

Abstract

Title of dissertation: AN INVESTIGATION ON THE MOLECULAR BASIS FOR DIMER FORMATION OF A BACTERIOPHAGE ENDOLYSIN POSSESSING ANTIMICROBIAL ACTIVITY AGAINST *STREPTOCOCCUS PNEUMONIAE*

Adit Bipin Alreja, Doctor of Philosophy, 2023

Dissertation Directed by: Professor Daniel C. Nelson
Institute for Bioscience and Biotechnology Research and
Department of Veterinary Medicine

The global rise of antibiotic resistance casts a shadow on treating infectious disease. An alternative to the use of antibiotics is bacteriophage-derived peptidoglycan hydrolases called endolysins. Endolysins, produced at the end of a bacteriophage replication cycle, cause bacterial cell lysis and virion release. When applied exogenously as recombinant proteins, they are also capable of cleaving the Gram-positive bacterial peptidoglycan. Various studies conducted *in vitro* and *in vivo* showcase the therapeutic potential of endolysins as the next generation of antimicrobials. *Streptococcus pneumoniae* is the most common cause of a variety of infections ranging from otitis media to invasive bloodstream infection (bacteremia) and meningitis (brain infection). While pneumococcal vaccination programs have proven to be effective, the high rates of antibiotic resistance reported for *S. pneumoniae* has led to the CDC classifying it as a “serious” threat. One of the most studied endolysins targeting *S. pneumoniae* is Cpl-1. This thesis represents an investigation into the molecular basis for dimer formation of the Cpl-1 endolysin which displays

antibacterial activity against *S. pneumoniae*. In addition to disproving a long-accepted mechanism of dimerization of Cpl-1 in the presence of choline, we have conclusively identified the residue involved in the formation of the Cpl-1 dimer. Our findings led to the discovery of a novel C-terminal consensus sequence shared by all pneumococcal endolysins that informs their propensity to form dimers in the presence of choline. Next, through a bioinformatics approach we identified a new endolysin containing this C-terminal consensus sequence, produced it, named it SP-CHAP, and showed that it forms a dimer in the presence of choline, indicative of the widespread dimerization phenomenon associated with pneumococcal endolysins. Of interest, SP-CHAP is the first endolysin with antimicrobial activity against *S. pneumoniae* that possesses a cysteine, histidine-dependent amidohydrolase/peptidase (CHAP) domain. SP-CHAP was subsequently characterized for its biochemical and antimicrobial properties and benchmarked against Cpl-1. SP-CHAP is active in all physiologically relevant conditions (pH, temperature) against various *S. pneumoniae* strains and displays no activity towards oral/nasal commensal organisms. This enzyme also displays pneumococcal biofilm eradication ability to a greater extent than Cpl-1, as visualized by confocal microscopy. To further translate the antimicrobial potential of this enzyme, the antimicrobial efficacy of SP-CHAP was validated in a *S. pneumoniae* mouse nasopharyngeal colonization model. Our results demonstrate the therapeutic potential of SP-CHAP as an attractive endolysin to treat *S. pneumoniae* infections and warrant further translational development of this enzyme.

AN INVESTIGATION ON THE MOLECULAR BASIS FOR DIMER FORMATION OF A
BACTERIOPHAGE ENDOLYSIN POSSESSING ANTIMICROBIAL ACTIVITY AGAINST
STREPTOCOCCUS PNEUMONIAE

by

Adit Bipin Alreja

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Professor Daniel C. Nelson, Chair
Professor James Culver
Adjunct Professor Zvi Kelman
Assistant Professor Gilad Ofek
Associate Professor Louisa Wu

© Copyright by
Adit Bipin Alreja
2023

Reflections

Ours is an ecosystem that is ever struggling, often peacefully, for space and survival within itself. Many a creature is driven by assuming unto itself the maxim “might is right” over peaceful coexistence. It falls to the lot of us biologists to unravel the energetic forces that drive life in all permutations of sizes and shapes, as they balance their intrinsic activities with virulent battles for species survival.

Nowhere is this more pronounced than in the relationship between humans and microbes. Our scientific forefathers such as Edward Jenner and Alexander Fleming, standing on the shoulders of Antoni van Leeuwenhoek, had an uncanny knack of brilliance, observational insight, and serendipity to combat microbial aggression.

Nevertheless, Lord Tennyson, with apologies to him, seemed to pen an ode to the lowly microbe, both virus and bacterium:

*“And out again I curve and flow
To join the brimming river,
For men may come and men may go,
But I go on for ever.”*

At no time in recent memory has the virulence of a microbe been so threatening to engulf humankind as during the pandemic of Covid-19.

However, not every virus spews its venom only against humankind. In his Nobel Banquet speech, Sir Alexander Fleming narrated isolating a contaminating mold and christening it ‘penicillin’, using the words “I had a clue that here was something good”. Coincidentally, our Nelson research team of biologists is committed to the optimistic song from Sweden “*something good in everything I see, I believe in angels...*”

We recognize that endolysins are that “something good” from a bacteriophage virus, our source of hope in combating bacterial infections.

Notwithstanding the poetic license, this is in essence the beginning of my tryst with the quest that this dissertation describes. It has been a most pleasurable and rewarding, albeit sometimes arduous, scientific journey motivated and inspired by

*“Knowing is not enough; we must apply.
Willing is not enough; we must do.”*

References:

Alfred Lord Tennyson, *The Brook*, 1886.
<https://www.nobelprize.org/prizes/medicine/1945/fleming/speech>
Benny Andersson and Björn Ulvaeus (ABBA), *I have a dream*, 1979.
Johann Wolfgang von Goethe, 1749-1832.

Acknowledgements

I would like to thank my advisor Dr. Daniel Nelson for his incredible guidance and support throughout my graduate career. My experiences and skills developed in the Nelson lab have made me the scientist I am today, and for that I am eternally grateful. His supervision, patience and generosity has had an everlasting impact on me and will continue to do so as I move forward in my scientific career.

I would like to thank my thesis advisory committee, Drs. Zvi Kelman, Louisa Wu, James Culver and Gilad Ofek for their support, helpful discussions and constructive feedback throughout my graduate school career.

Thanks to our collaborators both here and around the world. I am thankful to Dr. Norberto Gonzalez-Juarbe for his time, guidance and assistance with my research. Dr. Osnat Herzberg has provided invaluable insights and guidance. I would also like to thank Dr. Kinlin Chao for her help with accomplishing key experiments.

I would like to thank various lab members, both current and past, Harley, Hang, Cynthia, Irina, Xiaoran, Niels, Daniel and Urmil for their assistance and meaningful conversations.

Last but not the least, I am grateful to my family - my parents and my sister - for their unconditional love, support and encouragement. I would also like to thank my cousins, nieces and nephews present across India and the United States for helping me in different ways that they have.

Table of Contents

Reflections	ii
Acknowledgements.....	iv
List of Figures	viii
List of Tables	x
List of Abbreviations	xi
Chapter 1: Introduction and Literature Review	1
1.1 Streptococcal Infectious Diseases	1
<i>Streptococcus pneumoniae</i>	5
<i>Streptococcus pneumoniae</i> Virulence Factors.....	6
Choline-Binding Proteins	11
1.2 Antibiotics and Resistance	13
Antibiotic Discovery and Resistance Development	13
Antibiotic Classes and Mechanisms of Action.....	16
Mechanisms of Antibiotic Resistance	19
Antibiotic Resistant <i>S. pneumoniae</i>	25
<i>S. pneumoniae</i> Vaccines	27
1.3 Endolysins – Bacteriophage Derived Peptidoglycan Hydrolases	29
Bacteriophages.....	29
History and Discovery of Endolysins.....	34
Endolysin Domain Structure	37
Endolysins as Antimicrobials/Bacteriolytic Mechanism.....	38
Cell-wall Binding Domains	42
Advantages of Endolysins	42
Biofilms, Synergy, Safety & Immunogenicity	47
Methods to Determine Endolysin Activity	51
Turbidity Reduction Assay	51
Colony Forming Unit Counting Assay	52
Minimum Inhibitory Concentration Assay	52
Nasopharyngeal Colonization Assay	53
Biofilm Eradication.....	53

<i>S. pneumoniae</i> Endolysins	54
Unique <i>S. pneumoniae</i> Peptidoglycan Structure	55
Bacteriophage and Endolysin-Related Applications	56
Chapter 2: Understanding the Molecular Basis for Homodimer Formation of the Pneumococcal Endolysin Cpl-1	61
2.1 Abstract	63
2.2 Introduction	64
2.3 Results	68
The N-terminus of the CWBD is not Involved in Cpl-1 Dimerization but is Required for Catalytic Activity.....	68
Analytical Ultracentrifugation and Antimicrobial Activity of Cpl-1 16-residue C-terminal Deletion Mutant.....	70
Alanine Scanning Mutagenesis of the Cpl-1 Tail.....	74
Cpl-1 Dimerization Depends on Choline Concentration.....	79
Conserved C-terminal β -hairpin Sequence Identifies Pneumococcal Endolysins that Likely Function as Dimers	81
2.4 Discussion	84
2.5 Conclusion.....	88
2.6 Materials and Methods	89
Bacterial Strains and Growth Conditions	89
Cloning and Site-directed Mutagenesis	89
Protein Expression and Purification	89
Microbiology and Lytic Assays.....	90
Analytical Size Exclusion.....	91
Analytical Ultracentrifugation.....	91
Fluorescence Labeling of Proteins	92
Cell Wall Binding Assays and Fluorescence Microscopy.....	92
Bioinformatics and Discovery of Novel Enzymes	92
Chapter 3: SP-CHAP, an Endolysin with Enhanced Activity Against Biofilm Pneumococci and Nasopharyngeal Colonization	94
3.1 Abstract	96
3.2 Importance.....	97
3.3 Observation	98
3.4 Results	101

Biochemical Characterization of SP-CHAP	101
SP-CHAP has Greater Lytic Activity than Cpl-1 and can be Partially Inhibited by Pneumococcal Capsule	102
SP-CHAP has Greater Activity than Cpl-1 Against Biofilm Pneumococci <i>in vitro</i>	106
SP-CHAP Treatment Does Not Cause Host Toxicity and Decolonizes the Nasopharynx <i>in vivo</i>	106
3.5 Conclusion.....	110
3.6 Materials and Methods:	111
Bacterial Strains and Growth Conditions	111
Cloning	111
Protein Expression and Purification	111
Turbidity Reduction Assay	112
Colony Forming Unit Assay.....	112
Minimum Inhibitory Concentration Assay (MIC).....	113
Biofilm Assay	113
Confocal Microscopy	114
Acute Systemic Toxicity Testing	115
Mouse Colonization Studies.....	115
Chapter 4: Discussion and Future Directions	117
4.1 Summary of the Dissertation.....	117
4.2 Discussion & Future Directions	120
Appendix A: Sequence of 76 novel endolysins identified containing C-terminal consensus sequence	133
Appendix B: Supplementary data published in manuscripts	143
Appendix C: Sequence Information for Constructs used in Dissertation-Related Studies.....	144
Appendix D: List of Published, Submitted and Planned Co-authored Manuscripts.....	164
Bibliography	170

List of Figures

Chapter 1

Figure 1-1. Virulence factors of <i>S. pneumoniae</i>	9
Figure 1-2. Modular organization of phage-encoded pneumococcal choline-binding proteins.....	12
Figure 1-3. Timeline of antibiotic discovery and resistance development.....	15
Figure 1-4. An overview of the different classes of antibiotics.....	18
Figure 1-5. Mechanisms of antibiotic resistance.....	24
Figure 1-6. Comparison of the bacteriophage life cycles.....	31
Figure 1-7. Overview of the bacterial resistance mechanisms to phages.....	33
Figure 1-8. Structure of the Gram-positive bacterial peptidoglycan.....	35
Figure 1-9. Electron micrographs of bacteria treated with endolysins.....	36
Figure 1-10. Comparison of DAP-type vs Lys-type peptidoglycan structure.....	39
Figure 1-11. Basic structure of the bacterial cell wall peptidoglycan.....	41

Chapter 2

Figure 2-1. Size exclusion chromatography and cellular activity of wild-type Cpl-1 and the N-terminal CBR W210A:F218A:Y238A triple mutant.....	69
Figure 2-2. Analytical ultracentrifugation of Cpl-1 and Cpl-1 Δ tail.....	72
Figure 2-3. Size exclusion chromatography of Cpl-1 tail alanine scanning mutagenesis.....	75
Figure 2-4. Comparison between the dimers of LytA crystal structure (PDB code: 4X36) and the model of Cpl-1 generated using the AI program AlphaFold2.....	77
Figure 2-5. Bactericidal activity of Cpl-1 F328A against <i>S. pneumoniae</i> DCC1811.....	78
Figure 2-6. Choline concentration dependent formation of Cpl-1 dimers.....	80

Figure 2-7. Novel endolysins identified by the consensus sequence form a dimer in the presence of choline.....	82
---	----

Chapter 3

Figure 3-1. Endolysin SP-CHAP is highly active against planktonic-grown pneumococci and its activity is modulated by capsule polysaccharide availability.....	104
--	-----

Figure 3-2. SP-CHAP reduces pneumococcal biofilm biomass <i>in vitro</i> and nasopharyngeal colonization <i>in vivo</i>	108
--	-----

List of Tables

Chapter 1

Table 1-1. Human and animal streptococcal diseases.....	4
---	---

Chapter 4

Table 4-1. Cpl-7 CBD host range.....	126
--------------------------------------	-----

Appendix B

Table B-1. Primers used in Chapter 2.....	143
---	-----

List of Abbreviations

ABC	ATP-binding cassette
CBD	cell binding domain
CBM	choline binding module
Cbp	choline binding protein
CBR	choline binding repeats
CD	catalytic domain
CDC	centers for disease control and prevention
CFSAN	center for food safety and applied nutrition
CFU	colony forming units
CHAP	cysteine, histidine-dependent amidohydrolase/peptidase
CSE	cystathionine γ -lyase
CWBD	cell wall binding domain
EAD	enzymatically active domain
EPS	extracellular polymeric substances
FDA	Food and Drug Administration
FM	functional module
GAS	Group A Streptococcus
GBS	Group B Streptococcus
GCS	Group C Streptococcus
GES	Group E Streptococcus
GlcNAc	N-acetylglucosamine
GPCR	G-protein coupled receptor
GRAS	generally recognized as safe
IPD	invasive pneumococcal disease
LB	luria broth
LPS	lipopolysaccharide
MIC	minimum inhibitory concentration
MRSA	methicillin-resistant <i>Staphylococcus aureus</i>
MurNAc	N-acetylmuramic acid
NAG	N-acetylglucosamine
NAM	N-acetylmuramic acid
OD	optical density
PAF	platelet-activating factor
PAP	periplasmic adaptor protein
PBP	penicillin-binding protein
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PG	peptidoglycan
PsrP	pneumococcal serine-rich repeat protein
PTD	protein transduction domain
RND	resistance-nodulation-division

SDM	site-directed mutagenesis
SEC	size exclusion chromatography
SP	signal peptide
SPR	surface plasmon resonance
STSS	streptococcal toxic shock syndrome
TA	teichoic acid
THY	Todd Hewitt yeast broth
TRA	turbidity reduction assay
WHO	world health organization
WT	wild-type
Xcc	<i>Xanthomonas citri</i> pv. <i>citri</i>
YLL	years of life lost

Chapter 1: Introduction and Literature Review

1.1 Streptococcal Infectious Diseases

Few will disagree that the infectious disease burden is a long overdue compelling, urgent priority for global public health.

Less than a year ago, the Global Burden of Disease 2019 group of collaborators spread over 204 countries, published their findings on 11 infectious syndromes by studying 343 million individual records or isolates covering 11,361 study-location-years (GBD Antimicrobial Resistance Collaborators, 2022). They found that of the 13.7 million infection related deaths in 2019, 54.9% were due to five pathogens, viz. *Staphylococcus aureus*, *Escherichia coli*, *Streptococcus pneumoniae*, *Klebsiella pneumoniae*, and *Pseudomonas aeruginosa*, out of the 33 pathogens investigated. Of these five, only *S. pneumoniae* has been a focus of global surveillance and public health initiatives. The study found it to be the leading pathogen in terms of years-of-life-lost (YLL) in 69 of the 204 countries. *S. pneumoniae* was associated with the most deaths (225,000) among young children (post-neonatal to age 4 years) across the entire study.

The Global Burden of Disease data reported two infectious syndromes as responsible for more than 2 million deaths each in 2019: lower respiratory infections with 4 million deaths, and bloodstream infections with 2.91 million.

Delving into the last pandemic of the “Spanish flu” from a hundred years earlier, i.e. before antibiotics were discovered, Dr. Anthony Fauci and his team studied slides recut from preserved autopsy lung tissue specimens of influenza fatalities and revisited bacteriology and pathology records (Morens et al., 2008). Their conclusion was that most deaths in this influenza pandemic

resulted from secondary bacterial pneumonia rather than the etiologic agent of the time, since the changes observed were consistent with classical descriptions of extensive bacterial pneumonia.

There is now wide acceptance for pneumonia occurring when secondary bacterial infection is associated with acute respiratory viral infection. *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Staphylococcus aureus* are the most common coinfecting bacteria. Much of the studies of interactions between organisms of the human microbiome, particularly for chronic conditions, has been focused on the gut. Early studies point to a linkage between the respiratory system and the gastrointestinal tract, known as the “gut-lung axis” (Lane et al., 2023), with immune dysregulation caused by the antiviral response being a contributor to secondary bacterial infection setting in.

The *Streptococcus* genus comprises a diverse collection of Gram-positive bacteria that are responsible for causing a wide variety of infections. In humans, streptococcal infections vary in severity from mild dental caries to life-threatening meningitis and necrotizing fasciitis. Streptococci are classified into different groups based on their hemolytic properties. The most common groups of *Streptococcus* that cause disease in humans are: *Streptococcus pyogenes* (Group A streptococcus), *Streptococcus agalactiae* (Group B Streptococcus), *Streptococcus pneumoniae*, and *Streptococcus mutans*.

Streptococcus pyogenes (Group A Streptococcus, GAS), is a human pathogen that infects over 600 million people a year and causes over 1.78 million severe infections that lead to over half a million deaths each year (Sanyahumbi et al., 2016). It colonizes epithelial surfaces of the throat and skin causing strep throat (streptococcal pharyngitis), skin infections such as impetigo, scarlet fever and cellulitis. Strep throat is the most common form of infection; it is highly contagious among children and adolescents but can infect people of all ages. While these mild infections are

typically not life-threatening, when left untreated, can lead to more serious complications such as rheumatic heart disease which is an inflammatory condition affecting the heart, joints, nervous system and skin. Some of the most serious life-threatening diseases caused by GAS include necrotizing fasciitis (flesh-eating disease) and streptococcal toxic shock syndrome (STSS).

Streptococcus agalactiae (Group B Streptococcus, GBS), is found in the lower human gastrointestinal tract and causes illness in infants, pregnant women, the elderly and the immunocompromised. GBS is the main cause of neonatal sepsis, meningitis and pneumonia worldwide. As per a global study conducted in 2017 regarding the impact of *S. agalactiae*, it was found that approximately 21.7 million pregnant women, accounting for one out of every five harbor *S. agalactiae* in their vagina asymptomatically. The pathogen is transmitted to the newborn during the process of childbirth (Seale et al., 2017). *S. agalactiae* infections have been estimated to cause 3.5 million preterm births, 409,000 adverse maternal/infant outcomes and 147,000 preventable stillbirths annually.

Streptococcus mutans colonizes the human oral cavity causing dental caries (i.e. cavities). Demineralization of tooth enamel resulting from localized acidification due to fermentation of carbohydrate rich environments by oral *S. mutants* is the cause of dental cavities. *S. mutans* attaches to the surface of teeth forming biofilms, also known as “plaque”. When it spreads into the bloodstream, *S. mutans* can also cause bacteremia and endocarditis. Global estimates suggest that 2.3 billion people suffer from dental caries of their permanent teeth, and 530 million children suffer from caries of primary teeth (GBD 2017 Disease and Injury Incidence and Prevalence Collaborators, 2018). Fluoridated drinking water, intended to prevent dental caries, has shown to be ineffective with >90% of adults between 20 and 64 years of age having had dental caries and 27% experiencing untreated tooth decay (Dye et al., 2015).

