1870\(^{46}\). His work led to the discovery that two separate infections affected silkworms. These diseases were known as pébrine and flacherie and were caused by microsporidia and bacteria respectively\(^{47}\). Pébrine disease, also known as "pepper disease," is caused by the microsporidian parasite *Nosema bombycis*. It results in brown dots on silkworm larvae and makes them unable to spin silk\(^{48}\). Through his experiments, Pasteur suggested procedures to help breeders restore their silkworm stocks and revive silk production. These procedures included identifying and isolating uninfected silkworms, selective breeding, and improving the cleanliness of silkworm colonies. Although these steps might seem obvious today they were not at the time\(^{47}\).

His work saved the silkworm industry, and the microsporidian field was initialized and shaped over 150 years of research. However, despite the duration of microsporidian research, our current understanding of microsporidia at the molecular level is limited. There are still challenges in developing new antifungal drugs, and tools to understand the biology of microsporidia are still scarce.

2.2.2. General Introduction

Microsporidia are obligate intracellular parasites infecting all animals, including agriculturally and environmentally important species such as silkworms (sericulture)\(^{49}\) and fish, and shrimps (aquacultures)\(^{50}\). Maybe the most significant impact of microsporidia has been observed in honeybees, where microsporidia have been linked to the hive collapse of free-living and managed honeybees over the last decade\(^{50}\). The number of
insects is increasingly declining worldwide, causing an ecosystem imbalance that adversely affects crop pollination\textsuperscript{51}. Furthermore, at least 14 known species of microsporidia can infect humans\textsuperscript{52}, causing a variety of clinical syndromes of the digestive and nervous system, primarily but not exclusively in immunocompromised patients, the elderly, and children\textsuperscript{53,54}. On the other hand, microsporidia are also used as biopesticides against locus swarms\textsuperscript{55} and are explored as a tool against malaria transmission\textsuperscript{56}.

Microsporidia have evolved to live parasitically, which led to drastic genome reduction\textsuperscript{57}. This resulted in the smallest known eukaryotic genome\textsuperscript{58-60} and the compaction of various evolutionarily conserved macromolecules like their ribosomes or proteasome\textsuperscript{42,61-64} (studied in \textbf{Projects 3 and 4}). Furthermore, the loss of numerous biosynthetic pathways and the reduction of their mitochondria to rudimental mitosomes\textsuperscript{65} makes them reliant on their host’s cellular resources\textsuperscript{57}. Without functional mitochondria, they cannot generate energy independently, necessitating their replication within a host. Outside a host, they survive as dormant spores in environments like water and soil\textsuperscript{66}. This transition between an actively replicating cell and a dormant spore requires an efficient shutdown and reactivation of cellular processes and macromolecular complexes, as well as energy-efficient infection of a host, which is crucial for their life cycle and reproduction (Investigated in \textbf{Projects 2-4}).

\textbf{Figure 1: Overview of Microsporidia} (A) Environmental microsporidian spores (left) transport their infection material via a specialized infection apparatus called the Polar Tube (right) through a process called germination. (B) The Polar Tube comprises at least 5-6 different Polar Tube Proteins: PTP1-6. PTP1-3 makes up the elongated part of the tube, while PTP4 and 5 are located at the tip of the tube and are presumably involved in the host-cell interaction.
2.2.3. *A Specialized Infection Apparatus – The Polar Tube*

One of the most peculiar features of microsporidia is their mode of infection, which involves a long filamentous tube called the polar tube. The polar tube transfers the cytoplasm and nucleus, termed the sporoplasm, from the microsporidian spore into a host cell\(^{67,68}\).

During the environmental spore stage, the polar tube is coiled inside the cell\(^{69}\). Once the host takes up the spore, the surrounding conditions trigger the eversion of the tube, which most likely attaches to a host cell\(^ {70}\) (See Figure 1 A). After host recognition, the entire spore content is pushed through this polar tube into the cell, where microsporidia can replicate and produce new spores. How the sporoplasm enters the host cell has yet to be fully understood. Still, the most popular hypotheses suggest that the tube pierces the cell or attaches to the host cell surface, and the sporoplasm is endocytosed\(^ {71}\).

The complete composition of the tube remains to be explored (studied in Project 4). Based on current knowledge, the polar tube comprises five to six different proteins (as shown in Figure 1 B). Polar Tube Protein (PTP) 1 to 3 forms a proteinaceous repetitive element, which makes up the structural, elongated part of the infection tube. On the other hand, antibodies against PTP5 bind to the whole extruded part of the polar tube\(^ {72}\). At the same time, PTP4 is located only at the tip of the tube\(^ {73}\) and plays a role in host cell recognition\(^ {70}\). Recently, a new protein called PTP6 has been discovered\(^ {74}\), but its function as a polar tube protein is still being debated and researched.

Despite the importance and increasing impact of microsporidia on human health and the environment, very little is known about the structural components and molecular mechanisms of tube evasion, host cell recognition, and transport of the cellular content through the polar tube (Project 2).
2.3. From Microbial Adaptation to Antibiotic Resistance:

2.3.1. Origins and Functions of Antibiotic Resistance in Microbes:

A major global challenge that has drastically changed how we manage infectious diseases is Antimicrobial resistance (AMR). The natural resistance development process is accelerated by the inappropriate use of antibiotics in human, animal, and agricultural sectors. Therefore, treating once-manageable diseases becomes more complex, and medical procedures become more complicated as we run out of ways to treat infected patients.

Efforts to combat AMR have led to global strategies like the “One Health approach” and the “Global Action Plan on AMR.” These strategies aim to foster a collaborative response to AMR by integrating human, animal, and environmental health strategies. However, despite these initiatives, few new antimicrobial drugs have been developed. The WHO's priority pathogens list emphasizes the importance of understanding the origins and mechanisms of resistance to develop effective countermeasures.

Microbes can evolve mechanisms that help them resist the effects of antimicrobial agents. These mechanisms include a range of molecular tactics such as enzymatic degradation or inactivation of antibiotics, drug efflux, target site modification, or reducing drug affinity. Because of these mechanisms, bacteria can withstand the effects of antimicrobial agents, making it challenging to fight them. To manage and prescribe antibiotics effectively, it is crucial to identify microbial resistance phenotypes and the genetic mutations that cause this resistance.

The complexity of resistance mechanisms poses a significant challenge in treating infections. To make matters worse, some pathogens acquired several mechanisms. Enterococcus faecalis is a prime example of multidrug resistance that characterizes many defenses bacteria have developed against antimicrobial agents.

2.3.2. Multidrug Resistance in Enterococcus faecalis

2.3.2.1. Enterococcus faecalis

The Gram-positive bacterium E. faecalis, a typical resident of the gut microbiota, has ascended as a formidable multidrug-resistant (MDR) pathogen, particularly within nosocomial settings. This organism is
known for its intrinsic resilience and ability to acquire resistance mechanisms and is generally resistant to a broad range of antibiotics\textsuperscript{84}. It often causes severe infections such as bacteremia and endocarditis\textsuperscript{85}.

The resistance observed in \textit{E. faecalis} is a byproduct of chromosomal adaptations and horizontal gene transfer via the type IV secretion system\textsuperscript{86}. Whole-genome sequencing of a vancomycin-resistant \textit{E. faecalis} strain has revealed that more than one-quarter of the predicted protein-encoding open reading frames originate from mobile and exogenously acquired DNA\textsuperscript{87}. This genetic mobility complicates treatment options and raises the potential for resistance genes in other pathogens\textsuperscript{88}.

A critical contributor to multidrug resistance is the active efflux of antimicrobial agents, mediated by transport proteins that serve as bacterial defense mechanisms. Among these, ABC transporters are one of the main contributors.\textsuperscript{89}

\subsection*{2.3.2.2. \textit{ABC} Transporters as Multidrug Transporters:}

ABC transporters, short for ATP-binding cassette transporters, are a class of membrane proteins crucial in transporting various substrates across cellular membranes\textsuperscript{90}. They are part of one of the largest and oldest gene families, with their origins tracing back over a billion years\textsuperscript{91}. They are found in all living organisms and are involved in a wide range of biological processes\textsuperscript{92}.

The hydrolysis of ATP powers the transport process, hence the name ATP-binding cassette. The transporter`s mechanism is based on the ability to couple the energy derived from ATP to transport different substrates across the membrane via conformational change\textsuperscript{92}. Both ABC importers and ABC exporters exist, but for this thesis, I will only focus on exporters.