Streptococci are not only human pathogens, they also cause disease in animals. *Streptococcus uberis*, *Streptococcus suis*, Group C streptococci (GCS), Group E streptococci (GES) are responsible for causing disease in cattle, pigs, horses and sheep. Table 1-1 below shows the various streptococcal species with their associated host and disease they cause.

From a treatment approach, streptococcal infections largely manifest at exposed surfaces or accessible mucous membranes (i.e. skin, nasopharyngeal cavity, lungs, oral cavity, vaginal tract, etc.), which should be beneficial for protein therapeutics and affect efficacy, immune responses, pharmacokinetics/pharmacodynamics. This, coupled with the fact that most streptococcal disease is the result of infection by a single species as opposed to a polymicrobial infection, lends promise to a targeted approach to combat streptococcal infections.

Table 1-1. Human and animal streptococcal diseases.

Reproduced with permission from (Nelson et al. 2021).

Streptococcal species	Host	Disease
<i>S. pyogenes</i> (GAS)	Human	Pharyngitis, impetigo, necrotizing fasciitis
<i>S. agalactiae</i> (GBS)	Human	Neonatal meningitis
<i>S. agalactiae</i> (GBS)	Dairy cow	Bovine mastitis
<i>S. dysgalactiae</i> (GCS)	Dairy cow	Bovine mastitis
<i>S. equi</i> (GCS)	Equids (horse, donkey, pony)	Strangles disease
<i>S. mutans</i>	Human	Dental caries
<i>S. pneumoniae</i>	Human	Pneumonia, otitis media
<i>S. suis</i>	Human	Septicemia, meningitis, endocarditis
<i>S. suis</i>	Swine	Septicemia with sudden death, meningitis, endocarditis
<i>S. uberis</i>	Dairy cow	Bovine mastitis

Streptococcus pneumoniae

S. pneumoniae is a Gram-positive bacterium that is part of the normal flora that colonizes mucosal surfaces of the human respiratory tract. It is estimated that among children, about 25-60 % are asymptomatic carriers of *S. pneumoniae* in the nose, while about 10% of adults are asymptomatic carriers (Abdullahi et al., 2012; Nunes et al., 2005; Yahiaoui et al., 2016). *S. pneumoniae* can cause a wide range of disease severity. *S. pneumoniae* is the most common cause of community-acquired bacterial pneumonia worldwide (van der Poll & Opal, 2009); it also causes simple otitis media infections. It is the cause of other life-threatening diseases such as invasive bloodstream infection (bacteremia) and meningitis (brain infection) (Shenoy & Orihuela, 2016). Pneumonia affects people at the extremes of life. According to the World Health Organization (WHO) and the Centers for Disease Control and Prevention (CDC), *S. pneumoniae* infections were responsible for the death of almost 750,000 children under the age of 5 in 2019 alone, and caused a significant number of infections in adults aged over 50 (CDC, 2022; Durando et al., 2013; WHO, 2019, 2021). At least 1.2 million infant deaths are attributed to this pathogen every year worldwide (van der Poll & Opal, 2009), and in the US, about 10% of invasive pneumococcal disease cases are fatal (Robinson et al., 2001). A recent study comparing the global mortality associated with bacterial pathogens across the world reported the years-of-life-lost burden to be highest for *S. pneumoniae* compared to various other infectious pathogens while also being the leading bacterial cause of death in 24 countries (GBD Antimicrobial Resistance Collaborators, 2022). Geographically speaking, *S. pneumoniae* had the highest age-standardized mortality rate in Oceania, South Asia and western sub-Saharan Africa. Major problems of pneumococcal disease include the acquisition of antimicrobial resistance and the global spread of resistant clones. Because of the severity and wide-ranging infections caused by *S. pneumoniae*, the WHO has listed

S. pneumoniae in the global priority list of antibiotic-resistance bacteria for which new drugs need to be developed (Tacconelli et al., 2018).

In terms of morphology, pneumococci are lancet-shaped, capsule-forming diplococci. *S. pneumoniae* do not form spores and are 0.5-1.25 μm in size. Most *S. pneumoniae* are α -hemolytic but can cause β -hemolysis during anerobic incubation. Unlike staphylococci, all streptococci lack the enzyme catalase which is responsible for catalyzing the decomposition of hydrogen peroxide to water and oxygen and is important in protecting the cell from damage by reactive oxygen species.

Streptococcus pneumoniae Virulence Factors

S. pneumoniae's effectiveness in causing infections is directly related to the developmental stage of the host's immune system and possible deterioration with aging. *S. pneumoniae* produces various virulence factors that enable the bacterium to survive in different host environments. Despite pneumococcal pathogenesis having been studied for over 100 years, new insights on *S. pneumoniae* virulence factors and on mechanisms as well as host responses continue to be uncovered (Marquart, 2021). Advances in molecular, biochemical and bioinformatic methods are providing new insights into streptococcal pathogenesis.

The main virulence factor for *S. pneumoniae* is its polysaccharide capsule. To date, there are at least 100 capsular serotypes identified (Ganaie et al., 2020; Steel et al., 2013). The polysaccharide capsule is a fluid-like structure that mediates host cell attachment while providing protection from the host's immune system. The capsule inhibits phagocytosis by innate immune cells and protects the bacteria from recognition by complement factors. The various serotypes identified differ in the composition of the polysaccharides present in the capsule with each one

being pathogenic in their own way. For example, serotype 1 is associated with invasive infections which have lower fatalities, whereas serotype 3 is associated with colonization of the nasopharynx and more serious infections that can lead to fatalities (Brooks & Mias, 2018). The capsule's involvement in pathogenesis has been attributed to its electric charge. The capsule carries a predominantly negative net charge, primarily due to the presence of acidic polysaccharides and phosphates within its composition. Electrostatic repulsion has been proposed as a possible explanation for how *S. pneumoniae* is able to avoid being trapped by negatively charged mucus layers and phagocytic cells, such as macrophages (Li et al., 2013; Nelson et al., 2007). Notably, *S. pneumoniae* is able to undergo a phenomenon known as capsule switching in which mutations in the capsule polysaccharide synthesis genes promote serotype switching. The occurrence of serotype switching in strains is becoming more prevalent, frequently facilitated by recombination or polymorphisms driven by the selective pressures exerted by antibiotics and vaccines.

Some other virulence factors include pneumolysin, choline-binding surface proteins and phosphorylcholine. Pneumolysin is a hemolytic pore-forming toxin capable of lysing cells that have cholesterol in their membrane (Nishimoto et al., 2020). Its highly conserved sequence consisting of a tryptophan-rich loop and a threonine-leucine amino acid pair are responsible for recognition and binding to membrane-bound cholesterol (Farrand et al., 2010; Nollmann et al., 2004). Pneumolysin can also contribute to inflammation by activating the complement pathway after acting as a toll-like receptor ligand resulting in production of inflammatory mediators such as interleukin-8 as well as reducing phagocytosis by innate immune cells (Cockeran et al., 2002). Being pro-inflammatory, there is an increase in shedding and a higher rate of host-to-host transmission of the bacteria. Pneumolysin facilitates bacterial crossing of the blood-brain barrier contributing to pneumococcal meningitis by inducing loss of tight junctions (Ring et al., 1998;

Zysk et al., 2001). In addition, brain tissue modification through astrocyte reorganization has been observed to be caused by pneumolysin as well as loss of cilia and induction of apoptosis of brain cells (Braun et al., 2002; Hirst et al., 2000). Pneumolysin has also been shown to promote the formation of bacterial biofilms (Shak et al., 2013). Studies have also shown that pneumolysin can cause DNA damage by inducing double-stranded DNA breaks. Due to its pore-forming properties, pneumolysin can create ion channels that alter cell calcium levels, leading to overproduction of intracellular reactive oxygen species that then cause DNA damage (Rai et al., 2016). Pneumolysin which has cytotoxicity toward almost all human cell types and induces inflammation in the lungs, was shown to be upregulated by *S. pneumoniae* during coinfection with RSV (Nishimoto et al., 2020). Phosphorylcholine can induce platelet aggregation by binding to the platelet-activating factor (PAF) receptor. Choline-binding proteins (Cbp) are unique to *S. pneumoniae* and perform a variety of different tasks described below. LytA, the major pneumococcal autolysin digests the cell wall and promotes the release of pneumolysin while PiaA, pneumococcal iron acquisition A is involved in acquisition of iron needed for growth of the bacteria (van der Poll & Opal, 2009). Pneumococcal surface antigen A (PsaA), plays a role in metal ion uptake leading to protection from oxidative stress (Kadioglu et al., 2008).

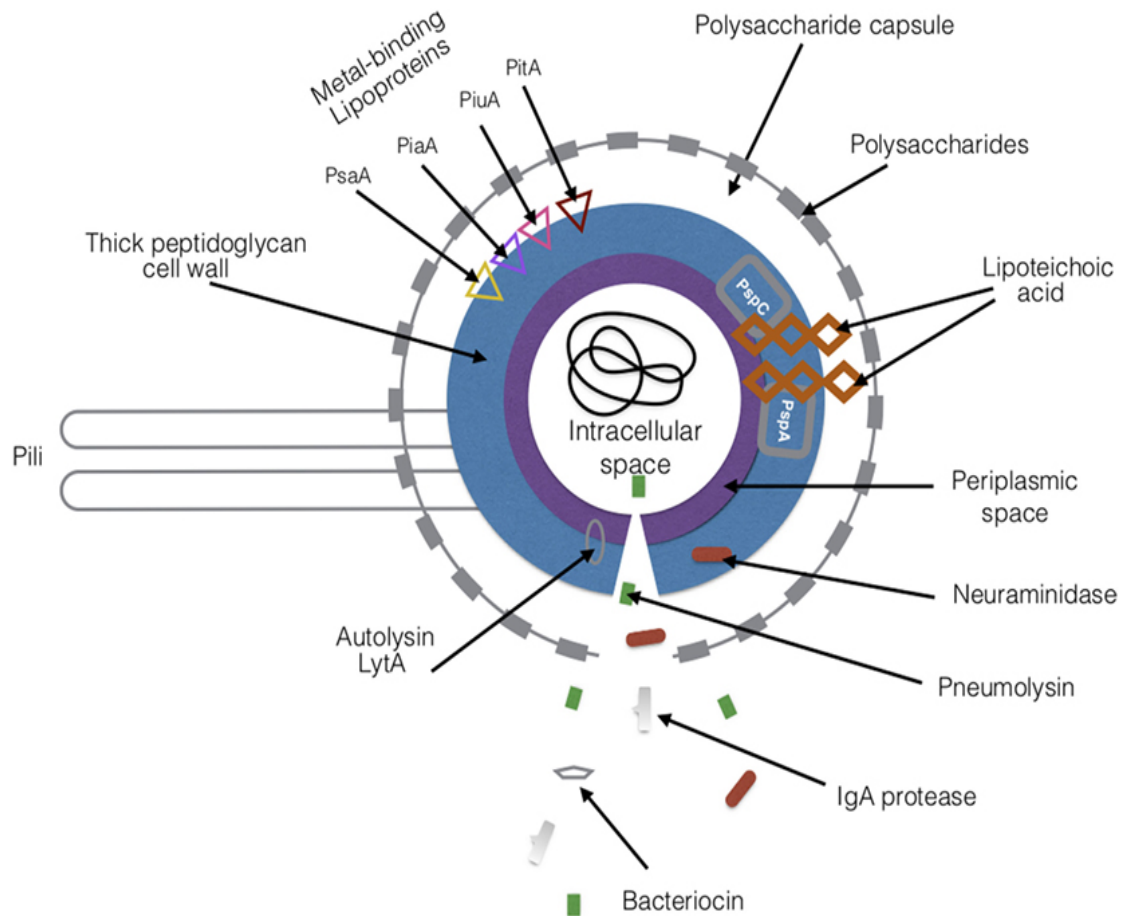


Figure 1-1. Virulence factors of *S. pneumoniae*. The various virulence factors of *S. pneumoniae* are shown. Image reproduced with permission from (Brooks & Mias, 2018).

Patients of sickle cell disease, one of the most common autosomal recessive disorders, have impaired immunity and are very susceptible to acute chest syndrome attended with *S. pneumoniae* infection, as its virulence factors devastate their lungs (Gonzales et al., 2021). Their primary defense against *S. pneumoniae* is the generation of reactive oxygen species during the oxidative burst of neutrophils and macrophages. Paradoxically, pneumococci themselves produce high levels of hydrogen peroxide, a reactive oxygen agent, as their virulence strategy along with pneumolysin which mediates particularly severe lung injury in sickle cell disease.

There are currently no known FDA approved anti-virulence therapeutics against *S. pneumoniae*, attractive as such an approach might seem. Extensive research has been conducted to design a vaccine targeting pneumolysin, without much avail. An anti-virulence approach would erode the ability of *S. pneumoniae* to colonize the host, thereby reducing the burden on patient immune systems.

The virulence of *S. pneumoniae* is a result of its ability to process carbohydrates presented on host cell surfaces. During the process of N-glycan trimming, various enzymes and transporters are deployed, and high throughput screening has identified virulence factors among these. A α -(1,2)-mannosidase, known as SpGH92, is involved in trimming high-mannose structures and has emerged as an enzyme of particular interest. Deletion of the gene for SpGH2 is known to attenuate the virulence of *S. pneumoniae*; as yet, inhibitors of SpGH2 have not been identified. Accordingly, recent investigation is being pursued on inhibiting the α -(1,2)-mannosidase SpGH92 as a potential anti-virulence therapeutic (Bhosale et al., 2023).

Choline-Binding Proteins

Choline-binding proteins (CBPs) are a family of polypeptides present in *S. pneumoniae* and related species, as well as in some of their bacteriophages. CBPs include peptidoglycan hydrolases, adhesins and bacteriophage encoded lytic enzymes. CBPs have a modular structural organization comprising a functional module (FM) and a choline-binding module (CBM) which serves to bind the protein to choline residues present in the bacterial cell wall. The FM is typically located at the N-terminus while the CBM is usually at the C-terminal end (except in the case of LytB and LytC). CBMs are comprised of multiple repeats, each about 20 amino acids long that are called choline-binding repeats (CBRs). Pneumococcal CBPs typically serve different roles to aid in bacterial cell viability and virulence (Maestro & Sanz, 2016). *S. pneumoniae* exhibits a unique nutritional requirement for choline and the pneumococcal cell wall is decorated with phosphocholine residues present in the wall teichoic acids. The number of CBRs can vary greatly from three (as in CbpG, CbpM and CbpK) to seventeen (as in LytB), although the number of CBRs does not necessarily correlate with the proteins' affinity for choline. All pneumococcal CBPs have been shown to form dimers in the presence of choline with the exception of Pce (phosphorylcholine esterase). Sedimentation velocity experiments have shown that Pce is a monomer in its free and choline bound state (Buey et al., 2007). Pce is responsible for the redistribution of choline in the cell wall by catalyzing the hydrolysis of phosphocholine residues in the teichoic acids in the cell wall. The modular organization of several pneumococcal CBPs is shown in Figure 1-2.

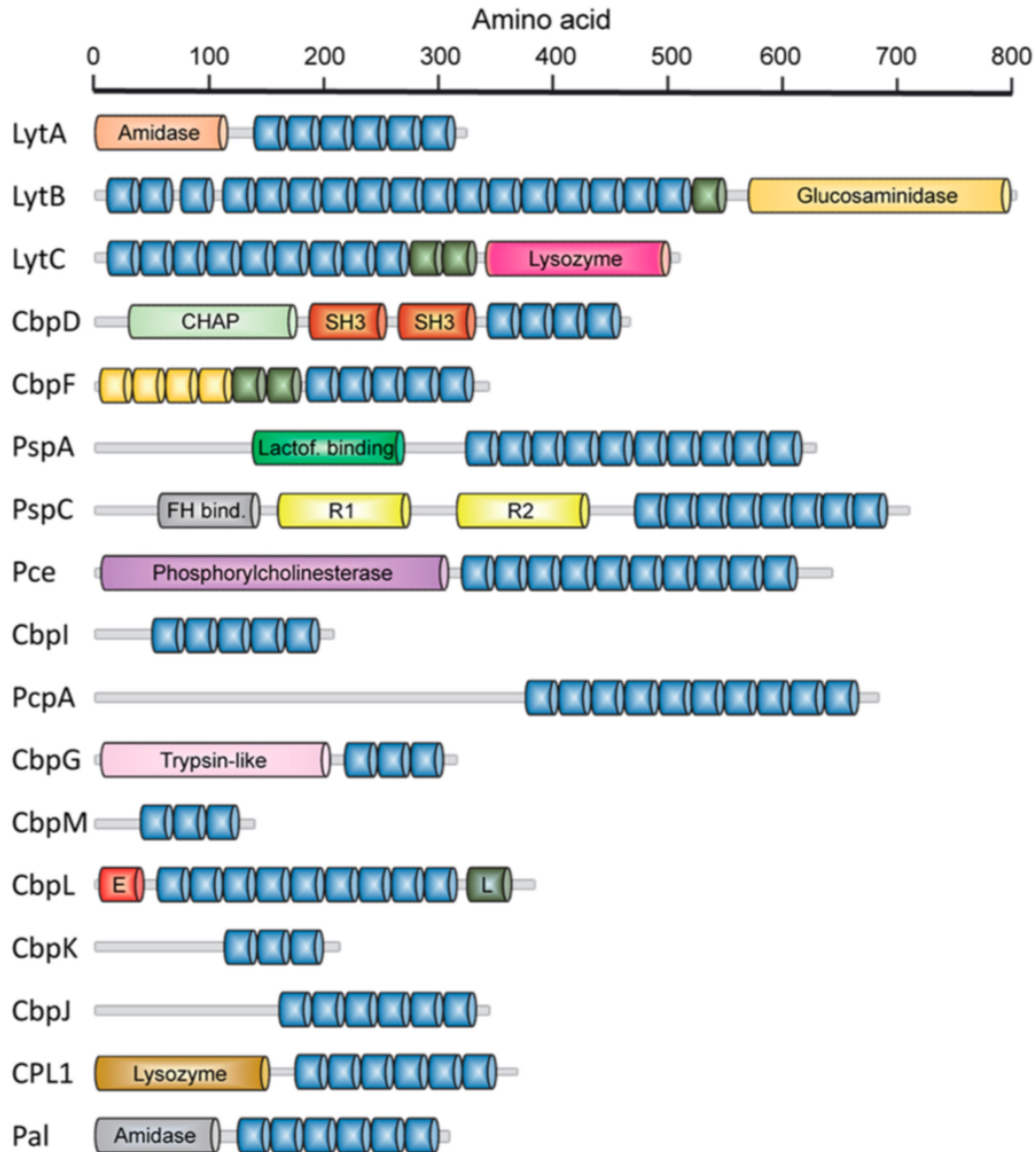


Figure 1-2. Modular organization of phage-encoded pneumococcal choline-binding proteins. Blue units represent choline-binding repeats (CBRs) that make up the choline-binding modules (CBM). Non-consensus CBRs are shown in green. Activities ascribed to various functional modules are shown. Image reproduced with permission from (Maestro & Sanz, 2016).

1.2 Antibiotics and Resistance

Before antibiotics, a relatively minor bacterial infection would prove to be difficult to treat or even deadly. Since their discovery, antibiotics have saved millions of lives. The discovery of antibiotics began in the 1920s while the golden age was from the mid-1940s to early 1960s as one-half of the drugs being used currently were discovered during this time period.

Antibiotic Discovery and Resistance Development

Antibiotics were discovered through serendipity and a keen eye. In 1928, Dr. Alexander Fleming had been cultivating *Staphylococcus* bacteria in a petri dish when he left for vacation. Upon returning from vacation, Dr. Fleming noticed mold growing in this petri dish. On closer observation, he saw that this mold seemed to be preventing the bacteria around it from growing. Investigation identified a substance that killed bacteria was being produced by the mold. This compound was named penicillin. After encountering several hurdles since initial discovery, including a very hard time to produce enough quantity of purified penicillin, it was only in 1940 that chemists Howard Florey and Ernst Chain from Oxford University were able to produce enough and animal trials began. Of eight mice that were infected with streptococci, only the four that were given penicillin survived while the remaining four succumbed to the infection.

Soon after penicillin was introduced into medicine, in 1943, the first aminoglycoside antibiotic, streptomycin, was discovered by Albert Schatz and Selman Waksman. Streptomycin gained popularity because it was active against *Mycobacterium tuberculosis* and *Yersinia pestis* (plague) infections.

The use of penicillin and streptomycin was greatly accelerated during World War II. Even though penicillin was discovered in the United Kingdom, the United States played a major role in

large scale production of the drug due to the World War. High demand and increasing bacterial infections during this time propelled the golden era of antibiotic discovery. Several antibiotic classes/groups were discovered between the 1940s to 1960s. In contrast, only three novel antibiotics (Daptomycin, Ceftaroline, Linezolid) have been introduced to the market in the last 30 years.

Antibiotic resistance has walked hand-in-hand with antibiotic discovery, as outlined in Figure 1-3. The first case of clinical resistance to penicillin was observed in 1945. Similar resistance was observed for other classes of antibiotics, while for tetracycline resistance was seen even before the drug went on to the market. Resistance to glycopeptides (vancomycin), rifamycins (rifampicin) and quinolones (ciprofloxacin) was seen soon after market introduction. Linezolid, approved by the FDA in 2000, encountered resistance early in clinical usage. It has been extensively studied through whole genome sequencing, and the mutations responsible for resistance and fitness compensation have been identified (Billal et al., 2011). Daptomycin was approved in 2003 despite resistance being observed only a year after its discovery in 1986. The alarming rate of resistance development soon after discovery and the emergence of multi-drug resistant bacterial strains compels us to the post-antibiotic era.

More recently, antibiotic resistance has been impacted by the COVID-19 pandemic (Feldman & Anderson, 2021; Sao-Jose et al., 2022). The prophylactic use of antibiotics, which reportedly surged at the time, including the use of antimicrobials as “repurposed drugs” to treat COVID-19 itself even without co-infection has become a concern. A US study documented early empirical antibiotic therapy was used in 56.6% of patients hospitalized with COVID-19, while a community-onset bacterial coinfection was confirmed in only 3.5% of patients (Vaughn et al., 2021).

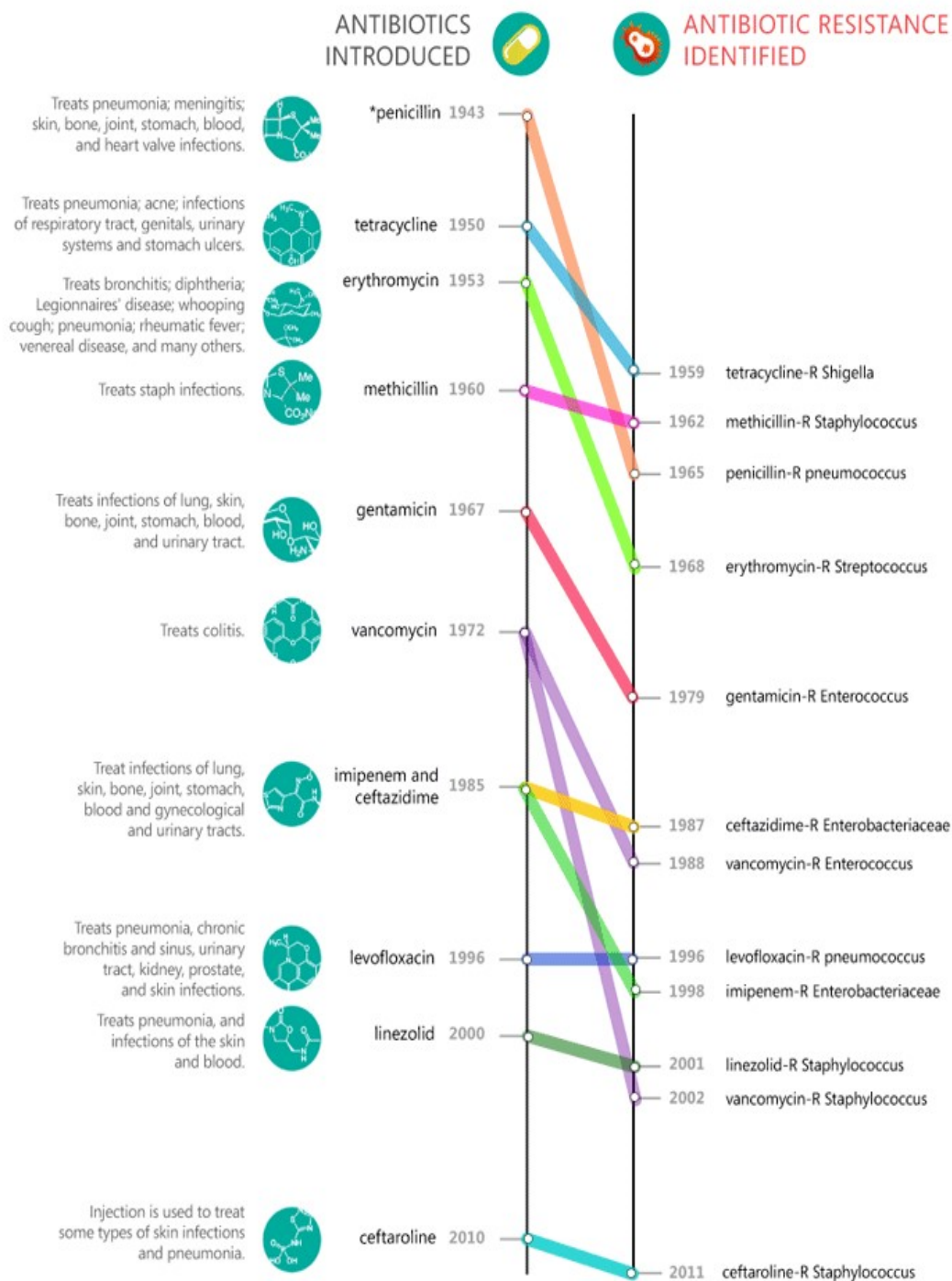


Figure 1-3. Timeline of antibiotic discovery and resistance development. Drug-resistant infections appear nearly as quickly as life-saving antibiotics are created. Image reproduced with permission from (Graham, 2017).

Antibiotic Classes and Mechanisms of Action

Antibiotics are classified into five categories based on their mechanism of action. In general, most antibiotics function by targeting processes such as cell wall synthesis, nucleic acid synthesis or protein synthesis (Kapoor et al., 2017; Kohanski et al., 2010). Antibiotics can be either bacteriostatic (prevent growth and reproduction) or bactericidal (result in bacterial cell death).

The largest class of antibiotics is the β lactams, which dates back to penicillin itself. These function by hindering bacterial peptidoglycan synthesis by preventing cell wall cross-linking directly or by interacting with various precursors. Examples include penicillins, cephalosporins, carbapenems, and monobactams. The antibacterial activity of penicillin, for example, is related to its β lactam moiety binding to penicillin binding proteins, which inhibits the normal function of these proteins by preventing cross-linking in the bacterial cell wall resulting in cell lysis and death. Vancomycin also inhibits cell wall synthesis by preventing transpeptidation after binding a D-alanyl subunit in the cell wall.

Inhibition of protein translation and synthesis is the mechanism by which the second largest group of antibiotics functions. Aminoglycosides (e.g., streptomycin, neomycin, kanamycin) and tetracyclines belong to this group, and function by binding with high affinity to the 16S rRNA of the 30S subunit ribosome resulting in premature termination of mRNA translation thereby preventing protein synthesis. Chloramphenicol and macrolides (e.g., erythromycin, azithromycin, clarithromycin), oxazolidinones (e.g., linezolid, cycloserine) and streptogramins suppress protein synthesis by binding the 23S rRNA of the 50S subunit ribosome, preventing t-RNA from binding.

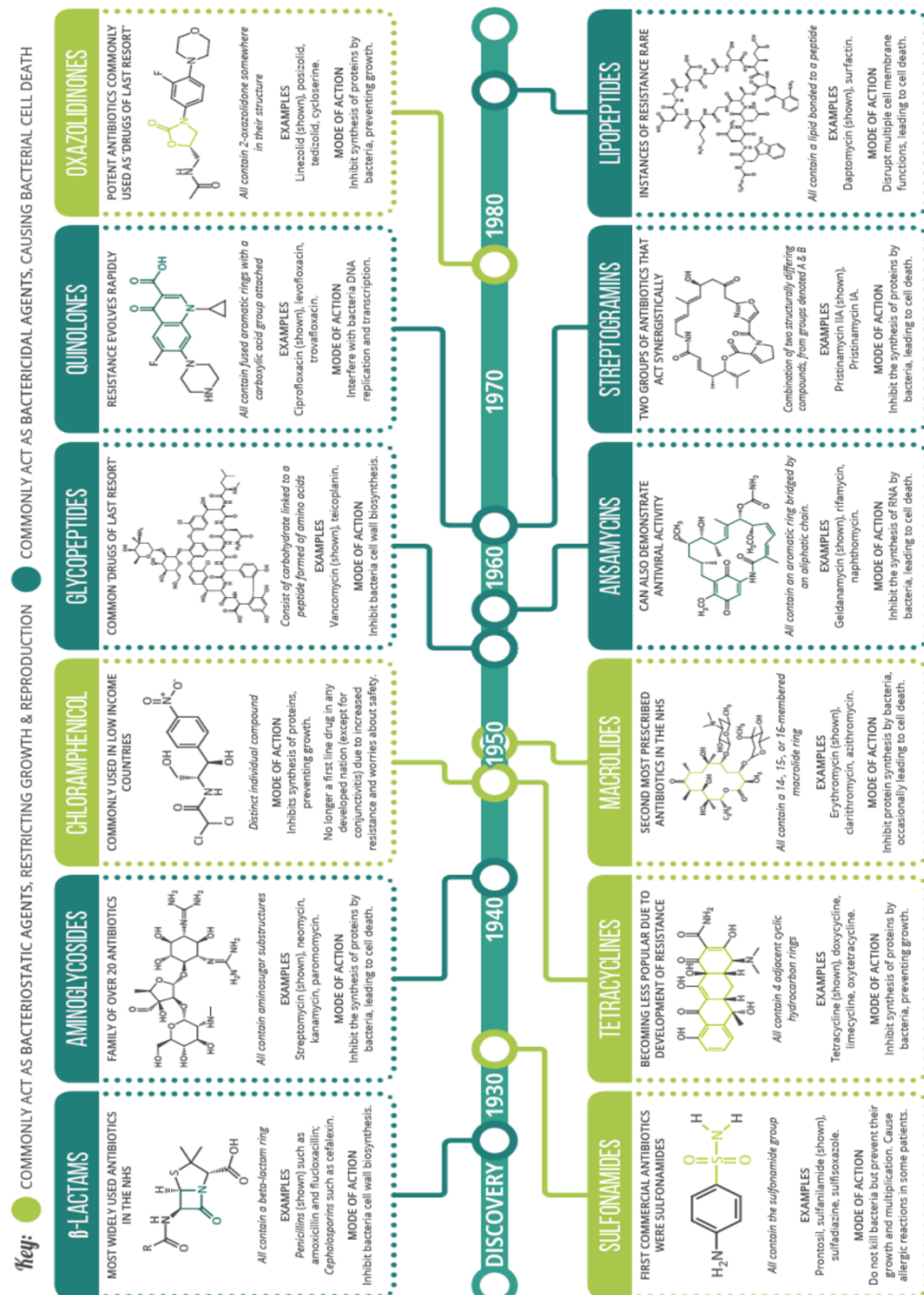
The third group of antibiotics inhibits DNA or RNA synthesis. Fluoroquinolones (e.g., ciprofloxacin, levofloxacin, moxifloxacin) target DNA gyrase and topoisomerase IV, inhibiting

supercoiling of DNA resulting in inhibition of DNA synthesis. Rifamycins (rifampicins) block RNA synthesis by binding and inhibiting DNA-dependent RNA polymerase.

The fourth group of antibiotics functions by altering cell membrane permeability and preventing bacterial cell growth. Colistin, polymyxin, daptomycin, and ionophores belong to this class.

The fifth and last class of antibacterials functions by interfering with bacterial metabolism and are bacteriostatic in nature. Sulfonamides (e.g. prontosil, sulfamethoxazole, trimethoprim) interfere with the folic acid metabolism cycle by competing with p-aminobenzoic acid in the synthesis of folic acid. Trimethoprim, a diaminopyrimidine, acts synergistically with sulfamethoxazole by inhibiting the next step of converting dihydrofolate to tetrahydrofolic acid, the physiologically active form of folic acid and a necessary cofactor in the synthesis of bacterial DNA. Among mycobacterial agents, isoniazid interferes with mycolic acid synthesis thus interfering with cell wall synthesis, while diarylquinolines, e.g. bedaquiline, function as ATP synthase inhibitors.

DIFFERENT CLASSES OF ANTIBIOTICS - AN OVERVIEW



Mechanisms of Antibiotic Resistance

Although there are multiple classes of antibiotics that function through different mechanisms of action, bacteria have evolved ways to develop resistance. Bacterial resistance can be either natural or acquired. Some bacteria are naturally resistant to certain antibiotics due to various structural or functional characteristics. On the other hand, other bacteria can acquire resistance due to selective pressure imposed by certain antibiotics that can mediate survival. This resistance can be acquired by horizontal gene transfer, mutations or vertical gene transfer. Mutations include single nucleotide polymorphisms (SNPs), frameshifts, insertions, and deletions. Horizontal gene transfer can occur by transformation (uptake of naked DNA), transduction (uptake of bacteriophage DNA during infection), and conjugation (DNA transfer through physical bacteria contact). Vertical gene transfer occurs by transfer of mutations from parental generation to offspring. For example, genetic variation has been observed in the dispensable genes for identical *S. pneumoniae* clones. Dispensable genes are not needed for growth but rather function to provide selective advantages such as antibiotic resistance (Weiser et al., 2018).

The first main mechanism that contributes to antibiotic resistance is reduced cell permeability. In order for many antibiotics to reach their target, they need to cross the bacterial cell envelope. While Gram-negative bacteria possess an outer membrane making them naturally more resistant to antibiotics, Gram-positive bacteria lack an outer membrane and are thus more permeable to the actions of antibiotics. Even so, alterations to the phospholipid and fatty acid content of the cytoplasmic membrane and loss of porins can reduce the ability of an antibiotic to penetrate the cell and result in the development of antibiotic resistance. Alterations in the structure of porins has been linked to the development of antibiotic resistance as seen in the case of *Klebsiella pneumoniae* where structural modification of porins OmpK35 and OmpK36 through

amino acid insertions led to a constriction of pore and increased resistance to carbapenems (Wong et al., 2019). Changes in electric charge within porins as a result of mutations in OmpC in the case of *E. coli* has been correlated with increased resistance to cefotaxime and gentamicin (Lou et al., 2011).

Efflux pumps are another common cause of antibiotic resistance. Efflux pumps are transmembrane proteins that transport toxic compounds and antibiotics from the cellular interior to the external environment in an energy dependent manner. Bacteria uses efflux pumps to pump antibiotics out of the cell. While all bacteria contain multiple efflux pumps, they are particularly important mediators of antibiotic resistance in Gram-negative bacteria. They function synergistically along with the outer membrane to make Gram-negatives intrinsically resistant to antibiotics. Bacterial efflux pumps have been classified into six families, with the major clinically relevant efflux pumps in Gram-negative bacteria belonging to the resistance-nodulation-division (RND) superfamily (Soto, 2013). They are typically composed of a cytoplasmic membrane pump, a periplasmic adapter protein (PAP), and an outer membrane protein channel. RND efflux pumps contribute to the multi-drug resistance observed for clinically relevant strains of Gram-negative bacteria such as *E. coli*, *P. aeruginosa*, *Neisseria gonorrhoeae*, and *A. baumannii*. The production of efflux pumps is subject to precise regulation, with the majority of these systems being governed by a repressor found in close proximity to the genes responsible for the structural components of the efflux system. Mutations in these repressor genes, which control transcriptional regulation, are often seen to result in the overexpression of efflux pumps. In the case of Gram-positive bacteria, ATP-binding cassette (ABC) transporters are the most common efflux pumps found and they use energy derived from ATP hydrolysis. One notable ABC pump is found in *Vibrio cholerae* (VcaM), and is capable of transporting fluoroquinolones and tetracycline.

Target modification and alteration is another mechanism whereby bacteria develop resistance to antibiotics. Structural alteration of the target can inhibit antibiotic binding conferring resistance. This is seen in the case of the quinolone antibiotics that function by inhibiting essential topoisomerase enzymes by binding near the active site. Amino acid substitutions in the target proteins result in lower antibiotic binding efficiency enabling the enzyme to function and thus confer antibiotic resistance (Bush et al., 2020). Resistance to β -lactam drugs occurs by preventing the interaction between the target penicillin-binding protein (PBP) and the drug, usually by modifying the ability of the drug to bind to the PBP (mediated by alterations to existing PBPs or acquisition of other PBPs). Target modification also plays a major role in the development of macrolide resistance. Ribosomal methyltransferases can methylate the 16S rRNA target preventing the binding of macrolides and streptogramins, thus enabling protein translation to continue. Methylation of the 16S rRNA is also a recently recognized method of developing high-level resistance to aminoglycoside antibiotics (Doi et al., 2016). Resistance to colistin has been observed to result from target modification as well. Colistin, an antibiotic active against Gram-negative bacteria functions by targeting lipopolysaccharide (LPS) resulting in membrane damage and cell death. Bacteria can develop resistance to colistin by decorating its LPS with moieties that alter the overall charge thereby preventing interaction between the antibiotic and its target.

Inactivation of antibiotics is a major mechanism of resistance seen in many bacteria wherein either the structure of the antibiotic is affected or it is subject to degradation which renders it ineffective. Hydrolysis of β -lactam antibiotics by β -lactamases is a commonly observed mechanism of resistance. β -lactamases inactivate β -lactam drugs by hydrolyzing the amide bond in the β -lactam ring structure causing the ring to open. These open-ring drugs are not able to bind their target PBP proteins and are ineffective. Currently there are over 7,000 β -lactamases identified

and the list continues to increase (Naas et al., 2017). Oxidation of tetracyclines by tetracycline-inactivating enzymes is another example of antibiotic resistance caused by drug inactivation.

The transfer of a chemical group facilitated by drug modifying enzymes can also cause antibiotic resistance. Aminoglycosides can be modified by aminoglycoside modifying enzymes such as acetyltransferases, phosphotransferases, and nucleotidyltransferases that modify the hydroxyl or amino groups of the antibiotic thereby reducing the affinity for its target (Ramirez & Tolmasky, 2010). Lincosamide antibiotics are susceptible to modification by nucleotidyltransferases, enzymes that append phosphate containing groups to the antibiotic reducing binding affinity. Phosphotransferases and esterases have the ability to alter macrolides, providing resistance, as these modified macrolides exhibit reduced efficiency in binding to the 50S ribosome.