Structurally, ABC exporters comprise two main domains: the highly conserved nucleotide-binding domain (NBD) and the transmembrane domain (TMD). The TMD, which shows low sequence conservation, is believed to reflect the need to accommodate a wide variety of substrate types, and at least seven types of ABC transporters are classified based on TMD folds and topologies\textsuperscript{93}. In recent years, there has been significant progress in understanding their structure through high-resolution structural analysis\textsuperscript{92}. 
Figure 2 The ABC Exporter Switch Model: Conformational changes in the TMD, induced by ATP hydrolysis in the NBDs, enable the transport of compounds across the membrane bilayer by an unknown mechanism.

The Switch model is probably the best-known model to explain how ABC transporters work. This Model suggests that the NBDs are separated and form an inverted "V" shape with the TMDs facing the inside of the cell (See Figure 2). According to this model, when ATPs and the substrate bind, the transporter changes from an inward-facing shape to an outward-facing one.

Despite this and other models, there remains uncertainty about the exact transport mechanism for specific ABC transporters. How substrate binding and transport is coupled to ATP binding and hydrolysis is not fully understood. Specifically, multidrug resistance in pathogens or cancer cells can be linked to the overexpression of ABC transporters such as P-glycoprotein (P-gp), breast cancer resistance protein (BCRP), multidrug resistance protein 1 (MRP1), and EfrCD is yet to be fully understood.

2.3.2.3. EfrCD

During research on the various ABC transporters in E. faecalis, one transporter was identified as a significant mediator of multidrug efflux. The 128 kDa heterodimeric transporter EfrCD consists of the two subunits EfrC and EfrD, harboring an N-terminal TMD and a C-terminal NBD.

Deleting the EfrCD encoding genes led to a notable increase in the susceptibility of E. faecalis to various antibiotics (See Figure 3 left panel). Furthermore, the transporter was compared with the previously believed main contributor to drug resistance in E. faecalis, EfrAB. Interestingly, under endogenous conditions, EfrAB was found to have only a marginal role, while when expressed in Lactococcus lactis, both EfrAB and EfrCD facilitated the efflux of fluorescent substrates and...
conferred resistance to multiple drugs, including fluoroquinolones⁸⁹ (See Figure 3 right panel). Those findings inspired us to proceed to the structural characterization of EfrCD and, in collaboration with the original authors of the discovery paper, shed light on its molecular mechanism (Project 1).

**Figure 3: EfrCD and its Role in Drug Resistance.**

A) Minimal inhibitory concentrations were experimentally determined for wild type and seven different ABC-exporter knock-outs. EfrCD knock-out showed a decreased level of resistance to Daunorubicin, Doxorubicin, Ethidium, and Hoechst 33342.

B) Increase in drug resistance upon over-expression of enterococcal transporters in L. lactis. Adapted from Lea M. Hürlimann et al. 2016.
3. Protein Life Cycle

3.1. The Molecular Machinery

Proteins are the main molecular machines of life. They start their life cycle through the translation of the genetic code into polypeptide chains. Ribosomes, which are cellular machines composed of rRNA molecules and proteins, play a crucial role in this process. They decode the genetic instructions that mRNA carries and translate them into polypeptide chains by adding amino acids represented by codons on the mRNA sequence. After a protein is synthesized, flawed, misfolded, or no longer needed, proteins are marked for degradation, a process often started via ubiquitination by the cell's quality control mechanisms for degradation. This is a crucial step in maintaining proteostasis. The proteasome, the counterpart of the ribosome, carries out the essential degradation of proteins within the cell. This interplay between protein synthesis and degradation is a delicate balance, ensuring that only correctly folded and functional proteins persist to prevent cellular toxicity and disease.

3.2. Protein Synthesis – The Ribosome

3.2.1. Composition, Function, and Regulation

Ribosomes are responsible for protein production in all living beings. They can make up nearly 30% of the cellular dry mass. Ribosomes are classified into two subunits: a small subunit and a large subunit. They comprise proteinaceous material and ribosomal RNA (rRNA); hence, they are called ribonucleoproteins.

The function of translation is reading messenger RNA (mRNA) from the genetic code and using it to create peptides. Translation has three major steps: initiation, elongation, and termination.

During Initiation, the ribosome assembles around the start codon of an mRNA strand. In eukaryotes, the small ribosomal subunit, along with initiation factors, binds to the 5’ cap of the mRNA and scans until it finds the start codon, which is typically AUG. Once the start codon is located, the initiator transporter RNA (tRNA) binds to it, carrying the amino acid methionine. The large ribosomal subunit joins this complex, forming a complete ribosome ready for elongation.
During **Elongation**, amino acids are added subsequently to the polypeptide chain. This process takes place in the three catalytic sites of the ribosome, namely A (aminoacyl), P (peptidyl), and E (exit) sites (see Figure 4). A transfer RNA (tRNA) carrying the following amino acid enters the A site and pairs up with the complementary codon in the messenger RNA (mRNA). The ribosome forms the peptide bond between the amino acids present at the A and P sites. Without the amino acid, the tRNA present at the P site leaves the ribosome and moves to the E site. The ribosome then translocates, shifting the tRNA to the A site and the mRNA to the P site, making the A site available for the next aminoacyl-tRNA to enter. This process repeats, elongating the polypeptide chain with each cycle (see Figure 4).

**Figure 4 Elongation of the Polypeptide Chain during Translation:** This cartoon depicts a ribosome’s elongation phase of protein synthesis. **(Left to right)** During translation, a tRNA molecule **(Pink)** carrying an amino acid **(Orange)** pairs with the corresponding codon on mRNA. The ribosome catalyzes the peptide bond formation, linking the amino acid to the polypeptide chain. The ribosome then moves the tRNA from the A site to the P site along with the mRNA, elongating the polypeptide chain, and eventually, the tRNA is released at the E site. This process continues until a complete polypeptide chain is formed.

Translation ends when a stop codon (UAA, UAG, or UGA) on the mRNA moves into the A site. Stop codon signals **Termination** without coding for an amino acid. After termination, release factors bind to the ribosome, and a newly synthesized polypeptide chain is released. Finally, the ribosome subunits and the mRNA and tRNA disassemble, completing the process.

This translation process is costly, with the synthesis and elongation of each amino acid requiring around 30 ATPs. Therefore, this process needs to be tightly regulated, and much of this energy can be saved by downregulating or completely shutting down translation. Complete
shut-down is specifically desired in nutrient-scarce environments and is where hibernation factors come into play (Projects 3 and 4).

Hibernation factors can shut down intact ribosomes instead of degrading them, which is cost-saving and allows for quick recovery when needed\textsuperscript{107,112}. This process is vital for organisms with an environmental dormant stage like microsporidia\textsuperscript{113}.

Overall, translation is highly regulated and energy-intensive. The process is highly efficient and stays conserved in its core mechanisms across the tree of life.
3.2.2. The Eukaryotic Ribosome

Although all known living organisms produce ribosomes with conserved core functions, the ribosome composition varies significantly between clades\textsuperscript{114}.

Eukaryotic ribosomes have nearly 30 additional proteins and numerous expanded rRNA elements compared to prokaryotic ones\textsuperscript{115}. These additional rRNA elements, known as expansion segments (ES), stabilize the additional protein layer\textsuperscript{116}. They expand the original core by adding further outer shells, which support the recruitment and organization of components of the eukaryotic ribosome biogenesis process\textsuperscript{117}. In addition, some segments recruit regulatory factors\textsuperscript{118}, while others recruit and stabilize mRNA during translation\textsuperscript{119}.

Expansion segments (ESs) are a unique characteristic of eukaryotic ribosomes located in specific, conserved positions across different eukaryotic species. Bacterial rRNAs lack ESs, with some exceptions in Archaea\textsuperscript{117,120}. Despite their importance, the full scope of the functions and implications of ESs is not yet fully understood. Microsporidia have reduced their ESs while keeping the newly acquired eukaryotic ribosomal proteins\textsuperscript{63}. The reductive evolutionary behavior of microsporidia may provide insights into the adaptation and function of eukaryotic ribosomes. (Project 3)

![Figure 5 Addition of Expansion Segments during Evolution: Ribosomal evolutionary progression from bacteria to higher eukaryotes. The addition of expansion segments (ES) in ribosomal RNA increases the complexity and functionality of eukaryotic ribosomes. These segments stabilize additional protein layers and recruit biogenesis and regulatory factors but are not yet fully understood.](image-url)
3.3. Protein Degradation – The Proteasome

The proteasome is an essential protein complex crucial in maintaining proteostasis by degrading unfavorable proteins\(^{121}\). It is tightly regulated and controls various cellular processes, including the cell cycle, transcription, signaling, trafficking, and protein quality control\(^{122}\). Any malfunction in proteasomal degradation can lead to multiple human diseases, such as cancer and neurodegeneration\(^{123}\).

The two main proteasomal complexes are the 20S and 26S proteasomes. The 20S proteasome is the core of this complex and is mainly responsible for breaking down proteins. It has a barrel-shaped structure with four stacked rings, each comprising multiple subunits. The inner rings contain the catalytic sites that are necessary for protein degradation. On the other hand, the 26S proteasome is a larger assembly formed by combining the 20S core particle with one or two 19S regulatory particles. The 19S regulatory units play a crucial role in recognizing ubiquitin-tagged substrates by unfolding them and helping their translocation into the 20S core for degradation\(^{124}\).

**Figure 6 Proteasome Mechanism:** The proteasome can be either in a 20S or 26S state while inactive. The 19S regulatory cap of the proteasome recognizes ubiquitin-tagged proteins and unfolds them with ATPase activity. Deubiquitinating enzymes associated with the 19S particle cleave the ubiquitin chains before the protein can be translocated into the 20S proteasome. Inside the 20S proteasome, catalytic $\beta$ subunits execute the proteolytic degradation of the proteins into smaller peptides, which are then released into the cytoplasm.

No longer needed or dysfunctional proteins are marked for degradation through the ubiquitin-proteasome system (UPS)\(^{124}\). The 19S regulatory cap recognizes ubiquitin-tagged proteins and has several subunits that
bind to ubiquitin chains and have ATPase activity. The ATPase activity is essential for unfolding substrate proteins and their transport into the core particle. Before the translocation of substrate proteins into the core particle, deubiquitinating enzymes (DUBs) associated with the 19S regulatory particle remove the ubiquitin chains from the substrate proteins. The ATPase activity of the 19S particle then unfolds the substrate protein and transports it into the core particle (20S proteasome). The core particle contains catalytic β subunits, which carry out the proteolytic activities of the proteasome. After degradation, the small peptides are released from the proteasome into the cytoplasm or nucleus of the cell. These peptides can be further broken down into amino acids by other proteases or serve as antigenic peptides presented by MHC class I molecules in the immune response. This selective degradation mechanism ensures the timely breakdown of proteins and prevents the accumulation of damaged or misfolded proteins in the cell\textsuperscript{125}.

The proteasome is integral to the protein life cycle and works closely with the ribosomal machinery to maintain cellular proteostasis\textsuperscript{126}. Proteasomes are highly abundant in cells, while their levels are tightly regulated. Due to their central role and abundance, they are an attractive target for drug development. Proteasome inhibitors are particularly promising in wildly divergent and unique pathogens like microsporidia. Microsporidia have also reduced the size and composition of their proteasomes, similar to their ribosomes, and proteasome inhibitors have become a recent focus for anti-microsporidian drugs\textsuperscript{127} (see more in Project 4).
4. The Tools to Investigate Protein Function and Mechanism - Structural Biology

4.1. General Introduction Structural Biology

Structural biology is a field of study that focuses on investigating the molecular structure of biological molecules. It employs techniques and principles from molecular biology, biochemistry, and biophysics to investigate the structure-function relationship of macromolecular complexes.

The origins of structural biology can be traced back to the early 20th century when X-ray crystallography was used for the first time to solve biological samples like myoglobin and hemoglobin. This technique determines the three-dimensional structures of molecules by passing X-rays through a crystal of the molecule and analyzing the resulting diffraction patterns. This method laid the foundation for modern structural biology.

Over time, various methods have been developed and improved to study the structure of proteins and nucleic acids. One such method is Nuclear Magnetic Resonance (NMR) spectroscopy. NMR shines in investigating the dynamic properties of proteins and nucleic acids in solution. It is advantageous in studying the structure and dynamics of small to medium-sized proteins.

Cryo-electron microscopy (cryo-EM) is a crucial method in structural biology. Biological samples frozen in ice are imaged using an electron microscope. Cryo-EM has transformed the field by enabling the determination of high-resolution structures of large and complex molecular assemblies that might be difficult or impossible to crystallize. Single-particle cryo-EM, in particular, has become a crucial method for determining the structure of such complexes. On the other hand, Cryo-electron tomography (cryo-ET) collects images from various angles to produce a 3D representation of large molecular assemblies within the cell in their native environment. This technique provides an exclusive perspective of structural features from the size of complete cells and with the development of Subtomography in recent years down to single proteins at mid-resolution range.

Structural biology employs various methods to understand the structure and function of biological macromolecules. Each method has strengths
and limitations, and a combination of techniques is often used to obtain a comprehensive understanding. With the ongoing development of these techniques and the advent of computational tools for data analysis and modeling, structural biology continues to expand its frontiers.
4.2. The Central Methods Employed in This Thesis

4.2.1. Protein Production and Purification

A common use case for protein production and purification is isolating a single protein or complex to examine its function using biochemical, molecular, and structural biology techniques. Before the 1970s, proteins had to be extracted from mainly endogenous sources. This was typically challenging due to the low levels of the protein of interest and the complex heterogeneity of the lysate. A limited amount of techniques are available to purify those proteins that exploit, for example, their natural traits. Such methods take advantage of, for example, their charge (Ion Exchange Chromatography), their size and shape (Size Exclusion Chromatography), or their density (centrifugation).

With the advancement of recombinant DNA technology in the late 1970s, these processes have become more efficient and controllable. Recombinant expression involves introducing the gene that encodes the protein of interest into an expression system, such as bacteria (E. coli), yeast, insect cells, or mammalian cells. This process aims to produce the desired protein in waste amounts and in an easily purifiable way. An advantage of recombinant expression is the ability to introduce tag sequences in the DNA sequence that is introduced into the host expression system. These tags can facilitate easier purification, improve solubility, aid in protein detection, and provide sites for subsequent attachment of molecules or substrates. Common tags include affinity tags such as polyhistidine (His-tag), which binds to metals like nickel or cobalt in immobilized metal affinity chromatography (IMAC), and the glutathione S-transferase (GST) tag, which can be used in affinity purification leveraging glutathione sepharose beads. Solubility tags, such as the maltose-binding protein (MBP), can enhance the solubility of proteins otherwise prone to forming insoluble aggregates. Additionally, epitope tags like the FLAG or c-Myc tag can be used for purification and detection with antibody-based methods. While those advantages of recombinant expression over endogenous purifications are still valid depending of the research question or for model organisms where genetic tools are either limited or unavailable (like microsporidia). Since the development of new genetic tools like CRISPR-Cas enabled the introduction of markers and tags into the endogenous source the technique has become more viable again.

Purifying membrane proteins presents additional challenges due to their hydrophobic nature. These proteins often require detergents or the
incorporation into lipid-based particles such as nanodiscs or micelles to maintain their structure and functionality. Detergents help to solubilize membrane proteins by mimicking the lipid bilayer environment. In contrast, nanodiscs provide a more native-like lipid environment, which can be essential for less stable proteins or to catch specific conformations\textsuperscript{140}.

In summary, protein production and purification involve strategic decisions and techniques tailored to the specific molecule of interest. Combining recombinant expression systems and sophisticated purification methods has dramatically advanced our ability to study proteins in detail. These advancements contribute significantly to our understanding of protein structure, function, and their roles in biological processes and diseases.

\textbf{4.2.2. Cryo-EM:}

Electron microscopy (EM) has been a crucial tool in exploring the ultrastructural world, impacting various scientific disciplines. The first electron microscope was developed in the 1930s by Ernst Ruska and Max Knoll\textsuperscript{141}. The technique exceeded the resolution limits imposed by the wavelength of visible light in traditional microscopes. Ruska's pioneering work marked a new era in microscopy, offering unprecedented resolution that revolutionized the field of physics by enabling the visualization of atomic structures\textsuperscript{142}.