Target bypass is another method employed by bacteria to develop resistance to antibiotics. In this method, the bacteria make the original target of the antibiotic redundant by producing an alternative pathway to bypass the antibiotic. Perhaps the most renowned instance of target bypass is the emergence of methicillin-resistant *S. aureus* (MRSA). β -lactam antibiotics like methicillin attach to PBPs and obstruct the transpeptidase domain, resulting in the disruption of cell wall synthesis. In the case of *S. aureus*, it has the capability to acquire an external PBP (PBP2a) that bears resemblance to the original target but exhibits diminished affinity for β -lactam antibiotics, thereby preventing the inhibition of cell wall synthesis, since transpeptidase activity of PBP2a is unaffected (Stapleton & Taylor, 2002). On the other hand, vancomycin binds to the terminal D-alanine-D-alanine from the pentapeptide precursors and inhibits peptidoglycan crosslinking, which inhibits cell wall synthesis. Enterococci can develop vancomycin-resistance by using a target bypass strategy. Enterococci can promote the synthesis of modified peptidoglycan precursors

leading to vancomycin now binding to D-alanine-D-lactate or D-alanine-D-serine with reduced affinity (Miller et al., 2014).

A relatively new identified resistance mechanism involves the bacterial hydrogen sulfide (H₂S) defense system which protects bacteria from oxidative stress (Shatalin et al., 2021). Cystathionine γ -lyase (CSE) was recently identified as the primary generator of H₂S in *Staphylococcus aureus* and *Pseudomonas aeruginosa*. Researchers have been able to block the source of bacterial H₂S by using small molecules that inhibit CSE and enable bacterial sensitivity to resistant antibiotics. CSE could thus be an attractive drug target to slow the spread of antibiotic resistance.

Finally, there has also been an interest in compounds that are called antibiotic resistance breakers. These compounds possess no antibacterial effect on their own but rather function to restore the efficacy of an antibiotic by inhibiting a specific mechanism of resistance. For example, clavulanic acid, sulbactam, and tazobactam are β -lactam compounds that can modify and inhibit β -lactamase enzymes thus creating an inactive inhibitor-enzyme complex that renders β -lactam antibiotics effective. While these are attractive options to inhibit the resistance mechanism, unfortunately, resistance has been observed for β -lactamase inhibitors as well.

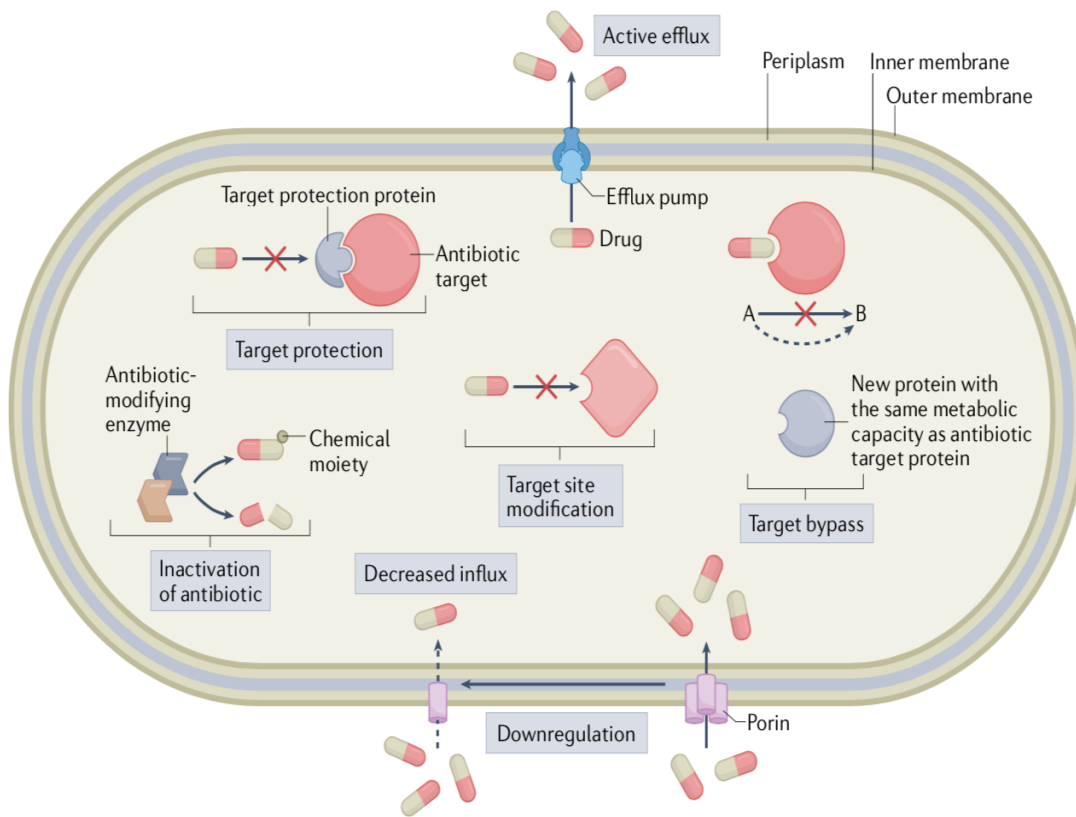


Figure 1-5. Mechanisms of antibiotic resistance. Antibiotic inactivation is mediated by reduced cell permeability, efflux pumps, target modification, inactivation, transfer of a chemical moiety and target bypass. Image reproduced with permission from (Darby et al., 2023).

Antibiotic Resistant *S. pneumoniae*

There has been a worldwide increase in antibiotic-resistant strains of *S. pneumoniae* (Butler et al., 1996; Song et al., 2013). Various studies report high rates of resistance for several antimicrobials such as penicillin (34.2%), trimethoprim-sulfamethoxazole (31.9%), and erythromycin (29.5%) (Doern et al., 2005). The European Antimicrobial Resistance Surveillance network (EARS-Net) reports infections with antibiotic resistance were 2.6 times higher than reported by the U.S. Centers for Disease Control (Cassini et al., 2019). Resistant infections imposed a burden of almost 900,000 disability-adjusted life years in the EU during the year 2015 alone, underscoring the rising urgent need for alternatives to antibiotics that are endowed with a much lower likelihood of developing resistance.

Penicillin binding proteins are the main reasons for β -lactam resistance in *S. pneumoniae*, with three of these (PBP1a, 2x, and 2b) being majorly involved (Li et al., 2023). Studies suggest that high-level penicillin resistance is a result of simultaneous alterations in penicillin-binding proteins 1a, 2x, and 2b. MurM, an enzyme involved in the biosynthesis of branched-stem cell wall muropeptides, has been reported to be required for penicillin and cefotaxime resistance. Mechanisms independent of penicillin binding proteins also contribute to β -lactam resistance; for example, a two-component signal-transducing system composed of histidine kinase CiaH and response regulator CiaR was found to be involved in the resistance phenotype of cefotaxime (Li et al., 2023).

The mechanism of *S. pneumoniae* resistance to macrolides involves erythromycin ribosomal methylase encoded by the *erm*(B) gene and macrolide efflux MFS transporter encoded by the *mef* (*mefE* or *mefA*) gene. A majority of erythromycin-resistant pneumococcal isolates

have been found to carry the *erm*(B) gene, with a minority carrying the *mefA* gene, and some strains having both *erm*(B) and *mefA* genes (Li et al., 2023).

Fluoroquinolones target type II DNA gyrase, encoded by *gyrA* and *gyrB* genes, and type IV topoisomerase, encoded by *parC* and *parE* genes, which are involved in DNA replication. Pneumococcal resistance to fluoroquinolones is explained by mechanisms of target mutation in the *gyrA/B* and/or *parC/E* genes, and the efflux mechanism reducing intracellular antibiotic accumulation.

Tetracycline resistance has been reported since the late 1950s. Ribosomal protection mediated by the genes *tet*(M) and *tet*(O) is the main mechanism for *S. pneumoniae* resistance to tetracyclines.

A majority of antibiotic prescriptions in the United States (up to 75%) are for upper respiratory tract infections in ambulatory patients, and many of these prescriptions are thought to be unnecessary because of the viral etiology of colds and acute bronchitis (Nuermberger & Bishai, 2004). As a consequence of antibiotic drug intervention reduction programs instituted in Finland and Iceland, prescriptions fell not only for macrolides but for other classes of antibiotics too. This was accompanied by a 48% reduction in the macrolide resistance rate, underscoring the benefit of judicious antibiotic prescribing (Nuermberger & Bishai, 2004). However, it is not clear if such a benefit could be obtained in decentralized medical systems such as in the U.S. Similarly, in a study conducted in Poland, isolated *S. pneumoniae* strains responsible for acute otitis media were found to be sensitive to antibiotics whose use is limited, as in the case of ofloxacin which is used only rarely in specific groups of children (Zielnik-Jurkiewicz & Bielicka, 2015). Suboptimal exposure to an antibiotic (i.e. inadequate dosing regimens) can selectively reduce the population of susceptible bacteria and favor the selection and outgrowth of drug-resistant clones.

S. pneumoniae Vaccines

Since their introduction into the market, pneumococcal vaccines have had a significant impact on reducing the incidence of pneumococcal disease. While vaccination initiatives have proven effective in reducing the occurrence of the most common serotypes, the emergence and dissemination of drug-resistant strains that were not initially covered by the vaccine formulations are currently threatening our ability to combat these infections. There are two main kinds of *S. pneumoniae* vaccines available in the United States, namely, the pneumococcal conjugate vaccines (e.g. PCV13, PCV15, PCV20) and the pneumococcal polysaccharide vaccine (PPV23; Pneumovax 23). PPV23 consists of polysaccharide from 23 different serotypes (1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19F, 19A, 20, 22F, 23F, and 33F) all of which represent the most common serotypes responsible for 85-90% of pneumococcal infections in the world. PPV23 protects against 23 serotypes and has shown efficacy against invasive pneumococcal disease (IPD) but poor effectiveness against pneumococcal pneumonia prevention. The pneumococcal conjugate vaccine, PCV13 (Prevnar 13®) uses polysaccharide antigens conjugated with carrier proteins (non-toxic diphtheria CRM₁₉₇ protein) that can elicit a T cell response, resulting in superior and longer lasting immunogenicity. PCV13 contains 13 valent capsular polysaccharides from serotypes (1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, 23F). In contrast to PPV23, PCV13 protects against only 13 serotypes but offers protection from pneumonia and IPD (Kim et al., 2017). The CDC recommends PCV13 or PCV15 for babies and children below 5 years of age and children between 5 to 18 years of age. PCV15 or PCV20 is recommended for adults aged between 19 to 65 years and those 65 years and older. PPV23 is recommended for adults 19 years or older who have previously received PCV15. The latest pneumococcal vaccine to be approved by the FDA is Pfizer's Prevnar 20®, which protects against 20 different pneumococcal serotypes (1, 3,

4, 5, 6A, 6B, 7F, 8, 9V, 10A, 11A, 12F, 14, 15B, 18C, 19A, 19F, 22F, 23F, and 33F) and received FDA approval recently in April 2023.

While the abovementioned vaccines exist and are effective, they offer limited protection against *S. pneumoniae* infection. This is mainly due to the fact that over 100 different *S. pneumoniae* serotypes exist and the current vaccines on the market protect only against certain serotypes (Ganaie et al., 2020). In addition, serotype switching is a healthcare concern as non-vaccine serotypes are being detected at higher rates compared to before vaccines were implemented (Chang et al., 2015). In developing countries, vaccines do not always protect from pneumococcal disease due to insufficient access to vaccination programs or because serotypes that are not covered by vaccines are dominant. In fact, vaccine candidates employing conserved *S. pneumoniae* surface protein immunogens remain an important area of research with the goal of serotype-independent coverage (Aceil & Avci, 2022). Towards this end, the Pneumococcal Serine-Rich Repeat Protein (PsrP), present in about 50% of clinical lung isolates, is a potential candidate which has not yet been worked on for direct vaccine studies although antibodies and passive immunization do provide protection against lung infection. It is believed that investigation of its extensive glycosylation may help yield a vaccine candidate with better clinical application.

The track of progress in development of vaccines, which has been undoubtedly helpful in reducing the prevalence of pneumococcal infections from a public health perspective, nevertheless points to the need for supplementing worldwide efforts with alternative approaches to combat the rise of *S. pneumoniae* infections.

1.3 Endolysins – Bacteriophage Derived Peptidoglycan Hydrolases

Bacteriophages

Bacteriophages are viruses that infect bacteria. They were discovered in 1915 by Frederick Twort and Felix d'Herelle in 1917 and are the most abundant microbial organisms on the earth (Summers, 2011). Soon after their discovery, bacteriophages were harnessed for their ability to treat bacterial diseases due to their bactericidal capacity, with the first reported case of a phage being used to treat a staphylococcus skin infection in 1921, with the infection clearing within two days of injection of the bacteriophage (Hermoso et al., 2007). While phages gained the attention of pharmaceutical companies in the 1930s, the discovery and manufacture of antibiotics in the 1940s slowed down the use of phages to treat bacterial infectious diseases and pushed them into the background. In Eastern Europe and the former Soviet Union, however, the use of phages for the treatment of bacterial infections has been pursued with success in the last century. Fast forward 60 years, and the rise of antibiotic resistant bacterial strains has brought attention back to phages as potential therapies for treatment of bacterial diseases. In 1968, human phage therapy studies were performed in Georgia, Russia, and Poland, including studies on *Shigella* phage against bacterial dysentery. In August 2006, the US FDA approved the use of a phage therapy cocktail that included six purified phages targeting *Listeria monocytogenes* for application on ready-to-eat meat and poultry products for decontamination purposes. The use of bacteriophages for the prevention and treatment of bacterial infections is now known as phage therapy.

In order to replicate, bacteriophages undergo a couple of steps. To begin, bacteriophages recognize and bind to various receptors present in the bacterial cell wall via their tail fibers. The tail fibers then flex and baseplate components penetrate the bacterial cell wall, a step that uses ATP

as its energy source. The phage then injects its viral DNA into the bacteria once the tail contracts and the tail tube passes through an opening in the baseplate that pierces the cell wall. Once inside the cell, the bacteriophage takes over the cell's replication cycle and protein synthesis enzymes to synthesize viral components. Once created and assembled, mature virions are released.

Bacteriophages are classified based upon their replication cycle. Lysogenic (temperate) phages enter their host, integrate their DNA with that of the host and replicate together (i.e. prophage), or exist as plasmids within their host, whereas lytic (virulent) phages infect their host and kill the host immediately after virion replication has completed. The lysogenic life cycle can be stable for thousands of generations and the bacteriophage may alter the phenotype of the bacterium by expressing genes that are not expressed in the usual course of infection in a process known as lysogenic conversion. Figure 1-6 outlines the two phage life cycles.

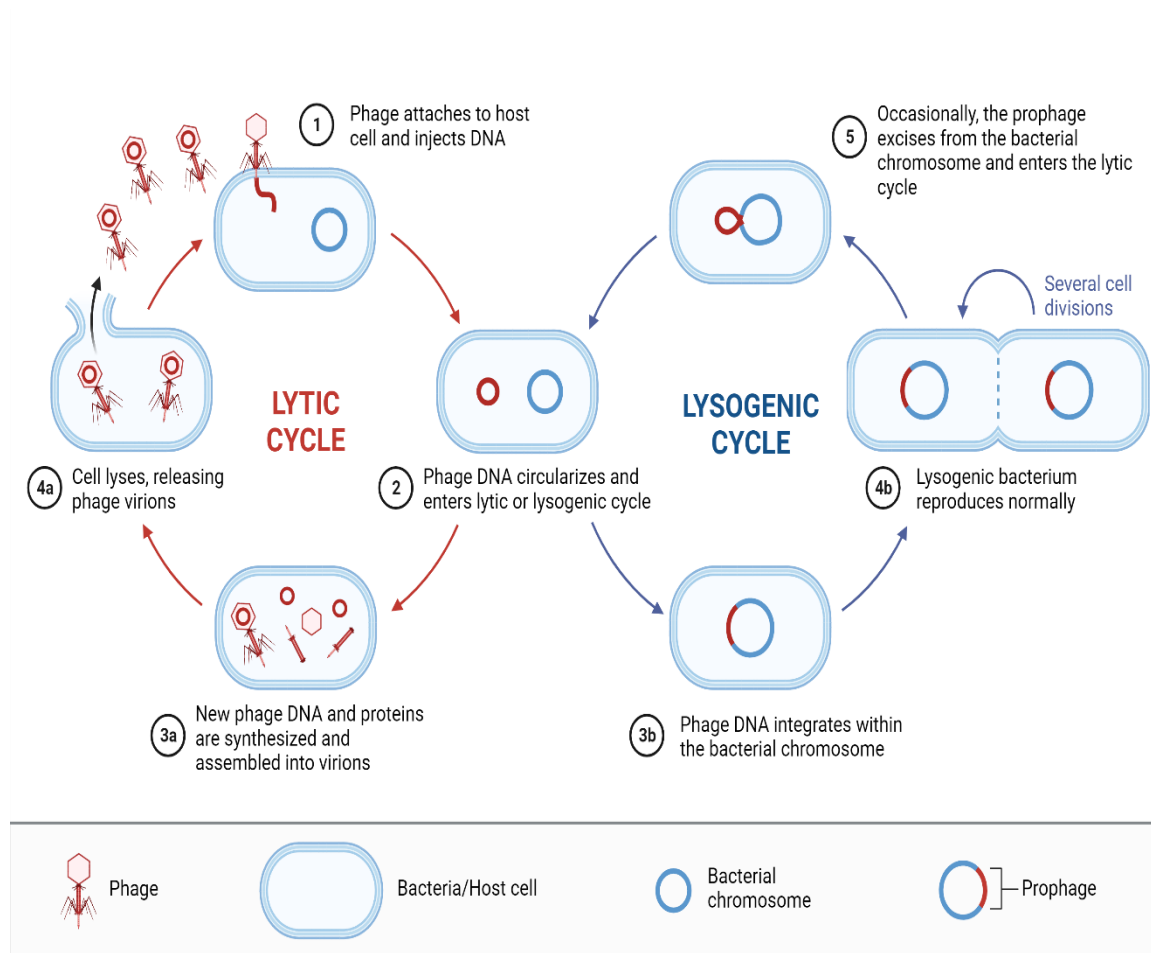


Figure 1-6. Comparison of the bacteriophage life cycles. Upon entering their hosts, bacteriophage can either enter the lytic cycle in which new phage virions synthesized are released resulting in bacterial cell lysis, or enter the lysogenic cycle in which phages integrate their DNA with that of the host and replicate together. (Image created with BioRender.com).

While beneficial, phage therapy does have certain drawbacks. First, phages usually have a very narrow host range and display strain specificity within a bacterial species thereby raising the need for multiple phages to treat multiple strains within a bacterial species. From a therapeutic point of view, patients would have to be strain-typed, which is a long and expensive process as compared to species identification. While the use of bacteriophage cocktails could circumvent this issue, the regulatory aspects associated with this are significant. Second, for the treatment of bacterial infections, only lytic phages must be selected to ensure killing of the bacteria as compared to lysogenic phages that would integrate with the host bacterium instead of killing it. Third, bacteria do develop resistance towards phages at every step of the phage infection cycle. Bacteria can do this in a couple of ways – by preventing phage binding and attachment by mutating or modifying bacterial surface receptor proteins (e.g. regulating receptor expression, receptor modification through interaction with another protein and sequestration of phages by outer membrane vesicles that can act as phage receptors thereby acting as decoys), by degrading the injected phage DNA via restriction-modification enzymes or the CRISPR-Cas system, by the bacteria committing cell death through the abortive infection system to reduce the spread of virions. Fourth, the pharmacokinetics of phage therapy is hard to measure because phages self-replicate and this presents problems with regards to phage therapy regulation. Thus, the use of an antimicrobial agent produced by phages, that possesses the ability to directly target and kill the bacteria, would be advantageous over phage therapy.