Adapting electron microscopy for biological samples was challenging due to the high vacuum environment and high energy of electron beams which induced damage\textsuperscript{131}. Between 1932 and 1934, Ladislaus Marton from Brussels adapted an electron microscope and used it to create the first image of a biological specimen. The sample was a 15-µm thick leaf from a sundew plant impregnated with osmium tetroxide\textsuperscript{143}. This technique would be further advanced in the 1940s and 1950s into still-used modern negative stain techniques.

A major breakthrough came with the introduction of modern cryo-EM developed in the late 1970s and early 1980s by Jacques Dubochet\textsuperscript{144}, Joachim Frank\textsuperscript{145}, and Richard Henderson\textsuperscript{146}. Together, they received the Nobel Prize in Chemistry in 2017 for this innovation. Cryo-EM involves flash-freezing biological specimens in a thin layer of vitreous ice, preserving them in a near-native state without the need for staining or dehydration. This technique revolutionized structural biology by allowing the study of biomolecules in their natural conformation, leading to
significant insights into molecular mechanisms underlying health and disease.

The evolution of cryo-EM has been remarkable over the years. Initially, it was primarily used to capture low-resolution images of large molecular complexes. However, with the development of direct electron detectors in the early 21st century, the achievable resolution has dramatically improved\(^1\). Cryo-EM can now resolve structures to the atomic level, providing detailed visualization of intricate molecular architectures, from small proteins to large cellular machinery.

Electron microscopy, from its rudimentary form to the advanced cryo-EM of today, represents a remarkable scientific journey. It has transformed our ability to visualize the microscopic world, from atoms to complex biological systems, and is an indispensable tool in fundamental research and applied sciences.

4.2.2.1. **Single Particle Cryo-Electron Microscopy - SPA**

Single particle cryo-EM allows scientists to visualize biological macromolecules with near-atomic resolution. Cryo-EM Single Particle Analysis SPA involves advanced microscopy, computational algorithms, and bioinformatics.

The protein sample is loaded on a support grid (**Figure 7, panel A**). It is rapidly frozen in a thin layer of glass-like ice, which preserves the natural structure of the biomolecules without the need for fixation or staining (**Figure 7, panel B top**). The vitrification process prevents the formation of ice crystals that can damage the structure and interfere with imaging. Once the sample is prepared, it is loaded into a cryo-electron microscope. Instead of using light, this microscope employs a beam of electrons to illuminate the sample (**Figure 7 panel B bottom**). While traveling through the sample, these electrons interact with the specimen and generate a two-dimensional projection image. However, these images are inherently noisy due to factors such as the low contrast of biological specimens in ice and the need to limit electron dose to avoid radiation damage (**Figure 8, panel A top**). To address this issue, thousands of these noisy images are collected, with many protein particles, each representing a different orientation. While screening for buffer composition, optimal protein concentration, freezing conditions, and sample quality is still essentially a manual process, the final image acquisition implements many automations. It can take several days to complete, depending on the size and complexity of the sample.
Figure 7 EM Grid Preparation and Data Collection: (A) This panel demonstrates the architecture of a cryo-EM grid. It starts with a full view of a 3mm diameter grid. When we zoom in, we can see the grid’s meshwork, and a more magnified view shows individual holes within the mesh that contain the sample. At the highest magnification level, we see a single hole ranging from 1 to 3 micrometers in diameter, where the sample is suspended. (B) (Top) The sequential steps of sample
vitrification for cryo-EM. The process starts with glow discharging of the grid to make the grid hydrophilic. Afterward, the sample is applied to the grid within a controlled environment, followed by blotting to remove excess liquid. Plunge freezing in liquid ethane results in vitrification, ensuring the sample is suspended in glass-like ice. 

(Bottom) Slab view of a frozen grid during data collection, where particles are embedded in vitreous ice. Electrons travel through and interact with the sample to produce micrographs, which are then used for further analysis and reconstruction of the three-dimensional structures of the particles.

The next stage involves image processing and analysis, which requires a lot of computational power. With the help of specialized software, the individual particles from each micrograph are identified and extracted (Figure 8, panel A middle). These particles are then grouped based on their similarity, which helps eliminate any noise. This process is called 2D classification (Figure 8, panel A bottom). This step is critical as it improves the quality of the dataset by enhancing the signal-to-noise ratio as well as impurities in our final dataset. After the alignment process, the images are used to create a three-dimensional (3D) density map of the molecule. This is done through a process called back projection, where the 2D images are merged mathematically to form a 3D representation (Figure 8, panel B top). Advanced algorithms and iterative refinement are used to increase the resolution of the 3D map, allowing researchers to identify finer details (Figure 8, panel B middle).

The final step is model building. During this phase, the atomic model of the molecule is fitted into the electron density map (Figure 8, panel B bottom). Even though huge advancements have been made in the last two years to automate this step further, it is still a largely manual process to ensure an accurate representation of the molecular structure. The resulting high-resolution models provide insights into the biomolecules' atomic arrangement and conformational states, which are crucial for understanding their function and interaction.

The power of single-particle cryo-EM lies in its ability to study molecules in their natural state. In recent years, cryo-EM technology has seen rapid advancements, notably in detector technology and image processing software. These advancements have allowed the visualization of increasingly complex and smaller molecular systems at unprecedented resolutions.
Figure 8 Data Processing and Model Building: (A) The first stage involves extracting individual particles from a micrograph and segmenting them from the noisy background (red circles). The following picture shows the extracted noisy particles, which are aligned with similar particles and are averaged to reduce noise and improve signal clarity. (B) Top to Bottom: Stages involve the 3D reconstruction process of the particles. The back projection technique merges 2D images and infers a 3D structure. The electron density map is then refined to enhance the resolution and accuracy of the 3D structure. Finally, atomic models are built into the electron density map.
SPA is a remarkable technique that combines state-of-the-art microscopy with sophisticated computational methods to turn noisy, low-resolution images into detailed, high-resolution models of biological molecules. This technique has revolutionized our understanding of molecular structures and mechanisms, significantly impacting fields such as drug discovery and molecular biology.

4.3. Experimental vs Computational Structural Biology:

Structural biology is a constantly evolving field that utilizes various experimental and computational methods. Traditionally, X-ray crystallography has been used to determine atomic-level structures. However, the field has rapidly advanced with the introduction of techniques such as NMR, cryo-EM, and computational approaches like AlphaFold. Each method has unique strengths, and comparing them highlights the diverse applications of structural biology.

SPA is a technique that enables the visualization of protein complexes in their near natural environment without the need for crystallization. This is a significant advantage that SPA shares with solution NMR. However, SPA can determine much larger protein complexes compared to solution NMR. Despite its usefulness, the SPA process requires extensive optimization for each specimen, which can consume much time and money. Moreover, the processing of the data obtained poses significant challenges.

Cryo-electron tomography is another application of cryo-EM that provides a three-dimensional perspective of molecules in their cellular context, allowing for a deeper understanding of the structural interplay at a cellular level. This technique, however, is even more demanding, requiring meticulous sample preparation and sophisticated data interpretation. It is also even more time-consuming and technically challenging than SPA.

Similarly, NMR spectroscopy is a crucial tool in structural biology that excels in studying biological molecules’ structure, dynamics, and interactions, especially in their natural states. It can provide atomic-level details of molecules in solution, offering valuable insights into their conformational flexibility and dynamics that other structural approaches often fail to reveal. However, NMR has limitations, including the restricted size range of molecules it can analyze and the requirement for extensive sample preparation and complex data interpretation.
On the other hand, protein structure prediction has seen a significant advancement with the advent of computational methods such as AlphaFold\textsuperscript{151}. These methods utilize deep learning algorithms to predict the 3D configuration of new proteins by analyzing patterns within known protein structures. AlphaFold's predictions have revolutionized the field, providing quick and accurate results. However, it is essential to note that they may only rarely capture the complete dynamics of proteins, their interactions with other molecules, and the conformational changes required for their function.

Combining experimental and computational methods is a powerful strategy for the structural biology workflow. Computational models can provide a starting point for structural hypotheses, which can then be validated or refined experimentally. This interplay is particularly beneficial when studying complex biological systems like infectious diseases, where a deeper understanding of the structural basis of pathogen-host interactions or drug resistance mechanisms is crucial. AlphaFold can be used to quickly identify potential structures and targets, which can then be confirmed using cryo-EM, NMR, or X-ray crystallography to provide the empirical evidence necessary for a comprehensive understanding. This integrated approach accelerates the pace of discovery and enables a deeper level of structural insight, ultimately facilitating advancements in drug development and a better understanding of biological processes.