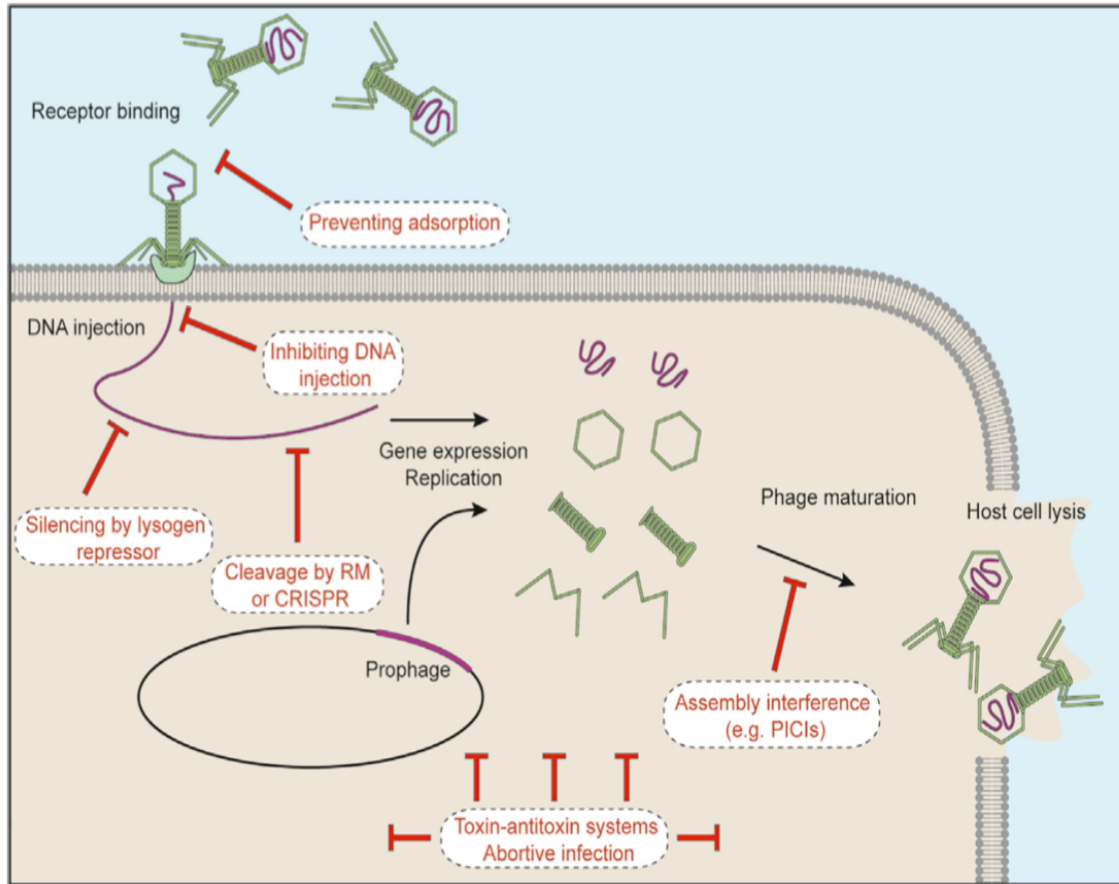


Figure 1-7. Overview of the bacterial resistance mechanisms to phages. Bacteria can develop resistance to phages at the level of phage binding and adsorption, by inhibiting DNA injection, by degrading injected phage DNA, by interfering with new phage assembly (e.g. PICs; phage inducible chromosomal islands) and by committing cell death. Image reproduced with permission from (Rostol & Marraffini, 2019).

History and Discovery of Endolysins

An alternative to antibiotics is the use of endolysins, also termed phage lysins or enzybiotics, which are bacteriophage-encoded peptidoglycan hydrolases (Nelson et al., 2006). Endolysins are produced at the end of a bacteriophage replication cycle resulting in cell lysis and virion release. These enzymes hydrolyze critical covalent bonds essential for maintaining cell wall structural integrity. While endolysins require a phage for the delivery of its DNA along with holins that form pores in the cytoplasmic membrane enabling access to the bacterial peptidoglycan to achieve “lysis from within”, extrinsic application of purified recombinant endolysin would remove these restrictions. Along these lines, these enzymes are capable of destroying the Gram-positive bacterial peptidoglycan (PG) rapidly and specifically when applied exogenously (Fischetti, 2006), known as “lysis from without”. Figure 1-8 compares both these lysis mechanisms. The cytoplasmic membrane is unable to maintain the structure of the cell due to the high internal turgor pressure of 20-50 atmospheres, resulting in cell death due to osmotic stress (Arnoldi et al., 2000). Endolysins have shown their efficacy against a variety of Gram-positive pathogens including *Streptococcus pneumoniae* (Entenza et al., 2005; Jado et al., 2003; Loeffler & Fischetti, 2003; Loeffler et al., 2001; McCullers et al., 2007), *Streptococcus pyogenes* (Hoopes et al., 2009; Nelson et al., 2001), *Streptococcus agalactiae* (Cheng et al., 2005; Pritchard et al., 2004; Pritchard et al., 2007), *Listeria monocytogenes* (Gaeng et al., 2000; Loessner et al., 2002), *Streptococcus uberis* (Celia et al., 2008), *Streptococcus suis* (Wang et al., 2009), *Staphylococcus aureus* (Donovan et al., 2006; O’Flaherty et al., 2005; Obeso et al., 2008; Rashel et al., 2007; Sass & Bierbaum, 2007; Takac et al., 2005), *Bacillus anthracis* (Porter et al., 2007; Schuch et al., 2002; Yoong et al., 2006), *Enterococcus faecalis* (Yoong et al., 2004) and *Clostridium perfringens* (Zimmer et al., 2002). Electron microscopy has enabled visualization of bacterial cell wall digestion caused by

endolysins, resulting in cell lysis as seen in Figure 1-9. Endolysins have thus been proposed as promising antimicrobial agents.

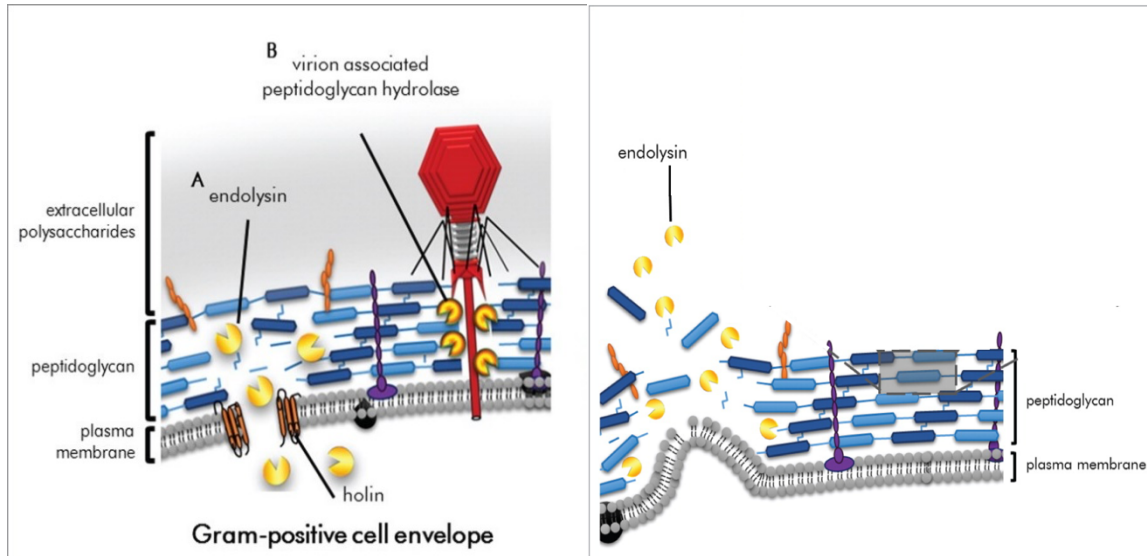


Figure 1-8. Structure of the Gram-positive bacterial peptidoglycan. (a) Endolysins can lyse the bacterial cell wall from inside (lysis from within) (b) Endolysins can also lyse the cell wall from outside (lysis from without). Image reproduced with permission from (Roach & Donovan, 2015).

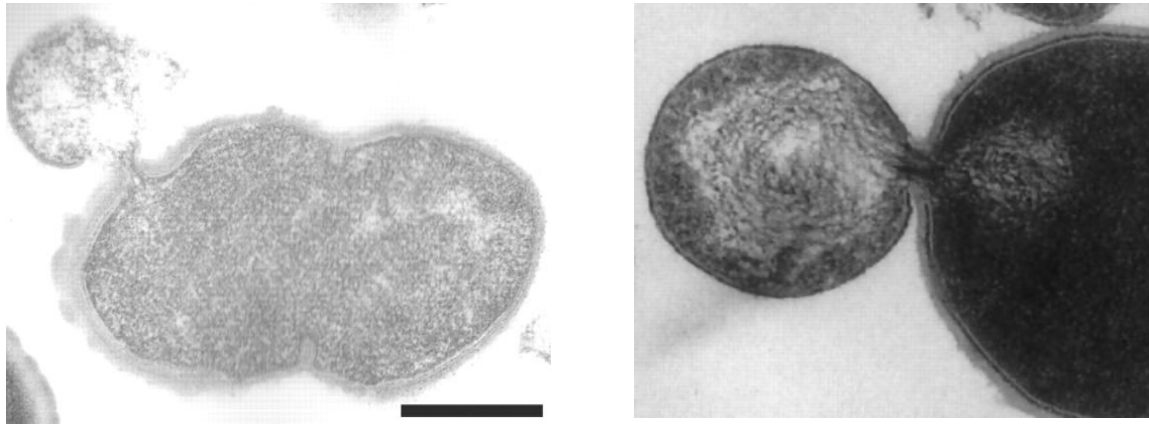


Figure 1-9. Electron micrographs of bacteria treated with endolysins. (a) *S. pneumoniae* treated with the Pal endolysin, causing the cell wall to rupture resulting in cytoplasmic leakage (b) *S. pyogenes* treated with the PlyC endolysin for 15 seconds. Total magnification is 50,000X, scale bar is 0.5 μm . Images reproduced with permission from (Loeffler et al., 2001) and (Nelson et al., 2001).

Endolysin Domain Structure

Endolysins derived from bacteriophages that infect Gram-positive hosts are typically 25-40 kDa and have very similar modular structures with one or more N-terminal enzymatically-active domains (EADs) and a C-terminal cell-binding domain (CBD) (Borysowski et al., 2006; Fischetti, 2010; Loessner, 2005). EADs are alternately abbreviated as CDs, which stands for catalytic domains and CBDs are alternatively abbreviated as CWBDs, which stands for cell-wall binding domains, both of which are seen in Chapter 2. The EAD is responsible for the hydrolytic activity of the enzyme by cleaving covalent bonds in the bacterial peptidoglycan. The peptidoglycan structure is comprised of a linear polysaccharide backbone of alternating units of β -1,4-linked N-acetylglucosamine (GlcNAc; NAG) and N-acetylmuramic acid (MurNAc; NAM). Enzymes with catalytic activity against the reducing ends of N-acetylglucosamine or N-acetylmuramic acid are referred to as N-acetylglucosaminidases or N-acetylmuramidases (also muramidase or lysozyme) respectively; collectively they are known as glycosidases. An N-acetylmuramoyl-L-alanine amidase (often referred to simply as “amidase”) cleaves the amide bond between MurNAc and L-alanine. Endopeptidases can hydrolyze bonds between two peptides, whether they are in the stem or the interpeptide bridge. In some cases, endolysins have two EADs connected to each other at the N-terminus, however, this does not necessarily correlate to increased activity compared to endolysins possessing a single EAD (Becker et al., 2009). In contrast, the CBDs structure is evolutionarily variable and possess no enzymatic activity but rather function to bind a specific substrate, usually a carbohydrate or teichoic acid attached to the host peptidoglycan. Due to its variable nature, the CBDs confer species, strains and even serotype specificity. Thus, the specificity of the endolysin as a whole is often dictated by the specificity of the CBD. While most EADs do require the presence of a CBD for binding to the bacterial surface and subsequent

activity, some EADs can interact independently of a CBD, such as the T4 lysozyme (Matthews & Remington, 1974).

Endolysins derived from bacteriophages that infect Gram-negative hosts are typically 15-20 kDa and consist of a singular globular EAD. Most of these endolysins lack a CBD. It has been hypothesized that the lack of wall teichoic acids/carbohydrates in the Gram-negative peptidoglycan could be the reason most Gram-negative endolysins lack a CBD, while their EADs bind directly to the peptidoglycan (Stark, 2010). There are however two exceptions to this, namely endolysins KZ144 and EL188 that consist of an N-terminal CBD and a C-terminal EAD.

Endolysins as Antimicrobials/Bacteriolytic Mechanism

The widespread distribution of antibiotic-resistance genes in bacteria compels us to the post-antibiotic era. Diseases that were once treatable through use of antibiotics are becoming deadly again (Brown & Wright, 2016). Moreover, the overuse of broad-spectrum antibiotics accelerates the dispersal of resistance genes not only to pathogens, but also to commensal organisms (Nelson et al., 2012). This makes development of alternative antimicrobial agents one of the highest global priorities for aggressive biological, pharmaceutical, and medical investigations (Veiga-Crespo et al., 2007). The bacterial peptidoglycan is essential for maintaining cell shape and counteracting the high intracellular osmotic pressure (Royet & Dziarski, 2007). As mentioned, the peptidoglycan structure consists of a linear polysaccharide backbone with alternating units of β -1,4-linked N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) where the lactyl group of every MurNAc unit is substituted with stem peptides that are typically four alternating L- and D-amino acids. With the exception of bacilli, for Gram-positive bacteria the four amino acids in order starting from MurNAc are L-alanine, D-iso-glutamic acid,

L-lysine and D-alanine. This is called the Lys-type peptidoglycan. For bacilli and Gram-negative bacteria, the L-lysine is replaced with meso-diaminopimelic acid (mDAP). This is called the DAP-type peptidoglycan. Figure 1-10 compares the DAP-type and the Lys-type peptidoglycan structure. Stem peptides that are adjacent to one another are cross-linked directly by mDAP to D-alanine (interpeptide bond, in the case of Gram-negatives and Gram-positive bacilli) or by an interpeptide bridge of amino acids (in the case of other Gram-positive bacteria). The interpeptide bridge in *Staphylococcus aureus* contains a pentaglycine cross-bridge while *Streptococcus pyogenes* has an interpeptide bridge connected by dialanine.

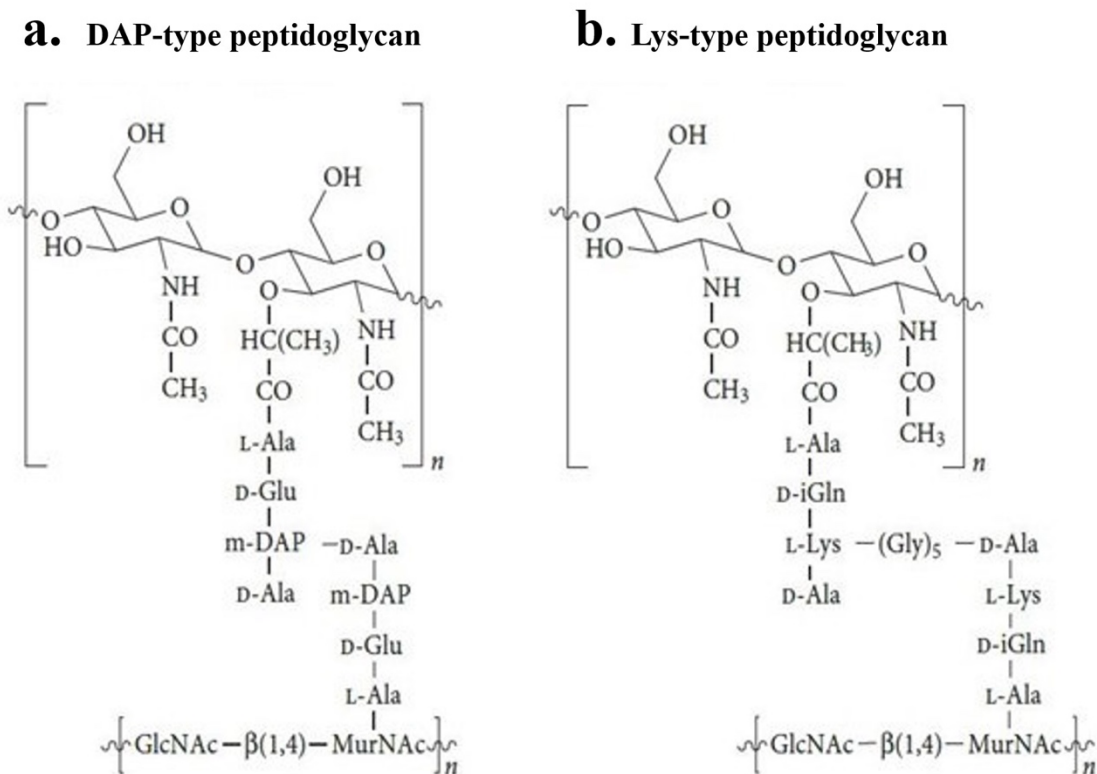


Figure 1-10. Comparison of DAP-type vs Lys-type peptidoglycan structure. (a) DAP-type peptidoglycan is specific to Gram-negative and Gram-positive bacilli species. (b) Lys-type is present in non-bacilli Gram-positive species. Image reproduced with permission from (Royet & Dziarski, 2007).

Endolysins can be classified into five groups based on the bonds they target: N-acetylmuramidases, endopeptidases, N-acetylmuramoyl-L-alanine amidases, lytic transglycosylases, and N-acetyl- β -D-glucosaminidases (Borysowski et al., 2006). Figure 1-11 shows the bacterial peptidoglycan structure and the bonds cleaved by various EADs. The cysteine, histidine-dependent amidohydrolase/peptidase (CHAP) domain is the most common EAD observed in endolysins. Some CHAP domains have been reported to display amidase activity (Nelson et al., 2006) while others have shown to possess endopeptidase activity (Becker et al., 2009; Pritchard et al., 2004). CHAP uses a cleavage mechanism where a cysteine is deprotonated by a nearby histidine promoting a nucleophilic attack by the cysteine on the scissile bond.

As antimicrobial agents, endolysins feature a narrow host spectrum and kill bacteria regardless of their antibiotic sensitivity. There is also evidence indicating that the development of resistance to endolysins is unlikely, the use of endolysins is safe *in vivo*, and will not cause a major immune response (Borysowski et al., 2006). In February 2020, the US FDA granted “Breakthrough Therapy” designation status to endolysins, thus validating endolysin technologies as antimicrobial biologics.

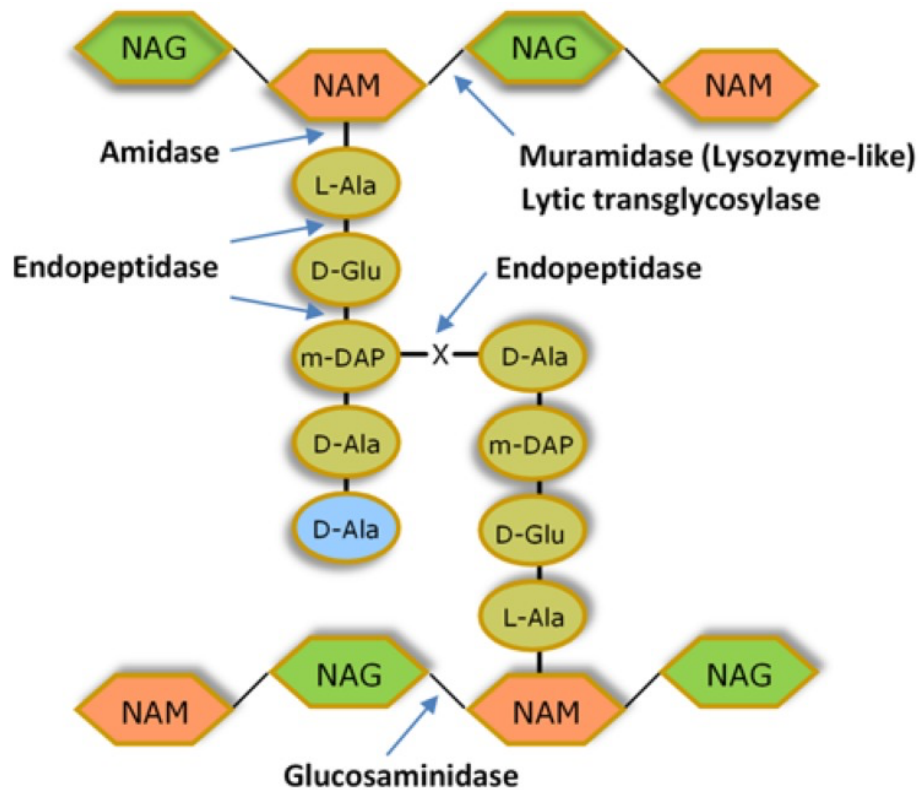


Figure 1-11. Basic structure of the bacterial cell wall peptidoglycan. The possible enzymatic activities of endolysins and the bonds they cleave are shown. NAG is the acronym for N-acetylglucosamine and NAM stands for N-acetylmuramic acid. Image reproduced with permission from (Sao-Jose, 2018).

Cell-wall Binding Domains

The cell-binding domains (CBDs) bind to the bacteria peptidoglycan, lipoteichoic acids or other cell wall components. Endolysin CBD binding has been shown by fusing to fluorescent markers, such as green fluorescent protein (GFP), and by conjugation to fluorophores, such as AlexaFluor®. In most cases, the presence of a CBD is essential for lytic activity as evidenced by a loss of activity upon deletion of the CBD (Garcia et al., 1994; Huang et al., 2015; Vander Elst et al., 2020). There is also a hypothesis suggesting that CBDs may a significant role in preventing the development of resistance (Becker et al., 2009). In the case of certain endolysins, research has demonstrated that the CBD strongly adheres to cell wall debris following cell lysis. This prevents the enzyme from potentially diffusing and harming other non-infected bacterial cells (Loessner et al., 2002).

Advantages of Endolysins

Endolysins have many benefits over classical antibiotics, as well as over the use of whole bacteriophages. Most importantly, despite studies conducted to isolate resistant mutants, no resistance to endolysins has been observed to date (Pastagia et al., 2011). The *Bacillus anthracis* endolysin, PlyG, was repeatedly added to *Bacillus* cultures in the presence of a mutagenic chemical methanesulphonic acid ethyl ester in order to expedite the evolution of PlyG resistant bacterial strains. No PlyG resistant strains of *Bacillus* were observed, but there was a 1000-10,000 fold increase in novobiocin and streptomycin resistance (Schuch et al., 2002). Similarly, the pneumococcal endolysin Pal was repeatedly added to *Streptococcus pneumoniae* cultures using increasing concentrations in liquid culture and no Pal resistant strains of *S. pneumoniae* were observed (Loeffler et al., 2001). Additionally, an engineered endolysin, ClyS targeting MRSA,

displayed no change in MIC after repeated treatment with endolysin, whereas the MIC of the antibiotic mupirocin increased ten-fold over an eight-day period (Pastagia et al., 2011). One possible explanation for the lack of resistance observed is that, given the fast-acting, high lytic efficacy of endolysins acting in a matter of seconds or minutes, this leaves no time for selective pressure to generate resistant mutants. Another likely explanation is that bacteriophages and bacteria have coevolved over millions of years, and that has led to endolysins binding and cleaving highly conserved targets in the cell wall peptidoglycan that are required for bacterial viability (Fischetti, 2005). This is exemplified by the epitopes targeted by pneumococcal endolysins that bind to choline on the pneumococcal surface; with choline being essential for cell viability (Garcia et al., 1983; Lopez & Garcia, 2004). Other streptococcal endolysins have been shown to bind to polyrhamnose in *S. pyogenes* which is critical for cell survival (Fischetti, 2003; Yamashita et al., 1999). While resistance to endolysins has not been observed, it is important to note that resistance to other peptidoglycan cell wall hydrolases, namely human lysozyme and bacteriocins has been observed. Various cell wall modifications such as O-acetylation and N-deacetylation of the peptidoglycan and D-alanylation of teichoic acids has rendered bacterial strains resistant to lysozyme (Guariglia-Oropeza & Helmann, 2011; Vollmer, 2008). *S. aureus* develops resistance against lysostaphin, a bacteriocin, by altering the structure of the peptidoglycan crossbridge by downsizing from a pentaglycine to a single glycine, which is the target of this enzyme (Dehart et al., 1995). In addition, resistance to bacteriophages has been observed (Labrie et al., 2010). Finally, the ease of accessibility to the bacterial cell wall makes extrinsic application of endolysins not susceptible to common resistance mechanisms such as degradation by intracellular antimicrobial proteins, expulsion by efflux pumps, etc.

Endolysins are typically species-specific as opposed to most bacteriophages that are strain specific. Some endolysins also display a broad host-range wherein they target multiple different species of bacteria. This wide spectrum of action makes endolysins preferable over bacteriophages since they can be used to treat a particular bacterial infection without the need to know the exact strain. Compared to antibiotics, the species-specificity displayed by endolysins is also a favorable characteristic. The misuse and overuse of antibiotics disturbs the normal flora resulting in the growth of opportunistic pathogens, such as *Clostridium difficile* in the gastrointestinal tract, which can result in acceleration of horizontal gene transfer of resistance within the bacterial community (Nelson et al., 2012). Whereas antibiotics can also cause nausea, diarrhea and vaginal/oral yeast infections due to eradication of the typical species that colonize these environments, resulting in the overgrowth of infectious bacteria (Tedesco et al., 1974), an endolysin kills a targeted pathogen. Antibiotics are also known to cause allergic reactions as one of their side-effects, ranging from mild (itchiness, redness, swelling of skin) to severe (anaphylaxis). Even though endolysins are proteinaceous in nature, they have been shown to not cause any of the allergic reactions typically seen with antibiotics.

Given their modular structure, endolysins are also amenable to various engineering approaches. Chimeragenesis (i.e. domain swapping), directed evolution, and peptide fusions are some of the approaches being investigated. Rationale-based methods (e.g. sequence comparison, site-directed mutagenesis (SDM), *in silico* computational modeling) and random methods (e.g. random mutagenesis, gene shuffling, directed evolution) can be used to engineer desirable properties in endolysins. Other pharmaceutically desirable attributes for engineering phage lytic enzymes, in addition to their bactericidal activity include expansion of the lytic spectrum, increased solubility, greater stability, and longer circulating half-life (Oliveira et al., 2018). Some

examples of endolysins engineered to possess improved activity include: PL3, a chimeric endolysin engineered to possess the EAD of the pneumococcal endolysin Pal and the CBD of the pneumococcal autolysin LytA, exhibited higher lytic activity and stability as compared to the parent enzymes (Blazquez et al., 2016). Cpl-711, a chimeric pneumococcal endolysin comprising the EAD from Cpl-7 and CBD from Cpl-1, displayed improved lytic activity and biofilm eradication ability compared to parent enzymes (Diez-Martinez et al., 2015). Another chimera, ClyJ, containing a CHAP EAD derived from PlyCA and a CBD derived from gp20, which shares a high degree of similarity with that of LytA was created (Yang et al., 2019). ClyJ was further modified by altering the length of the linker sequence between the EAD and CBD to create ClyJ-3, which possesses higher antimicrobial activity than ClyJ (Yang et al., 2020). Subsequently, ClyJ-3 was altered by shortening the C-terminal tail, yielding ClyJ-3m which had higher activity than ClyJ-3 and Cpl-1 (Luo et al., 2020). Other than activity, chimeric endolysins can also be engineered to expand their lytic spectrum. For example, the CHAP domain of PlyC was fused to the CBD of the PlySs2 endolysin creating the enzyme ClyR. ClyR was active against streptococcal, staphylococcal and enterococcal species, similar to PlySs2, but it was also active on *S. mutans*, a species the wild-type (WT) PlySs2 showed no activity against (Yang et al., 2015).

Charge has also been shown to play an important role in endolysin activity and engineering approaches. A study conducted by Low et al. demonstrated that an EAD can operate independently of its CBD due to its inherent positive charge. This independence is likely facilitated by ionic interactions with the bacterial surface, which typically carries a negative charge owing to the presence of surface carbohydrates (Low et al., 2011). Endolysin engineering studies took advantage of this concept to increase lytic activity and broaden the host range. Modulation of the charge of the CBD can also improve endolysin activity as seen in the case of Cpl-7. Cpl-7 is a

pneumococcal endolysin displaying bacteriolytic activity against *S. pneumoniae*, albeit lower than other pneumococcal endolysins such as Cpl-1 and Pal (Diez-Martinez et al., 2013). The wild-type CBD of Cpl-7 possesses a negative charge at neutral pH. In light of the findings by Low et al., Diez-Martinez et al. altered the charge of the CBD of Cpl-7 from -14.93 to +3 by conducting site-directed mutagenesis of 13 non-structural amino acids. This new mutant was named Cpl-7S and it displayed improved bacteriolytic activity on pneumococcal and non-pneumococcal species when compared to its parent enzyme Cpl-7 (Diez-Martinez et al., 2013). While it is certainly possible to engineer bacteriophages for improved lytic potential, it is comparatively facile to modify individual proteins.

The efficacy of endolysins is not significantly compromised by the host's production of neutralizing antibodies, presumably because the kinetics of lytic action is much faster. Even so, the short plasma half-life of Cpl-1 was increased by introducing cysteine residues at the C-terminal end, to facilitate disulfide bond formation resulting in a more stable homodimer that was twice as active *in vitro*, with a 10-fold higher bioavailability in mice (Resch et al., 2011b). By contrast, PEGylation of Cpl-1 abolished bacteriolytic activity (Resch et al., 2011a) while PEGylation of other peptidoglycan hydrolases such as lysostaphin and bacteriocin improved pharmacokinetics while not affecting activity (Walsh et al., 2003). This suggests that bioengineering approaches to improve endolysin properties may prove to be more successful as compared to chemical approaches such as PEGylation that are otherwise commonly used to slow down systemic clearance of molecules by reducing antibody binding affinity.

Finally, endolysins have shown efficacy in clearing intracellular infection. Intracellular pathogens are typically troublesome because they can evade the host immune system and antimicrobials cannot penetrate the eukaryotic cell membrane. Endolysins have been fused to

cationic protein transduction domains (PTDs), also called cell-penetrating peptides to facilitate internalization. An *S. aureus* endolysin fused to a PTD was shown to kill both drug-sensitive and drug-resistant intracellular *S. aureus* in cultured bovine mammary epithelial cells and murine osteoblasts (Becker et al., 2016). The PlyC endolysin was fused to PTDs and was able to kill intracellular *S. pyogenes* (Shen et al., 2016). Interestingly, wild-type PlyC was also able to cross the epithelial cell membrane and kill intracellular *S. pyogenes* on its own, without a PTD. The authors of the study determined this to be due to its PlyCB binding subunit interacting with phosphatidylserine of the cell membrane facilitating extracellular to intracellular entry. This highlights the potential for endolysins to be used to treat intracellular pathogens either by fusing them to PTDs or on their own and this can also aid in engineering approaches that could be applied to different endolysins to target other bacteria.

Biofilms, Synergy, Safety & Immunogenicity

Biofilms have been associated with up to 80% of chronic infections (Wolcott & Ehrlich, 2008), particularly those associated with surgical implants. Biofilms are sessile microbial communities that elaborate extracellular polymeric substances (EPS) and are typically composed of DNA, proteins and polysaccharides. Biofilms are formed in response to stress and harsh conditions to promote bacterial survival. Compared to planktonic cells, many cells present in the biofilm become metabolically inactive and the EPS based matrix also contributes to community recalcitrance to antibiotics and immune cells (Lewis, 2008). Various factors affect the formation of biofilms, such as nutrients, pH, osmotic pressure, and antibiotics (Sun et al., 2023). *S. pneumoniae* is known to form bacterial biofilms, with researchers having found *S. pneumoniae* biofilms in the nasal mucosa of patients with chronic sinusitis. While the formation of

pneumococcal biofilms on the mucosal epithelium has been documented, the composition of the pneumococcal biofilm EPS has not been elucidated (Moscoso et al., 2009). Studies have indicated that *S. pneumoniae* biofilms are not effectively cleared during antimicrobial treatments due to increased antimicrobial resistance, raising the need for alternatives. Endolysins have been shown to be effective in eradicating bacterial biofilms. The pneumococcal endolysins Cpl-1, Cpl-7 and the pneumococcal autolysin, LytA, have all demonstrated biofilm disintegration ability. Synergy with respect to biofilm eradication has been shown when LytA and Cpl-1 used in combination led to an augmented eradication in biofilm compared to when each endolysin was used in isolation (Domenech et al., 2011).

Endolysins have shown promise when used synergistically with other endolysins or antibiotics, both *in vitro* and *in vivo*. Synergy has a couple of benefits. The first being, in order to achieve synergy between two or more antimicrobials means that the effective dose for each individual antimicrobial is lower and this reduces the likelihood of development of resistance. Second, synergy between two endolysins that have two different peptidoglycan cleavage sites but act simultaneously has been shown to have enhanced destabilizing effects (Schmelcher et al., 2012). It is also likely that the first endolysin may cleave a bond in the peptidoglycan giving the second endolysin better substrate access to the bond that it cleaves, thereby acting on multiple unique targets and ultimately resulting in lysis of the bacterial cell. The synergistic action of endolysins has been demonstrated for the pneumococcal endolysins Cpl-1 and Pal. Cpl-1, a muramidase and Pal, a L-alanine amidase, displayed synergy both *in vitro* and *in vivo* (Jado et al., 2003; Loeffler & Fischetti, 2003). In a murine sepsis challenge, the bacteremic titer was reduced to a greater extent when both Cpl-1 and Pal were used together compared to when each endolysin was used alone. Synergy has also been shown between the *S. aureus* endolysin LysK (L-alanine

amidase) and lysostaphin (glycyl-glycine endopeptidase) *in vitro*, even on strains that are normally resistant to lysostaphin (Becker et al., 2008). Not only has synergy been demonstrated between endolysins, it has also been shown when combining the use of endolysins and antibiotics. Cpl-1 in combination with antibiotics like daptomycin, penicillin, gentamycin, or azithromycin shows synergistic effects (Djurkovic et al., 2005; Vouillamoz et al., 2013). The combination of Cpl-1 and penicillin was shown to have synergistic killing effects on *S. pneumoniae* strain 8249, a strain that has a high penicillin MIC. When Cpl-1 and gentamicin were administered together, the final bacteremic titer was more than 1000-fold-lower as compared to when they were used in isolation. In a mouse model of pneumococcal bacteremia, the percentage of mice that survived infection was significantly increased when daptomycin and Cpl-1 were delivered in combination compared to those that either did not receive any treatment or those that received daptomycin and Cpl-1 in isolation. The chimeric endolysin ClyS, also displayed synergy with oxacillin in an *in vivo* mouse model of MRSA-induced septicemia (Daniel et al., 2010). This combination of antibiotics with endolysins represents a promising therapeutic strategy to treat and control infectious diseases.

Obtaining comprehensive preclinical safety and toxicity assessments using mammalian cells and tissues is of utmost importance to pave the way for the future translational advancement of endolysins. A study conducted in 2018 successfully demonstrated that concerning immune responses, alternations in the microbiome and inflammatory responses, the pneumococcal endolysins Cpl-1 and Pal exhibited an overall safe profile (Harhala et al., 2018). N-Rephasin® (SAL200), an endolysin with antimicrobial activity against *S. aureus*, has undergone evaluation in Phase 2a clinical trials. In rodents and dogs, SAL200 displayed no signs of toxicity in the respiratory system and central nervous system (Jun et al., 2014). Likewise, no adverse events were

observed in monkeys (Jun et al., 2016) and SAL200 was well tolerated in a Phase 1 human trial (Jun et al., 2017).

Because endolysins employ a bacteriolytic mechanism, their sustained use in humans or animals may potentially lead to the release of proinflammatory bacterial components like peptidoglycan and teichoic acids. This, in turn, could give rise to conditions such as septic shock or multiple organ failure (Nau & Eiffert, 2002). *In vivo* studies suggest varying endolysin concentration could lessen this problem. In one sepsis study conducted in mice, it was observed that the continuous intravenous administration of the Cpl-1 endolysin led to an increase in the levels of proinflammatory cytokines (Entenza et al., 2005). However, in a comparable study, lower concentrations of Cpl-1 were found to decrease cytokine levels compared to infected mice that were not treated (Witzenrath et al., 2009). To provide an explanation for this observation, it was theorized that employing higher concentrations of endolysins triggers a more thorough breakdown of the bacterial cell wall. Consequently, this could result in the production of elevated levels of proinflammatory cell wall pieces that circulate in the bloodstream. As such a proinflammatory response would be likely due to lysed bacterial cells rather than the endolysins themselves. Immunogenicity studies have also been conducted *in vitro* and *in vivo*. Antibodies have been successfully developed against endolysins targeting *S. pneumoniae*, *S. aureus*, *S. pyogenes*, and *B. anthracis* (Fischetti, 2010) however they did not neutralize their lytic efficacy (Jado et al., 2003; Rashel et al., 2007). The hyperimmune serum obtained after mice were administered Cpl-1 resulting in generation of antibodies did slightly decrease Cpl-1 activity *in vitro*, however *in vivo* they did not reduce the efficacy of Cpl-1 in protecting mice from a *S. pneumoniae* infection. It has been hypothesized that the binding affinity of CBD for bacterial surface being higher than that of

IgG for endolysin (*L. monocytogenes* endolysins have nanomolar binding affinities) and kinetics of lytic activity being so rapid, enable endolysins to evade the host's immune system.

Methods to Determine Endolysin Activity

To assess and evaluate the bacteriolytic activity of endolysins, several *in vitro* assays have been developed. Assays that have been used in this dissertation research have been described below.

Turbidity Reduction Assay

The turbidity reduction assay (TRA) as the name implies, is a technique used to measure the ability of an enzyme to reduce the turbidity or cloudiness of a liquid sample. Turbidity is due to the presence of microorganisms or suspended particles in the liquid. This assay measures the decrease in optical density (OD) over time using a spectrophotometer at a typical wavelength of 600nm (OD₆₀₀). Peptidoglycan hydrolysis caused by the endolysin is measured which results in osmotic lysis and cell death thereby resulting in a reduction in turbidity of the liquid that can be quantified. Results are typically compared to an untreated negative control. The TRA is usually performed in a 96-well plate. An overnight bacterial culture is started and the next day the cells are washed and resuspended in desired buffer to a starting OD₆₀₀ of 1. Cells are mixed with endolysin treatment and the OD₆₀₀ is monitored every 15 s for 30 min at 37°C. Lytic activity is then quantified as the reduction in turbidity and measured as the difference in OD₆₀₀ between control and endolysin treated cells (Nelson et al., 2012). It is important to note that in a TRA, cells that are in an exponential phase of growth are more sensitive to endolysin treatment compared to cells in stationary phase, presumably due to thinner cell walls.

Colony Forming Unit Counting Assay

The Colony Forming Unit (CFU) assay is a very sensitive and accurate assay because it involves the counting of individual bacterial colonies. In this assay, sterile protein is mixed with bacteria in a 96 well plate and incubated at 37°C for an hour. Then, five 10-fold dilutions in PBS are made and 10 µl from each dilution is plated on THY agar plates, air-dried and placed in a 37°C incubator overnight. The individual colonies are counted and the data is reported as the log-fold killing compared to the untreated control.

Minimum Inhibitory Concentration Assay

The Minimum Inhibitory Concentration (MIC) assay is a microbiological assay used to determine the lowest concentration of an antimicrobial agent that can effectively inhibit the growth of a specific microorganism. The MIC assay is important in assessing the susceptibility or resistance of a microorganism to a particular treatment. In an MIC assay, an overnight bacterial culture is started and is diluted (typically 1:100) in fresh media the next day. Serial dilutions of endolysin are prepared and the diluted bacterial cells are mixed with each endolysin concentration in a 96-well flat-bottom plate. The plate is incubated overnight in a 37°C incubator and observed the next day for the lowest concentration at which there is a clear well with no bacterial growth. This represents the MIC for the particular treatment.