Therefore, while computational methods have added an invaluable dimension to structural biology, they by no means render experimental techniques obsolete. Instead, the most efficient path forward involves leveraging the strengths of both computational predictions and experimental validations. This collaborative approach is poised to accelerate scientific breakthroughs, driving innovation and deepening our understanding of life at a molecular level.
5. Objectives of the Thesis

The initial chapters of this thesis frame the historical and present-day importance of infectious diseases and introduce our model organisms, *E. faecalis*, and microsporidia. The context underscores the importance of understanding *E. faecalis* and microsporidia on a molecular level. Using cryo-electron microscopy, the research investigates the mechanisms of multi-drug resistance in *E. faecalis* (Project 1) and the unique biology of microsporidia (Projects 2-4). This fills critical knowledge gaps and emphasizes the crucial role of such molecular insights in addressing infectious disease challenges.

**Project 1: A Molecular Basis of Multi-Drug Export in *Enterococcus faecalis***

**Project 2: The Structural Characterization of the Microsporidian Infection Mechanism***

**Project 3: Structural Characterization of a new Microsporidian Ribosome Hibernation Mechanism***

**Project 4: Structural Characterization of the Microsporidian Proteasome – Reduction and Regulation.***
6. Publications and Supplementary Data

6.1. Project 1: A Molecular Basis of Multi-Drug Export in Enterococcus faecalis (partially published)

6.1.1. Background and aim

As introduced, E. faecalis is a concerning multidrug-resistant pathogen involved in hospital-born diseases and the main contributor to its resistance is the transporter EfrCD.

To understand the mechanistic basis for drug resistance, the aim was to solve the cryo-EM structure of the primary drug efflux pump in E. faecalis, the heterodimeric ABC transporter EfrCD, and subsequently use the structural data to decipher the protein function in vitro studies, which our collaborators have done in the group of Dr. Markus Seeger.

6.1.2. Results and Discussion of Supplementary Data

6.1.2.1. Summary of Unpublished Results

EfrCD was shipped to our lab by the Seeger lab either as wild type or as an NBD E512Q active site mutation, and both were available in either micelles or nanodiscs. The frozen aliquots were used for a final purification and binding partner incubation step. The obtained samples were analyzed at the Umeå University EM facility on a Titan Krios equipped with a Gatan K2 electron detector. To assess sample quality conformational homogeneity and to obtain an overall structural model of EfrCD, we have prepared, analyzed, and processed the following samples:

**Sample 1:** EfrCD in nanodiscs
**Sample 2:** EfrCD-EQ in nanodiscs
**Sample 3:** EfrCD-EQ in nanodiscs with nanobody NB193
**Sample 4:** EfrCD in micelles with ATPγS
**Sample 5:** EfrCD in micelles with NB193 (dataset collected in Zuerich)
**Sample 6:** EfrCD in micelles in the presence of NB193 and ATPγS

In all nanodiscs reconstituted samples, one of the two NBDs appears dynamic and cannot be resolved. The EQ mutation (rendering the protein unable to hydrolyze ATP), presence of ATPγS (a non-hydrolyzable analog of ATP), or a raised nanobody NB193 does not stabilize a conformation with both NBDs resolved. In contrast, EfrCD’s nucleotide-binding domains seem less dynamic in micelle-reconstituted proteins. The
addition of ATPγS and a nanobody stabilizes an inward-facing open conformation with both NBDs resolved (sample 6). This sample was used in Publication 1.

**Figure 9 Towards a Stable and Wholly Resolved Cryo-EM Structure of EfrCD:** 8 Å low-pass filtered maps obtained from datasets of EfrCD in different lipid/detergent environments or in complex with ATPγS or/and nanobody NB193. The volume obtained from sample 5 is not shown but represents an intermediate state between conformation A/B and D (nanobody bound, one NBD less well resolved).

6.1.3. Analyzed Datasets and Obtained Conformational States of EfrCD

6.1.3.1. Overall Conformational Differences

A total of three nanodisc reconstituted samples (samples 1-3) and three micelle reconstituted samples (samples 4-6) have been analyzed by cryo-EM (Figure 9). Data collection and processing of grids frozen with EfrCD in nanodiscs resulted in 2 distinct conformations with the NBD of EfrC (conformation A) or EfrD not resolved (conformation B). The transmembrane regions of both conformations are structurally similar despite differences in NBD behavior. In those conformations, the substrate binding pocket is not well accessible. In contrast, the three micelle reconstituted samples (sample 3-6) adopt an inward-facing open conformation (similar to PDB 4Q4A152.), with improved densities for both NBDs and more accessible substrate binding site than in conformation A and B.
6.1.3.2. Nanobody NB193 Supplemented Samples 3 and 6

No density was observed for the nanobody in sample 3 (EfrCD-EQ in nanodisks with nanobody NB193). In the micelle reconstituted sample, the apparent density for NB193 was visible on the extracellular site of EfrCD. A micelle environment in combination with ATP\textsubscript{γS} and NB193 appears to stabilize EfrCD’s nucleotide-binding domains. The most complete map was used to build a model of EfrCD (sample 6, conformation D).

6.1.3.3. Cryo-EM Data Collection and Processing of Sample 6 (EfrCD\textsuperscript{wt}:ATP\textsubscript{γS}: NB193)

Detailed data collection and processing information can be found in the method section of Publication 1.

6.1.4. The Structure of EfrCD:ATP\textsubscript{γS}: NB193

6.1.4.1. Overall Structure

The general conserved structural features of heterodimeric ABC exporters can also be observed in EfrCD. The transporter comprises two subunits, EfrC and EfrD (Figure 10). The structure contains twelve TMD-helices, six for each monomer, two NBDs, two elbow helices at the N termini, and two coupling helices interacting with the NBD of the other subunit.

Figure 10 The Cryo-EM Structure of EfrCD:ATP\textsubscript{γS}: NB193 at 4.25 Å Resolution: (A) Four 90° related views of the cryo-EM map of EfrCD colored according to subunit (EfrD; purple, EfrC; cyan, NB193; yellow). An additional unexplained density is colored in red.
An additional density colored in red (Figure 10) was observed below the N-terminal elbow helix of EfrD. This density is only present in EfrCD reconstituted micelles and the presence of ATPγS and NB193 and presumably represents a bound small molecule. The resolution in this region did not allow for the identification and placement of a model. A similar density has also been observed in the structure of the ABC importer YbtPQ and might represent a copurified lipid\textsuperscript{153}.

Figure 11 Additional Densities: The nucleotide-binding site (orange) of EfrC (cyan), EfrD (purple), NB193 (yellow), and an additional density (red) below the elbow helix are indicated in an 8 Å low-pass filtered map of EfrCD sample 6.

6.1.5. Conclusion of Supplementary Data

In conclusion, the supplementary data shows that the nucleotide-binding domains (NBDs) display incredible flexibility, which is likely crucial to the functioning of ABC exporters. Understanding the transport mechanisms involved requires an investigation of the dynamic nature of NBDs. Based on observations of NBDs in various interaction states, the transporter’s activity appears to be linked to a complex choreography. Additionally, extra density beneath the elbow helix requires closer examination. Although the theory that these densities may be lipids is compelling, confirming this could lead to further mechanical knowledge.

Conducting a comprehensive investigation into the flexibility of NBDs and the mysterious density beneath the elbow helix could enhance our
understanding of the ABC transporter function. This investigation could uncover new regulatory mechanisms or therapeutic targets within these critical cellular components.

6.1.6. **Summary of Publication**

Our publication presents a detailed analysis of the bacterial ABC transporter EfrCD using deep mutational scanning (DMS). The study aimed to determine the drug efflux activity profile of over 1,430 single EfrCD variants and uncover the relationship between the transporter's structure and function.

We showed that introducing negative charges at different locations within the large substrate binding pocket significantly increased efflux activity towards positively charged ethidium. However, additional aromatic residues, which historically have been linked to transporter polyspecificity, did not show the same effect. Furthermore, specific substitutions resulted in Hoechst influx while not affecting the efflux activity for ethidium and daunorubicin. This demonstrated that single mutations could convert EfrCD into a drug-specific ABC importer.