Nasopharyngeal Colonization Assay

In the nasopharyngeal colonization assay, mice are infected with 1×10^5 CFU of bacteria in each nostril. After 48 hours of inoculation, a certain number of mice receive PBS (control) treatment, and others receive endolysin treatment. Each nostril of every mouse receives 25 μ l of either PBS or endolysin treatment. The mice are then treated for five hours, sacrificed, and their nose tissue is ground. This is then plated on blood agar plates and the CFU's are counted. Endolysins used in this assay are sterile and are passed through an endotoxin removal column prior to administration.

Biofilm Eradication

The biofilm eradication assay is performed in 96-well flat-bottom polystyrene plates. Bacteria are grown to an OD₆₀₀ of 0.5 to 0.6, sedimented by centrifugation and resuspended in an equal volume of fresh media. The cells are then diluted 1:2 into a 15 ml tube containing fresh media and dispensed (200 μ l) into each well of the 96-well plate. The plate is then incubated at 37°C overnight and the next day, the media is carefully pipetted out from each well and the biofilm formed gently washed twice with PBS to remove unattached cells. Serial dilutions of endolysin are added to treatment wells and the plate is incubated at 37°C for 1 hour. The liquid is then pipetted out and the wells washed with distilled water and left to air-dry. The formed biofilm is stained with 50 μ l of a 0.2% crystal violet solution and incubated at room temperature for 15 minutes. After pipetting out excess crystal violet, the wells are washed with distilled water and air dried. The biofilm is solubilized in 10% SDS to extract crystal violet from the biomass, transferred to a fresh 96-well plate and then quantified by determining the A₅₉₅ using a SpectraMax M5 plate reader.

S. pneumoniae Endolysins

The first *S. pneumoniae* endolysins to be studied *in vivo* were Pal, derived from the Dp-1 phage, and Cpl-1, derived from the Cp-1 phage. Pal was shown to protect mice in a nasopharyngeal colonization model (Loeffler et al., 2001), whereas Cpl-1 showed efficacy in a mouse intravenous bacteremia model (Loeffler et al., 2003), a mouse otitis media model (McCullers et al., 2007), a severe mouse pneumonia model (Doehn et al., 2013), a rat endocarditis model (Entenza et al., 2005), and by intracisternal administration in a rat pneumococcal meningitis model (Grandgirard et al., 2008). Aside from Pal and Cpl-1, both native endolysins, the majority of research efforts in the past 5-10 years have focused on chimeragenesis and/or engineering of native phage endolysins. In one example, Cpl-7, a native pneumococcal endolysin with an *N*-acetylmuramidase EAD, was engineered to create Cpl-7S, a variant that contained inversion of charge in 15 amino acids of the CBD (Diez-Martinez et al., 2013). Cpl-7S showed improved lytic activity when benchmarked against Cpl-7 *in vitro* and *in vivo*. In a separate study, Cpl-7 was engineered to replace the native CBD with the CBD from Cpl-1. This new variant, called Cpl-711 showed improved antimicrobial activity *in vitro* and *in vivo* over either of the parental enzymes (Diez-Martinez et al., 2015). Synergy between endolysins was first described in pneumococcal endolysins. The EAD of Pal is an *N*-acetylmuramoyl-l-alanine amidase, which cleaves between the glycan strand and the stem peptide of the peptidoglycan whereas the EAD of Cpl-1 is an *N*-acetylmuramidase, cleaving the glycan strand directly. A synergistic effect was demonstrated *in vivo*, with 200 µg of either Pal or Cpl-1 required to protect mice intraperitoneally from a fatal pneumococcal infection, but a cocktail containing 2.5 µg of each endolysin displayed the same efficacy (Jado et al., 2003).

Unique *S. pneumoniae* Peptidoglycan Structure

The *S. pneumoniae* peptidoglycan structure consists of alternating units of N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) linked together by β (1,4) bonds. Three to five amino acid long short peptide chains are bound to MurNAc residues. The pneumococcal cell wall is also composed of layers of peptidoglycan that contain teichoic acids (TA) covalently bound to wall teichoic acids or to the membrane lipoteichoic acids (Denapate et al., 2012; Gisch et al., 2013; Seo et al., 2008). The basic repeating unit for TA is comprised of 2-acetamido-4-amino-2,4,6-trideoxy-D-galactose, D-glucose, ribitol-phosphate and two residues of N-acetylgalactosamine (GalNAc) in which the GalNAc residues are decorated with phosphorylcholine, a unique feature only observed on *S. pneumoniae* Gram-positive bacteria (Brundish & Baddiley, 1968; Fischer et al., 1993; Karlsson et al., 1999; Seo et al., 2008). Phosphorylcholine residues are involved in a variety of physiological functions. Phosphorylcholine plays a role in the attachment of bacteria to G-protein coupled platelet-activating factor enabling adherence to the host cell demonstrated by the fact that pneumococcal strains lacking phosphorylcholine are unable to colonize the upper respiratory track in mice and have diminished virulence in a murine sepsis model (Cundell et al., 1995; Kharat & Tomasz, 2006). Choline residues serve as binding sites for various surface proteins and also for bacteriophages (Lopez & Garcia, 2004). In addition, phosphorylcholine groups are recognized by immune system components, C-reactive protein (Volanakis & Kaplan, 1971) and L-ficolin (Vassal-Stermann et al., 2014), that induce complement activation leading to host protection against infection. The importance of choline in *S. pneumoniae* cell viability is evidenced by choline-independent pneumococcal mutants that failed to grow normally. Finally, ethanolamine, a choline analog can also be used to grow *S. pneumoniae*, however, the bacteria are unable to

undergo autolysis, genetic transformation, form long chains and are resistant to bacteriophages (Tomasz, 1968).

Bacteriophage and Endolysin-Related Applications

The first *in vivo* endolysin study was conducted in 2001 by the Fischetti group to explore the therapeutic potential of endolysins (Nelson et al., 2001). The authors first administered a single 250 U dose of the PlyC endolysin and then infected mice with 10^7 colony forming units (CFU) of live *S. pyogenes*. Their results showed that mice that were pre-treated with PlyC were protected from bacterial colonization in the oral cavity compared to untreated control mice. Following these attractive *in vivo* antimicrobial results, endolysins started to be actively explored for their commercial use.

A clear and evident use for endolysins exists within the realm of medicine. Staphitekt™ was marketed by Micros in 2013 and is the world's first endolysin available for human use on intact skin. Staphitekt™ targets *Staphylococcus aureus* including antibiotic resistant strains such as MRSA (methicillin-resistant *Staphylococcus aureus*). Staphitekt™ is available in the form of a cream and gel to treat skin conditions with an infectious component, like acne, eczema, rosacea and skin irritation. Staphitekt™ is a chimeric lysin comprised of the M23 glycyl-glycine endopeptidase domain from lysostaphin, the amidase domain from Ply2638 and the SH3b cell-wall binding domain from staphylococcal phage endolysin Ply2638.

In the United States, Contrafect Corp. is the first company to enter an endolysin for clinical trials. Exebacase (CF-301) is their lysin with potent antimicrobial activity against *Staphylococcus aureus*. Contrafect Corp. recently completed a Phase 2 clinical trial for Exebacase evaluating its safety efficacy and tolerability in patients with *S. aureus* bacteremia and proceeded with Phase 3

trials, although they were stopped halfway due to the power of the study being below the pre-specified threshold for futility. Tonabacase (SAL200), a novel endolysin based drug targeting *S. aureus* developed by the Korean company iNtRON Biotechnology, has completed successful Phase 1a, 1b, and 2a clinical trials in Korea and the IND application for the U.S. Phase 2 clinical trial has been approved by the FDA in January 2022.

Endolysins have also been used to treat ocular infections. The standard treatment for ocular infections typically consists of application of topical broad-spectrum antibiotics, a practice which has led to an increase in antimicrobial resistance as well as disruption of the beneficial ocular microbiome. Recently, the pneumococcal endolysin MSlys was used to target and kill *S. pneumoniae* that caused ocular infections such as conjunctivitis, keratitis, and periocular cellulitis (Silva et al., 2022). The authors found that MSlys was able to successfully and rapidly kill all *S. pneumoniae* strains isolated from various ocular infections. This type of targeted approach that spares the beneficial ocular microbiome is highly desirable in nature.

Endolysin CBDs can be used to create a targeted immunotherapeutic by fusing them to the Fc region of human IgG. Raz et al. created a fusion consisting of CBDs from endolysins targeting MRSA attached to the Fc region of human IgG (Raz et al., 2017). They termed this construct a “lysibody”. These lysibodies displayed high affinity binding to the bacterial surface, induced fixation of complement on the bacterial surface of *S. aureus* and promoted phagocytosis by macrophages and neutrophils. This approach utilizing the high binding affinity of endolysin CBDs to clear bacterial infections is an attractive therapeutic and could be expanded to target a range of pathogenic bacteria.

Another application where endolysins can be used is in the detection and control of foodborne pathogens. The CDC estimates that 1 in 6 people each year are infected by foodborne

pathogens after consuming contaminated foods resulting in several hospitalizations and deaths. According to the FDA's Center for Food Safety and Applied Nutrition (CFSAN), the most common foodborne pathogens responsible for causing disease in humans in the United States are *E. coli*, *L. monocytogenes*, *B. cereus*, *S. aureus*, *C. perfringens*, and *C. botulinum*. Of note, several lytic phages and their endolysins targeting these bacteria have been identified and produced to validate their effectiveness in eradicating these foodborne pathogens. With regards to milk contamination, the *S. aureus* PhiH5 phage endolysin LysH5 reduced staphylococcal cell populations growing in pasteurized milk to undetectable levels after a 4-hour incubation at 37°C (Obeso et al., 2008). The streptococcal endolysins Ply700 (Celia et al., 2008) and B30 (Donovan et al., 2006) have also been shown to display bacteriolytic activity in milk. Listeriosis, caused by *L. monocytogenes*, can be devastating especially in the elderly, and in pregnant women. While various disinfecting chemicals used in food processing facilities kill *L. monocytogenes* effectively, they also kill beneficial bacteria thereby creating an unsustainable sterile environment. A more targeted approach wherein the beneficial bacteria are unaffected would be beneficial as they could prevent the pathogenic bacteria from re-establishing themselves. Due to their remarkable ability to maintain substantial lytic activity even after exposure to 90°C for 30 minutes, the kinetically stable endolysins Ply511, Ply500, and Ply118 targeting *Listeria monocytogenes* may find practical use as disinfectants for food items subjected to high-temperature processing, such as pasteurized milk (Schmelcher et al., 2012). These three endolysins were also successful in reducing *L. monocytogenes* survival on iceberg lettuce by as much as 2.4 log units after being stored for six days. Endolysin CBDs have been bound to paramagnetic beads and shown to effectively detect pathogens in food by fluorescent imaging, polymerase chain reaction (PCR), and surface plasmon resonance (SPR) (Kong et al., 2015; Kretzer et al., 2007; Walcher et al., 2010). Currently,

companies such as Intralytix (ListShield™) and Microcos (PhageGuard) have bacteriophage products targeting *E. coli*, *Listeria*, and *Salmonella*, for use on food equipment and products that are FDA “generally recognized as safe” (GRAS). ListShield™, consists of six naturally occurring bacteriophages that significantly reduces *Listeria monocytogenes* contamination when sprayed on ready-to-eat meat products. Intralytix has also developed EcoShield™, which is effective against *E. coli* in reducing contamination of various foods such as fruits, vegetables, and SalmoFresh™, which targets *Salmonella spp* in red meat and poultry.

Animal health is another area where endolysins could find application. As such, endolysins have been evaluated as disinfectants. An endolysin could be applied as a spray within animal housing facilities or on equipment, serving as a safer substitute for the strong chemicals currently employed to decontaminate surfaces. This would help prevent the transfer of bacteria from surfaces to animals or humans. *Streptococcus equi* is responsible for causing a highly contagious disease called equine strangles in the lymph nodes of horses. When left untreated, horses can suffocate to death due to obstruction of their airways by enlarged lymph nodes. The streptococcal endolysin PlyC was compared to Virkon-S, a disinfecting agent, for its potential to eliminate *S. equi* on horse stable equipment and surfaces. Just 1 µg of PlyC sterilized a 10⁸ CFU/ml culture of *S. equi* in half an hour, with it being 1000 times more active on a per weight basis compared to Virkon-S. PlyC also showed efficacy at sterilizing horse stable equipment contaminated with *S. equi* when applied as an aerosol (Hoopes et al., 2009), demonstrating efficacy on multiple surfaces such as cotton, leather, neoprene, wood, polyester, glass, and stainless steel.

S. aureus is the primary pathogen associated with causing bovine mastitis in cattle. Bovine mastitis is an economically significant disease in which the udder gets inflamed due to bacterial infections. Mastitis can lead to decreased milk production, more discarded milk, increased

veterinary costs and harm to the affected cows. Chimeric endolysins such as λ SA2-E-LysK-SH3b and λ SA2-E-Lyso-SH3b have displayed synergy along with lysostaphin in killing *S. aureus* *in vitro* as well as in a murine model of bovine mastitis (Schmelcher et al., 2012). Moreover, transgenic cows expressing lysostaphin in the mammary glands were protected against mastitis (Wall et al., 2005).

Another application for endolysins is in agriculture. Potatoes that were genetically modified to produce lysozyme demonstrated resistance to black leg and soft rot caused by *Erwinia carotovora* (During, 1993). Transgenic plants can also serve as bioreactors for the large-scale and inexpensive production of endolysins, as evidenced by the *S. pneumoniae* endolysins Cpl-1 and Pal and the *S. agalactiae* endolysin PlyGBS that were produced by the tobacco plant *Nicotiana tabacum* chloroplasts in high amounts (Oey et al., 2009). The USDA lab in Beltsville, MD has produced bacteriophage PlyCP41 endolysin in systemically infected plant tissue, with potential for use of crude plant sap as an effective antimicrobial agent against *Clostridium perfringens*, which is the third leading cause of human food-borne illnesses (Hammond et al., 2019). Recently, endolysins have also been used to improve resistance of citrus plants to Huanglongbing and citrus canker, both diseases that affect citrus plants and are caused by *Candidatus Liberibacter asiaticus* and *Xanthomonas citri* pv. *citri* (Xcc). The LasLYS2 endolysin demonstrated bactericidal activity against Xcc and protected citrus plants from being affected by citrus canker disease (Xu et al., 2023).

Chapter 2: Understanding the Molecular Basis for Homodimer Formation of the Pneumococcal Endolysin Cpl-1

This chapter is published as:

Alreja, A. B., Linden, S. B., Lee, H. R., Chao, K. L., Herzberg, O., & Nelson, D. C. (2023). Understanding the Molecular Basis for Homodimer Formation of the Pneumococcal Endolysin Cpl-1. *ACS infectious diseases*, 9(5), 1092–1104.

Images reproduced with permission.

Author Contributions

A.B.A. and D.C.N. conceived the study and designed the research plan. A.B.A. performed most experiments, analyzed the data, interpreted the results, wrote, reviewed and edited the original draft of the manuscript. S.B.L. assisted with discussions. H.L. assisted with routine experimental prep. K.L.C. performed analytical ultracentrifugation experiments. O.H. reviewed and edited the manuscript. D.C.N. contributed with reviewing and editing.

Data Contributions

A.B.A. obtained all data except for the following:

K.L.C. obtained data for and provided Figs. 2-2 B-D.

O. H. generated Fig. 2-4.

2.1 Abstract

The rise of multi-drug resistant bacteria to traditional antibiotics has prompted the search for alternatives to combat bacterial infections. Endolysins, which are bacteriophage-derived peptidoglycan hydrolases, are an attractive tool in this fight. Several studies have already demonstrated the efficacy of endolysins in targeting bacterial infections. Endolysins encoded by bacteriophage that infect Gram-positive bacteria typically possess an N-terminal catalytic domain (CD) and a C-terminal cell-wall binding domain (CWBD). In this study, we have uncovered the molecular mechanisms that underlie formation of a homodimer of Cpl-1, an endolysin that targets *S. pneumoniae*. Here, we use site-directed mutagenesis, analytical size exclusion chromatography (SEC), and analytical ultracentrifugation to disprove a previous suggestion that three residues at the N-terminus of the CWBD are involved in the formation of a Cpl-1 dimer in the presence of choline in solution. We conclusively show that the C-terminal tail region of Cpl-1 is involved in formation of the dimer. Alanine scanning mutagenesis generated various tail mutant constructs that allowed identification of key residues that mediate Cpl-1 dimer formation. Finally, our results allowed identification of a consensus sequence (FxxEPDGLIT) required for choline-dependent dimer formation, a sequence that occurs frequently in pneumococcal autolysins and endolysins. These findings shed mechanistic light on Cpl-1 and related enzymes and can be used to inform future engineering efforts for their therapeutic development against *S. pneumoniae*.

2.2 Introduction

Streptococcus pneumoniae is included in the global priority list of antibiotic-resistant bacteria identified by the World Health Organization (WHO), which require new drugs to be researched and developed (Tacconelli et al., 2018). Despite the successful global launch of pneumococcal vaccines, the rate of invasive pneumococcal disease is still high in children under five and in adults aged over 50 (CDC, 2022; Durando et al., 2013; WHO) with almost 750,000 deaths in children under the age of five in 2019 alone (WHO, 2019, 2021). Pneumococcal infections can range from simple otitis media or respiratory tract infections to invasive bloodstream infections or meningitis (Shenoy & Orihuela, 2016). Complicating the management of pneumococcal disease is the remarkable variability among *S. pneumoniae* strains, with at least 100 capsular serotypes identified to date, (Ganaie et al., 2020; Steel et al., 2013) whereas the available vaccines target just a fraction of these serotypes. Furthermore, there has been an increase in antibiotic-resistant strains of *S. pneumoniae* (Butler et al., 1996; CDC, 2019; Song et al., 2013). High rates of resistance for several antimicrobials such as penicillin, trimethoprim-sulfamethoxazole, and erythromycin have been reported (Cillóniz et al., 2018; Doern et al., 2005). Thus, the widespread distribution of antibiotic-resistance genes in bacteria compels us to the post-antibiotic era and makes development of alternative antimicrobial agents one of the highest global priorities for aggressive biological, pharmaceutical, and medical investigations (Veiga-Crespo et al., 2007).

The use of endolysins (also termed enzybiotics), which are bacteriophage-encoded peptidoglycan hydrolases, provides an alternative to small molecule antibiotics (Fischetti, 2010; Nelson et al., 2006). Endolysins and holins are produced by the invading bacteriophage to lyse the

host cell and permit the release of new bacteriophage virions. Endolysins can also be applied exogenously and are alone capable of destroying the Gram-positive bacterial peptidoglycan both rapidly and specifically (Fischetti, 2006; Linden et al., 2021; Roach & Donovan, 2015). Endolysins derived from bacteriophage that infect Gram-positive bacteria share a similar modular structure with one or more N-terminal catalytic domains (CDs) and a C-terminal cell-wall binding domain (CWBD) (Ajuebor et al., 2016; Fischetti, 2005, 2010; Loessner, 2005). The CDs carry the catalytic activity of cleaving various bonds of the peptidoglycan. The CWBDs bind to specific substrates, which are usually carbohydrates or teichoic acids attached to the host peptidoglycan, (Fischetti, 2008) thus providing host specificity.