In the context of the inward-facing cryo-EM model of EfrCD, the data analysis uncovered a high-affinity drug binding site that releases bound drugs through a peristaltic transport mechanism as the transporter transitions to its outward-facing conformation.

The research provides valuable insights into understanding the structure-function relationship of EfrCD and its potential applications in drug resistance and efflux activity.

6.1.7. **Additional Methods for Publication**

6.1.7.1. **Sample Preparation**

All samples were produced according to Lea M. Hürlimann *et al*. 2016. Sample preparation for EM studies, as well as processing, was performed according to Meier, G. *et al*. 2023.
6.2. Project 2: Structural and Functional Characterization of the Microsporidian Polar Tube Architecture

6.2.1. Background and Aim
As introduced in the chapter “A Unique Fungal Eukaryotic Parasite – Microsporidia,” microsporidia are an ideal model organism for comparative evolutionary structural biology due to genome reduction and the resulting compaction of various conserved macromolecules (Projects 3 & 4). However, despite minimizing conserved macromolecules or biochemical pathways, microsporidia also acquired new molecular mechanisms to adapt to a parasitic lifestyle. One of the most peculiar features of microsporidia is their mode of infection, which involves a long filamentous tube called the polar tube, as introduced in “A Specialized Infection Apparatus – The Polar Tube.”

Initially, I aimed to reconstitute, purify, and analyze PTP1, 2, and 3, which are the significant elements of the elongated part of the polar tube (Figure 1 B). The initial results for PTP1 and 2 were promising, but several considerable hurdles remained. Due to the closure of the lab and the change in principal supervisor, I focused on PTP3, which had the highest chances of success. Vairimorpha necatrix PTP3 is a 151 kDa (1’353 amino acids) protein with an N-terminally predicted and reported 22 amino acid long signal sequence. The protein is expected to consist of two alpha-helical and one beta-sheet domain, with the N- and C-termini largely unstructured (Figure 12 A). V. necatrix was mainly chosen for its availability in the lab.

Initially, this chapter will address the structural and functional work centered around PTP3. A summary will follow up on my contribution to the endogenous extracted polar tube manuscript, which has just been accepted.

6.2.2. Results and Discussion of Supplementary Data

6.2.2.1. Purification of PTP3 Constructs
Bacterial overexpression constructs for full-length PTP3 and several truncated versions were cloned (Figure 12 E). Constructs were designed according to secondary structure predictions, signal sequence predictions, and AlphaFold models.
Figure 12 PTP3 Prediction, Cloning, and Purification: (A) AlphaFold prediction of the full-length V. necatrix PTP3. The different domains are color-coded. (B)-(D) SDS pages of the protein purifications are explained in detail in the method section and in the chapter “Purification of PTP3 Constructs”. (E) Library of designed expression plasmids. Domains are colored according to the AlphaFold prediction. In this summary, the discussed constructs are marked with blue boxes.

Expression tests in Escherichia coli yielded no protein production for full-length PTP3, while truncated versions of PTP3 were successfully produced and purified (Figure 12 B-D). This suggests that some removed regions contain elements preventing proper protein expression in E. coli. Purification of H14-3C-VnPTP3_201-1’175 revealed that degradation
occurred during production. Furthermore, size exclusion chromatography indicated that PTP3 forms different complexes, most likely with truncated versions of itself. A new construct was designed to investigate this behavior and optimize the purification protocol: GFP-3C-VnPTP3_His_201-1’175. This construct was utilized for a sequential purification protocol, first using the N-terminal GFP and later the C-terminal His tag. Several truncated versions of PTP3 remained in the final sample. N- or C-terminal degradation should result in losing at least one purification tag. This means that no truncated versions should stay after the sequential pull-down. Unexpectedly, this was not the case (Figure 12 C). This could have two explanations: 1. The protein degrades from the middle of the protein while the domain architecture prevents the protein from falling apart. 2. PTP3 forms complexes with its truncated versions and pulls them down. As for now, explanation 2 seems more likely, but this remains to be investigated.

So far, the full-length construct (which did not lead to protein production) and constructs with a large part of their N- and C-terminal regions removed (predicted unstructured region) have been discussed. A construct was designed to investigate further the possibility of producing a more complete and highly pure protein sample. Only the nucleotides encoding the first 22 amino acids were removed, as they were predicted to be part of a signal sequence. An initial overexpression of this plasmid in E. coli showed that removing the potential signal sequence enabled the production of intact protein (Figure 12 D). However, a later western blot also revealed a degradation pattern of this construct.

6.2.2.2. Initial Insight Into Endogenous Purified PTP3

Apart from recombinant expression strategies, the possibility of endogenous protein purification has also been explored. V. necatrix PTP3 (VnPTP3) has seven consecutive Histidines in the N-terminal region (His81 – His87), which we identified and used as an affinity tag for purification (Figure 13 B&C). Initial experiments were promising as a band of roughly the expected size was observed, and only a limited number of additional protein bands were observed. Mass spectrometry of the sample confirmed the presence of PTP3, and further interaction partners were observed (see Chapter “Summary of Publication” and Publication 2).

Because of limited access to spores, the amount of PTP3 from the endogenous source was too low for extensive protein purification.
Therefore, the sample was used for initial structural characterization directly after the IMAC purification.

6.2.2.3. Structural Characterization of PTP3

Initial analysis using negative stain (NS) electron microscopy (EM) revealed a rather heterogeneous sample (Figure 13 B). While some particles seem to have the expected size according to the mass of PTP3 and its AlphaFold predictions (Figure 12 A), others are much larger. Pulldown and mass spectrometry have already confirmed that we would expect higher-order complexes. Still, due to the low sample amount, we could not use methods like SEC, which would have allowed us to sort the potentially different complexes into separate fractions to obtain more homogeneous samples. Therefore, we prepared cryo-EM grids with the sample at hand.

Ice quality and thickness allowed for a small data collection even though the overall particle number per micrograph was low. This allowed us to generate initial 2D classes to understand the sample's particle shape quality and usability (Figure 13 C: right). Even though some of the classes look very promising, the particle number needed to be higher to draw conclusions with high confidence.

I prepared a new sample according to the same procedure to get cryo-grids with more particles per micrograph. The main difference was that continuous carbon cryo-grids were used for freezing this time. Those grids have a thin layer of carbon throughout the grid. Since the particles showed some affinity to carbon, the plan was to increase the number of particles on the grid. Unfortunately, this carbon layer comes at the cost of having a worse signal-to-noise ratio, making processing generally more challenging. Screening and a small dataset collection show that particles did not increase significantly. Interestingly, the classes seem different, and we obtained more globular particles than before (Figure 13 C). However, it is difficult to draw any conclusions since the number of particles is too low, and the noise, due to the continuous carbon, makes it difficult to align the particles properly.
Figure 13 Purification of PTP3: (A) The purification method is represented in Publication 2 (B)-(C) left to right: SDS page of the elution fraction used for electron microscopy, representative micrograph of negative stain screening, representative micrograph of cryo-EM screening, and representative 2D Classes of the initial processing.

6.2.3. Conclusion of Supplementary Data

While the supplementary data on polar tube proteins (PTPs) may not have produced any conclusive results, it is crucial that we recognize the significance of these trials. They represent a critical stepping stone towards further research to enhance our understanding of the structural and functional properties of microsporidian PTPs. This includes not just PTP3 but also other significant elements like PTP1, PTP2, PTP4, and PTP5.
A crucial aspect will be exploring alternative expression systems or modifying the current expression and purification strategies for the full-length PTP3 to produce a non-degraded sample. Additionally, it's essential to examine the effect of post-translational modifications (PTMs) on PTP3. Mainly to investigate how these modifications influence the protein's stability and functionality, but also how they might affect the whole PT complex either during assembly or in its assembled stage.

Finally, it's vital to advance the structural characterization of endogenous PTP3. Improved sample preparation methods that increase particle numbers and achieve higher resolution structures will offer valuable insights. Such an approach could reveal more about the heterogeneity observed in the PTP3 sample and the formation of higher-order complexes and may lead to a sample usable for structural characterization.

### 6.2.4. Summary of Publication

In this research paper we investigated the dynamics and ultrastructure remodeling of the polar tube in microsporidia during germination. We used light microscopy and cryo-electron tomography to visualize the polar tube's multiple ultrastructural states and characterize its germination's kinetics. We also observed two distinct states of the polar tube: a cargo-filled state and an empty post-translocation state. We found that the PTPs form two layers, with the outer layer being thicker in the empty state than in the cargo-filled state. Additionally, we performed proteomic analysis of the endogenously affinity-purified polar tubes using a native affinity tag that we discovered. We identified potential protein-protein interactions of the PTPs and found that PTP3 interacts with other PTPs and Ricin-B lectin domain-containing proteins (RBLs). These findings provide comprehensive data on the structure and composition of the polar tube in microsporidia during germination and shed light on the mechanisms of polar tube remodeling and cargo delivery.

Overall, the research contributes to a better understanding of the mechanisms involved in cargo delivery by the polar tube and its compositions, as well as the structural dynamics of the microsporidian infection apparatus during host cell invasion.
6.2.5. Methods not Described in Publication.

6.2.5.1. *Cloning of Overexpression Plasmids*

Constructs for full-length VnPTP3 (1-1’354), VnPTP3-20-1’179, VnPTP3-201-1’175, and VnPTP3-22-1’354 have been cloned into IPTG inducible pRSFDuet vectors with a variety of N-terminal and C-terminal tags and a kanamycin resistance cassette. A representation of some prepared plasmids is summarized in Figure 12 D. Here, the discussed constructs are highlighted in blue.

6.2.5.2. *Production and Purification of PTP3.*

The expression vector RSF_H14-3C-VnPTP3 for the whole-length protein was transformed into *E. coli* Rosetta (DE3) cells via heat shock. Positive strains were selected overnight at 37°C on LB-Kanamycin plates, and a single colony was picked to start an LB-Kanamycin preculture. The preculture was grown shaking overnight at 37°C and then introduced into a 5 L baffled Erlenmeyer flask containing 2 L of LB-Kan. The expression culture was grown shaking overnight at 37°C and then the temperature was changed to 28°C, induced with 0.5 mM IPTG, and then grown overnight at 28°C. Culture was harvested the next day at 6’200 g for 10 min at 4°C, washed with 50 mM Tris-HCl pH8 (4°C), 250 mM NaCl, and centrifuged at 4’000 g for another 10 min. The remaining liquid was discarded, and the obtained pellet was flash-frozen in liquid nitrogen and stored at -80°C until further use.

The pellet was dissolved in a Lysis Buffer (50 mM Tris-HCl pH 8 (4°C), 250 mM NaCl, 30 mM Imidazole, 2 mM DTT); furthermore, 5 mM MgCl₂, DNase I, and a protease inhibitor mix were added. The solution was then sonicated on ice for 13 minutes with an amplitude of between 35% and 42% and a 3-second long pulse with 9-second breaks in between. The homogeneous lysate was centrifuged for 1h at 48’000g at 4°C to separate the soluble fraction.

The soluble fraction was loaded onto a 5 mL Ni Hi load Column (GE) and washed with 3 times 10 Column Volumes (CV) Wash Buffer (50 mM Tris-HCl pH 8 (4°C), 150 mM NaCl, 2 mM DTT) supplemented with an increasing Imidazole concentration (50 mM/80 mM/100 mM) and then eluted with elution buffer (50 mM Tris-HCl pH 8 (4°C), 250 mM NaCl, 250mM imidazole, 2mM DTT) in 2 mL fractions. Intermediate purification steps and elutions were loaded on an SDS-PAGE gel.
6.2.5.3. **Further Purification of VnPTP3_201-1’175.**

Initial purification was performed according to “Production and Purification of PTP3”. Elution samples containing protein were pooled, and 90 µL 3C Protease [2 mg/mL] was added to the 14 mL elution samples and was dialyzed overnight at 4°C against 600 mL lysis buffer. 2M imidazole was added to the sample up to a concentration of 40 mM and then loaded on a 5 mL Ni Hi load Column (GE). Flowthrough fractions were immediately collected. Re-eluting the column showed that only minimal amounts of sample remained on the column. 500 µl of sample were centrifuged for 10min at 4°C loaded on a Superose6I 10/300 for size exclusion using SEC Buffer I (50 mM Tris-HCl pH 8 (4°C), 150 mM NaCl). 500 µl fractions were collected and loaded on an SDS page gel.

6.2.5.4. **Further Purification of VnPTP3_201-1’175.**

Protein production, cell lysis, and separation of the soluble and insoluble fraction were performed according to “Production and Purification of PTP3”. However, instead of directly performing a Ni-pulldown, 50 µl of in the lab-produced GFP beads were added and incubated, rotating overnight at 4°C. The next day, the beads were washed 4 times with 10 mL of SEC Buffer I. The protein was then eluted at 4°C for 1 hour with 3 times 50 µl SEC Buffer I with an addition of 3 µl 3C protease. Following that procedure, a Ni-pull down was performed using previously described buffers.

6.2.5.5. **Negative Stain**

Copper grids coated with a thin carbon film were glow-discharged for 30 sec at 15 mA. 4µl of elution fractions (at different dilutions in the final wash buffer) were applied and incubated at RT for 30s. The sample was removed via filter paper, and the grid was washed 3 times with ddH2) water droplets. The water was removed with filter paper, and the grid was passed with one drop of 1.5% uranyl-acetate. It was immediately removed and incubated at RT with a second drop for 35 seconds. Uranyl-acetate was removed, and the grid dried and stored for imaging on an FEI Talos L120C transmission electron microscope (Thermo Fisher Scientific), operating at 120 kV, and equipped with a Ceta 16M CCD camera.

6.2.5.6. **Grid Freezing and Data Collection**

A Quantifoil R 2/2 +2nm C Cu 300 grid (Quantifoil Micro Tools, Prod. No. 658-300-CU) was glow-discharged for 30 seconds at 15mA before adding a 4-µl endogenously purified sample. Grids were then flash-frozen in liquid ethane using an FEI Vitrobot (Thermo Fisher Scientific) set to 100%
humidity, 4°C, blot force of −5, waiting time of 1 second, and blot time of 3.5 seconds. Micrographs were collected at the Umeå Core Facility for Electron Microscopy on Glacios (Thermo Fisher Scientific) operated at 200kV, equipped with a Falcon 4i Direct Electron Detector (Thermo Fisher Scientific). EPU (Thermo Fisher Scientific) was used for the automated data collection.

**6.3. Project 3: Structural Characterization of the Microsporidian Proteasome – Reduction and Regulation.**

**6.3.1. Background and Aim**

Proteasomes are expressed in all known eukaryotes, and although they conserve nutrients by recycling amino acids, degrading undesirable and misfolded proteins is energy intensive\(^\text{156}\), for organisms like microsporidia, which lost the ability to generate ATP using the tricarboxylic acid cycle and oxidative phosphorylation\(^\text{157,158}\), and which are almost entirely reliant on their hosts for ATP\(^\text{57}\), tight regulation of this process is essential. One mechanism by which this can be achieved is through interactions with inhibitory proteins\(^\text{159,160}\).

The 31-kDa proteasome inhibitor PI31 is a conserved eukaryotic protein that inhibits 20S proteasome activity and the ATP-dependent assembly of human 26S proteasomes\(^\text{159-164}\). Our work co-discovered and demonstrated that PI31 binds to the internal cavity of the proteasome and inhibits all six active sites simultaneously.

![Figure 14 The Cryo-EM Structure of the Microsporidian Proteasome, Isolated from Spores or Sporoplasms](image)

*Slab-view of the full proteasome and zoom-in sections of the central proteolytic cavities. The peptide-like density (red), corresponding to PI31L, is present in the spore (left) but absent in the sporoplasm-isolated proteasome (right). The central circular inlet shows the difference in*
density (green) between 4-Å low-pass-filtered cryo-EM maps from spores and sporoplasm proteasomes.

6.3.2. Summary of Publication

In this publication, we explored the proteasomes of microsporidia by presenting the cryo-electron microscopy structures of the microsporidian 20S and 26S proteasome isolated from dormant or germinated Vairimorpha necatrix spores.

The microsporidian proteasomes underwent significant reductive evolution, resulting in the loss of at least two regulatory proteins and the compaction of nearly every subunit. This highlights the unique reductive evolution and biology of these medically and economically important pathogens. The study identifies PI31-like peptides occluding the active sites of the microsporidian proteasome, indicating reduced activity in the dormant spore stage and reactivation in nutrient-rich conditions. The highly derived structure of the microsporidian proteasome and the minimized version of PI31 presented in this study reinforce the feasibility of developing specific inhibitors and provide insight into these pathogens’ unique evolution and biology.

We investigated the conservation and variation in the 20S proteasome active sites and the gate access mechanism in microsporidia, revealing a different closed-gate conformation compared to that of yeast proteasomes.

In conclusion, the study highlights microsporidia’s unique proteasome structure and function. It highlights the potential for developing specific inhibitors and sheds light on the evolution and biology of these medically and economically important pathogens. The research has important implications for understanding proteostasis in microsporidia and the potential development of therapeutics targeting their proteasomes.
6.4. Project 4: Structural Characterization of a new Microsporidarian Ribosome Hibernation Mechanism

6.4.1. Background and Aim

Extended rRNA elements called expansion segments are a critical difference between prokaryotic and eukaryotic ribosomes. Most ribosomal ESs are involved in stabilizing an additional protein layer. In contrast, others have been proposed to provide an interaction surface for regulatory factors or assist in ribosome biogenesis. At the same time, eukaryotes' evolutionarily acquired ESs and microsporidia tend to remove those again. In the case of Vairimorpha necatrix, this resulted in the smallest known eukaryotic cytoplasmic ribosome. However, it remains unknown how microsporidia compensate for the loss of those seemingly essential ESs.

Ribosome biogenesis and protein translation are energy-intensive processes. Therefore, organisms operating under nutrient limitations depend highly on energy conservation via ribosomal hibernation and recycling. However, the mechanisms of hibernation in microsporidia are poorly understood.

6.4.2. Summary of Publication

In this study, we solved the cryo-EM structure of the Pararosea locustae ribosome, bound by the recently discovered and conserved eukaryotic hibernation and recycling factor Lso2. Lso2 binds to the mRNA channel in a V-shaped conformation, blocking multiple functional regions simultaneously and providing a reversible ribosome inactivation mechanism.

By comparing the obtained structure to the Saccharomyces cerevisiae and Vairimorpha necatrix ribosomes, we showed that the P. locustae ribosome retains several ES absent in other microsporidia. Surprisingly, microsporidian ribosomes display a dramatic reduction in expansion segment content. In one case, an expansion segment absent in V. necatrix has been reduced in P. locustae to its most minimal version, a single nucleotide. This single nucleotide acts as an architectural co-factor stabilizing a protein-protein interface.

Our research highlights the reductive evolution in microsporidia and sheds light on understanding ribosomal hibernation. Additionally, we identified a novel mechanism of translational shutdown in extracellular stages of microsporidia, which contributes to a better understanding of
these emerging pathogens. The cryo-EM analysis of the ribosome from *P. locustae* demonstrated an intermediate state of rRNA reduction between yeast and *V. necatrix*.

**Figure 15 Cryo-EM Structure of the Microsporidian P. locustae Ribosome.** A: Overall view of the solved ribosome. The small subunit of the ribosome is presented in yellow, the large subunit in blue, and Lso2 in red. The zoom-in illustrates a single nucleotide, which represents the most minimal version of an ES. B: Lso2 blocks the binding site of the P-tRNA and A-tRNA simultaneously by adopting a V-shaped conformation.
7. Summary of Research Findings & Discussion

This thesis explores the complex mechanisms of infectious diseases and the evolutionary competition between host and pathogen, focusing on *E. faecalis* and microsporidia. The research employs advanced cryo-EM techniques to uncover the molecular basis of multidrug resistance in *E. faecalis*, shedding light on the structure and function of the primary drug efflux pump EfrCD. Additionally, This thesis provides crucial insights into the unique biology of microsporidia, emphasizing the significance of understanding at the molecular level in addressing infectious disease challenges. The research findings reveal the dynamic nature of essential cellular components, such as the proteasome, ribosome, and infection mechanism, offering potential new avenues for therapeutic targets and regulatory mechanisms.

In the broader context, the findings of this thesis contribute to the ongoing narrative of infectious disease research. As outlined in earlier chapters, the history of infectious diseases demonstrates a continuous battle between human innovation and pathogen adaptation. The insights from studying *E. faecalis* and microsporidia add another layer to our understanding of this complex interplay. They highlight the importance of continuous research and adaptation in our approaches to disease management, especially in an era where the rise of multidrug resistance poses a formidable challenge.

In conclusion, this thesis not only provides valuable knowledge to the infectious disease research field but also serves as a call to action. It emphasizes the need for innovative research methodologies, such as the use of cryo-EM, in understanding and combating the ever-evolving threat of infectious diseases. The fight against pathogens like *E. faecalis* and microsporidia symbolizes the broader struggle against infectious diseases and AMRs, which requires a multidisciplinary approach and a global cooperative effort.
8. Acknowledgment

At the end of this four-and-a-half-year journey, I want to express my deep gratitude to the many people who have supported me. Although there are too many to mention, I want to tell you how much your support has meant to me.

To my family, thank you for being my unwavering support system throughout my entire education—your belief in me, even in the face of doubts from others. I wouldn’t be here without you.

I want to give a special thanks to Jonas, who was the reason for my move to the north of Sweden. Although our journey did not pan out as we had hoped, I learned invaluable lessons from you. Ronnie, you started as my co-supervisor and became much more than that - my principal supervisor. Your help during the transitions, amidst the complexities of bureaucracy and lab work, was a huge help. Thank you.

I’m grateful to my old lab family - Himanshu, Nathan, Alex, Léon, Rahel, Lucia, and Dennis. Our collective spirit and the environment we created made it a phenomenal place to work. Cheers to our honorary lab members Michelle, Eliana, and Anna, too. Thanks to all of you for the adventures inside and outside the lab. The same goes for my new lab mates, Anais, Wei-Sheng, Josy, Annika, and Krishna. I felt very welcome, and even though our time together was short, I enjoyed it.

Thanks to my committee, examiners, and co-supervisors for their efforts. Special thanks to the whole EM facility. I learned a lot from you and always had a great time.

Karim, one of my first friends here, living with you and our hangouts have been excellent.

Victor, Selma, Max, and many others, you are great people, and I’m happy that we met.

A big shoutout to the Molbio crew! Our times together, from after-work gatherings to cake club, have been unforgettable. Likewise, to the Medchem folks and our Fridays, Events going out or DnD seasons were terrific.

To my friends back in Switzerland, thanks for still being around after those four years.

A special mention to the Volley folks, Cake people, EM fanatics, Vasa crew, Whiskey conference gang, Skiers, both Andrés, the Rodrigo crew, Hikers, Snowshoe Hikers, Headbangers, music enthusiasts, Heidelbergers, Boardgamers, the old-old-lab, and the River Rafters. Each of you has added unique experiences and memories to my Ph.D. journey

– Thank you all.
9. References


24. Campbell-Lendrum, D., Manga, L., Bagayoko, M. & Sommerfeld, J. Climate change and vector-borne diseases: what are the implications for public health research and policy?


Han, B. *et al.* The role of microsporidian polar tube protein 4 (PTP4) in host cell infection. *PloS Pathogens* 13 (2017). https://doi.org/10.1371/journal.ppat.1006341


https://doi.org/10.1186/s13071-023-05908-9


https://doi.org/10.1111/jeu.12751


https://doi.org/10.1186/s13071-020-04348-2


https://doi.org/10.3389/fpubh.2014.00145


https://doi.org/10.3390/tropicalmed4020088


https://doi.org/10.1016/j.jiph.2021.10.020


https://doi.org/10.3390/pathogens10101310


https://doi.org/10.1038/s41467-021-21749-5


https://doi.org/10.1186/s13756-019-0508-4


https://doi.org/10.3389/fmicb.2016.00788


https://doi.org/10.1126/science.1080613


https://doi.org/10.1128/Aac.00661-16


113 Microsporidia. (2014). [https://doi.org/10.1002/9781118395264](https://doi.org/10.1002/9781118395264)


134 Jackson, D. A. *et al.* Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing Lambda Phage Genes
Helmut Ruska and the visualisation of viruses. The Lancet 355 (2000/05/13). https://doi.org/10.1016/S0140-6736(00)02250-9
https://doi.org/10.1017/S003583500004297