S. pneumoniae has a unique nutritional requirement for choline and displays this moiety on its wall teichoic acids. Not surprisingly, pneumococcal endolysins and autolysins contain choline binding repeats (CBRs) within their CWBD that bind to the teichoic acids via phosphocholine, and as such, lend specificity of these enzymes to pneumococci and other choline displaying organisms. Pal is the endolysin from the pneumococcal Dp-1 bacteriophage and contains an *N*-acetylmuramoyl-L-alanine amidase CD and a CWBD comprising six CBRs followed by a short (17 amino acid) C-terminal tail region (Garcia et al., 1984). Another pneumococcal endolysin, Cpl-1, is derived from the Cp-1 bacteriophage and possesses an *N*-acetylmuramidase (lysozyme) CD and a CWBD containing six CBRs and a 16 amino acid C-terminal tail (Garcia et al., 1987). Various *in vivo* studies investigated the potential therapeutic efficacy of Pal and Cpl-1. For example, Pal was shown to reduce *S. pneumoniae* to undetectable levels in a mouse nasopharyngeal colonization model, (Loeffler et al., 2001) while Cpl-1 was shown to be effective in killing *S. pneumoniae* in a mouse otitis media model, (McCullers et al., 2007) in a mouse

intravenous pneumococcal bacteremia model, (Loeffler et al., 2003) in a rat endocarditis model, and in a mouse pneumonia model (Doehn et al., 2013).

Similar to endolysins, pneumococcal-encoded autolysins also adopt modular structures with a CD and a CWBD that contains CBRs. LytA is one such autolysin that hydrolyzes the pneumococcal cell wall at the end of the stationary growth phase (Howard & Gooder, 1974). The crystal structure of full-length LytA, containing an *N*-acetylmuramoyl-L-alanine “amidase” CD and a CWBD that contains six CBRs, each comprising a β -hairpin, has been published (PDB code: 4X36) (Li et al., 2015). Notably, LytA forms a dimer in solution in the presence of choline (Usobiaga et al., 1996). The structure shows this dimerization occurs through a C-terminal β -hairpin 7 and β -hairpin 6 of the last CBR (Fernandez-Tornero et al., 2002). This dimer structure is known as the “tail-to-tail” association of two monomers. The LytA dimer interface is governed by interactions of hydrophobic and aromatic residues. Various LytA truncations have been produced to investigate the importance of the CBRs and the C-terminal tail (Garcia et al., 1994). Constructs lacking the last ~11 amino acids do not form a dimer in the presence of choline and display a significant decrease in catalytic activity (Garcia et al., 1994; Varea et al., 2000).

It has long been known that the Cpl-1 and Pal endolysins also form homodimers upon binding choline in solution, although the molecular mechanisms involved in the dimerization have remained elusive (Monterroso et al., 2008; Varea et al., 2004). Buey et al. modeled the three-dimensional structure of Cpl-1 in solution in its free as well as in its choline-bound states. Small angle X-ray scattering (SAXS) was used in combination with the high-resolution three-dimensional crystal structure of the monomer (PDB code: 1H09) to model a solution dimer of Cpl-1 (Buey et al., 2007). Based on these results, the authors hypothesized the dimer interface involves three aromatic residues, Trp210, Phe218, and Tyr238, located at the N-terminus of the CWBD.

This “head-to-head” association model of the Cpl-1 dimer contrasts with the C-terminal “tail-to-tail” association observed in the experimental dimer structure of LytA (Fernandez-Tornero et al., 2002). The “head-to-head” interpretation of the Cpl-1 SAXS envelope also disagrees with the recent SAXS-based interpretation of the homologous Pal protein, despite the similar shapes of the two envelopes, (Gallego-Paramo et al., 2021) underscoring the inherently low resolution of the SAXS method. Lastly, Resch et al. hypothesized that the Cpl-1 C-terminus mediates dimer formation and constructed a homodimer by introducing cysteine residues at the C-terminal end. The dimer stabilization by disulfide bond formation increased anti-pneumococcal activity and bioavailability in blood (Resch et al., 2011b). Thus, there is ambiguity on the precise nature and location of Cpl-1 dimer formation upon binding choline. In this manuscript, we dissect the molecular mechanisms governing formation of the Cpl-1 dimer and conclusively identified the exact residue(s) required for dimer formation.

2.3 Results

The N-terminus of the CWBD is not Involved in Cpl-1 Dimerization but is Required for Catalytic Activity

Superose 12 analytical size exclusion column showed that Cpl-1 forms dimers in the presence of choline. Consistent with the published data, (Monterroso et al., 2008) wild-type Cpl-1 forms a dimer in the presence of choline (Fig. 2-1A). A clear shift in the peak is visible when comparing Cpl-1 alone (light blue curve) to Cpl-1 in the presence of 50 mM choline chloride (dark blue curve). Monomeric Cpl-1 (39.8 kDa, inclusive of a 6X His tag) migrated between the 17 and 44 kDa molecular mass standards while the dimer in choline co-migrated with the 158 kDa molecular mass marker (standards shown in grey), which suggests an extended conformation of the dimer, consistent with previous studies. Buey *et al.* proposed that dimer formation of Cpl-1 in the presence of choline occurs via three aromatic residues, Trp210, Phe218 and Tyr238 (Buey et al., 2007). To probe these residues, we prepared three individual single alanine mutants (W210A, F218A, and Y238A), three double mutants (W210A:F218A, W210A:Y238A, and F218A:Y238A), and a triple mutant (W210A:F218A:Y238A, referred to as WFY). Six of these mutants were produced and purified to homogeneity while one of the double mutants (W210A:Y238A) was insoluble. All three single mutants, both soluble double mutants, and the WFY mutant were able to form dimers in the presence of 50 mM choline. The WFY mutant data are shown in Fig. 2-1B. These data refute the Buey *et al.* hypothesis and suggest other residues are likely involved in the formation of Cpl-1 dimers upon binding choline.

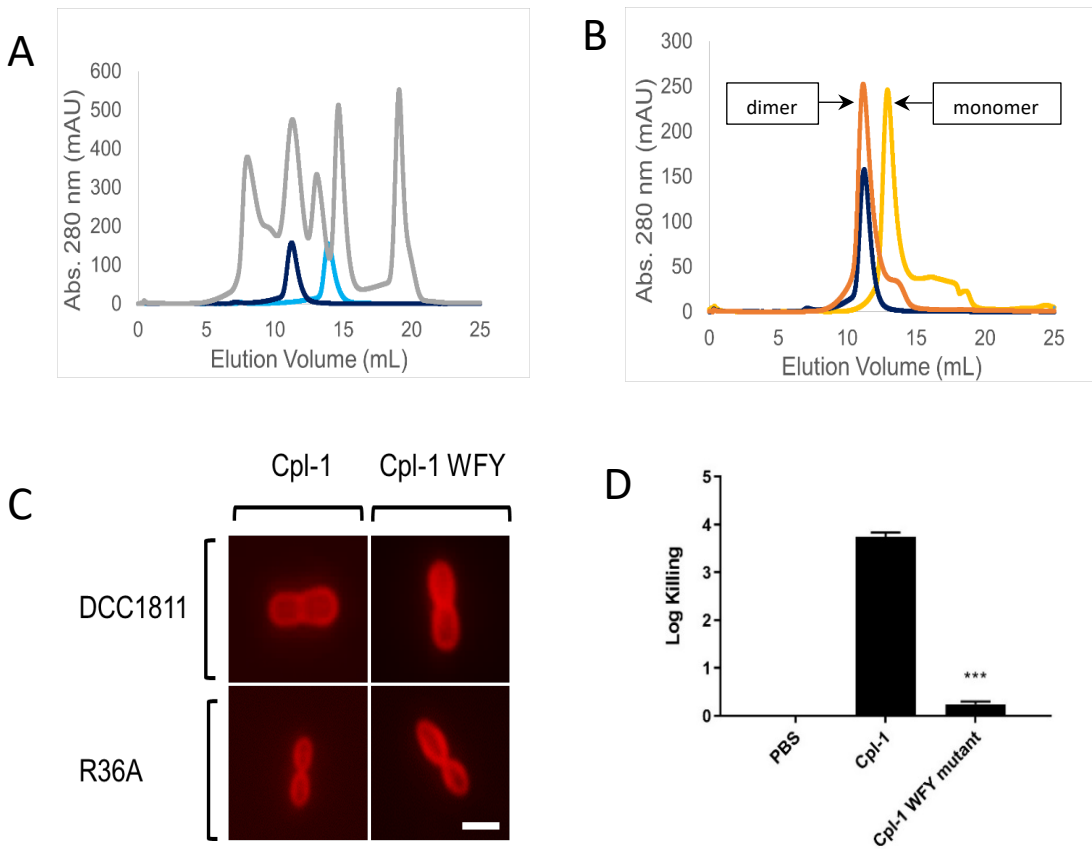


Figure 2-1. Size exclusion chromatography and cellular activity of wild-type Cpl-1 and the N-terminal CBR W210A:F218A:Y238A triple mutant. (A) Wild-type full-length Cpl-1 was applied to a Superose 12 analytical size exclusion column in the presence or absence of 50 mM choline chloride. Cpl-1 formed dimers with 50 mM choline chloride (dark blue curve), but was a monomer in PBS with no choline chloride (light blue curve). Standards (grey curve) are 670, 158, 44, 17, and 1.35 kDa. (B) The Cpl-1 W210A:F218A:Y238A (WFY) triple mutant was applied to a Superose 12 column in the presence and absence of 50 mM choline chloride. The triple mutant formed a dimer in the presence of choline (orange curve) and eluted at the same volume as the choline-containing wild-type Cpl-1 sample (dark blue curve). The yellow curve corresponds to the WFY triple mutant in PBS with no choline chloride. (C) Fluorescence microscopy of *S. pneumoniae* strains DCC1811 (serotype 11) or R36A (capsule-free) when treated with a fluorescent derivative of the Cpl-1 CWBD or the Cpl-1 WFY triple mutant CWBD. Scale bar: 0.5 μ m. (D) Cpl-1 WFY mutant displays loss of bactericidal activity against *S. pneumoniae* DCC1811. Experiments were done in triplicate, and the error bars represent the standard deviations. Asterisks represent level of significance ($P < 0.0005$) between Cpl-1 and Cpl-1 WFY mutant.

To investigate whether the WFY mutant can bind to the bacterial surface, the CWBDs (residues 200-339) of both wild-type Cpl-1 and its WFY mutant were produced and labeled with AlexaFluor-555. Fluorescence microscopy showed both CWBDs bound with equal intensity to the surface of *S. pneumoniae* DCC1811, a serotype 11 strain, and R36A, a capsule-free strain (Fig. 2-1C). Yet, a *S. pneumoniae* DCC1811 lysis assay showed that the Cpl-1 WFY mutant had reduced lytic activity ($P < 0.0005$) with a reported killing of ~ 0.2 logs at 100 $\mu\text{g/ml}$ whereas the wild-type Cpl-1 at the same concentration reduced the number of surviving bacteria by ~ 4 logs (Fig. 2-1D). Taken together, the results show that none of the single, double, or WFY mutants affected the ability of Cpl-1 to form dimers in the presence of choline or to bind the pneumococcal surface. However, there is a noted defect in antimicrobial activity of the WFY mutant. Whether this defect induces structural changes, which in turn affect catalytic activity, is unknown and requires further investigation.

Analytical Ultracentrifugation and Antimicrobial Activity of Cpl-1 16-residue C-terminal Deletion Mutant

Because Cpl-1 dimer formation via the N-terminus of the CWBD (i.e., the head-to-head model) can be ruled out, we speculate that the dimer interface is mediated by the C-terminal tail region of the protein by analogy to the LytA mode of dimerization (i.e., the tail-to-tail model). The crystal structure of Cpl-1 (PDB code: 1H09) shows that the CWBD comprises two regions, CI and CII (Hermoso et al., 2003). The CI region contains the first four CBRs while the CII contains the remaining two CBRs followed by a sixteen amino acid C-terminal tail. Markedly, the C-terminal tails of Cpl-1, LytA, and Pal share high sequence identity (Fig. 2-2A). To test this hypothesis, a Cpl-1 deletion mutant was produced, Δtail , which lacks only the 16 C-terminal residues (residues

324-339). Analytical size exclusion experiments measured formation of dimers in the presence of 50 mM choline. However, despite being a pure protein, this mutant displayed multiple retarded peaks on chromatograms, suggesting possible interactions with the column matrix. The same matrix interaction phenomenon was reported previously for C-terminal truncations of LytA (Varea et al., 2000). Therefore, analytical ultracentrifugation was employed to circumvent the matrix effect.

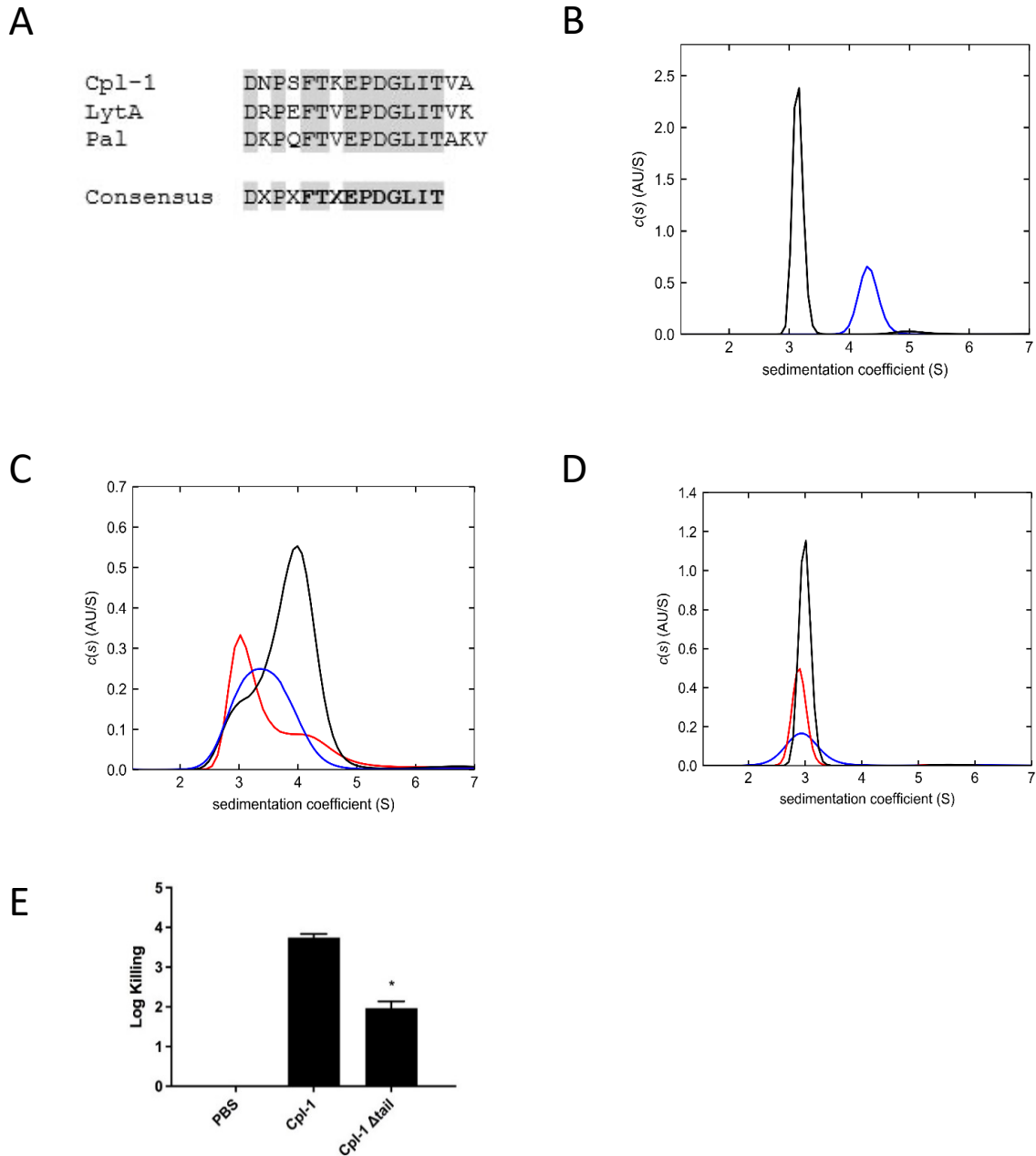


Figure 2-2. Analytical ultracentrifugation of Cpl-1 and Cpl-1 Δ tail. (A) C-terminal tail sequence alignment of Cpl-1, LytA, and Pal shows a consensus sequence (grey boxes). Residues in bold indicate consensus sequence after comparing to 76 novel putative endolysins. (B) The $c(s)$ distributions of 14.3 μ M Cpl-1 in PBS (black) and 7.5 μ M Cpl-1 in PBS plus 50 mM choline (blue) shows dimer formation only in the presence of choline. (C) The $c(s)$ distributions of Cpl-1 Δ tail in PBS at different concentrations: 10.7 μ M (blue), 11.2 μ M (red), and 17.8 μ M (black). (D) The $c(s)$ distributions of Cpl-1 Δ tail at different concentrations in PBS plus 50 mM choline: 4.5 μ M (blue), 5.6 μ M (red), and 9 μ M (black). (E) The Cpl-1 Δ tail mutant displays reduced lytic activity compared to wild-type Cpl-1 on *S. pneumoniae* DCC1811. Data is represented as log-fold killing

compared to untreated PBS control. Experiments were done in triplicate, and the error bars represent the standard deviations. Asterisks indicate level of significance ($P < 0.05$) between Cpl-1 and Cpl-1 Δ tail data.

Sedimentation velocity (SV) experiments were conducted in PBS or PBS supplemented with 50 mM choline chloride. In PBS alone, the Cpl-1 wild-type protein had an experimental weight-average sedimentation coefficient of 3.1S with a MW_{app} of 40 kDa, consistent with its monomeric mass (Fig. 2-2B). In PBS with 50 mM choline chloride, wild-type Cpl-1 sedimented as a single species of 4.4S, with a MW_{app} of ~76 kDa, suggesting formation of a stable dimer. By contrast, the SV analyses of Cpl-1 Δ tail protein in the absence of choline chloride displayed broad sedimentation boundaries with shifting peaks between monomeric 3.1S and dimeric 4.3S forms at all three protein concentrations examined (Fig. 2-2C), indicative of a reversible monomer-dimer equilibrium. In the presence of 50 mM choline chloride, the Cpl-1 Δ tail protein exists only as monomeric 3.0S species (Fig. 2-2D) with a MW_{app} of 33 kDa at each of three tested concentrations. Although these results prove that the C-terminal 16 residues of Cpl-1 are required for dimer formation in the presence of choline, it remains unclear why the Cpl-1 Δ tail protein forms transient interactions in the absence of choline.

Log-fold killing assays were also performed with the Cpl-1 Δ tail mutant to determine the effect of the deletion on antimicrobial activity. The Cpl-1 Δ tail mutant had significantly lower activity compared with the wild-type Cpl-1 ($P < 0.05$). In a one-hour log-fold killing experiment, the Δ tail mutant caused a ~2 log decrease in CFU of *S. pneumoniae* strain DCC1811, whereas wild-type Cpl-1 killed the bacteria more efficiently and reduced the CFU's by ~4 logs (Fig. 2-2E).

Alanine Scanning Mutagenesis of the Cpl-1 Tail

To pinpoint the exact residue(s) involved in the formation of the Cpl-1 dimer, five triple alanine point mutants were made along the 16 amino acid C-terminal Cpl-1 sequence (DNP·SFT·KEP·DGL·ITV·A). All but the ITV construct, which was insoluble, were purified and analyzed on a Superose 12 column in PBS or PBS supplemented with 50 mM choline chloride. The DNP→AAA and KEP→AAA mutants retained the ability of the wild-type protein to form dimers in the presence of choline (Fig. 2-3A). In contrast, the SFT→AAA and DGL→AAA mutants did not form dimers (Fig. 2-3B). The chromatograms for both the SFT mutant (red curve) and the DGL mutant (gray curve) in the presence of choline superimpose with the peak for wild-type Cpl-1 (light blue) in the absence of choline, clearly demonstrating that these mutants do not form dimers. Taken together, the data obtained for the Cpl-1 Δ tail mutant and the triple alanine mutants establish the critical role that the 16 C-terminal residues play in choline-bound Cpl-1 dimer formation, supporting the tail-to-tail association model. Furthermore, both the SFT and DGL triple alanine mutants displayed reduced lytic activity compared to wild-type Cpl-1, suggesting formation of the dimer is required for full lytic activity (Fig. 2-3C).

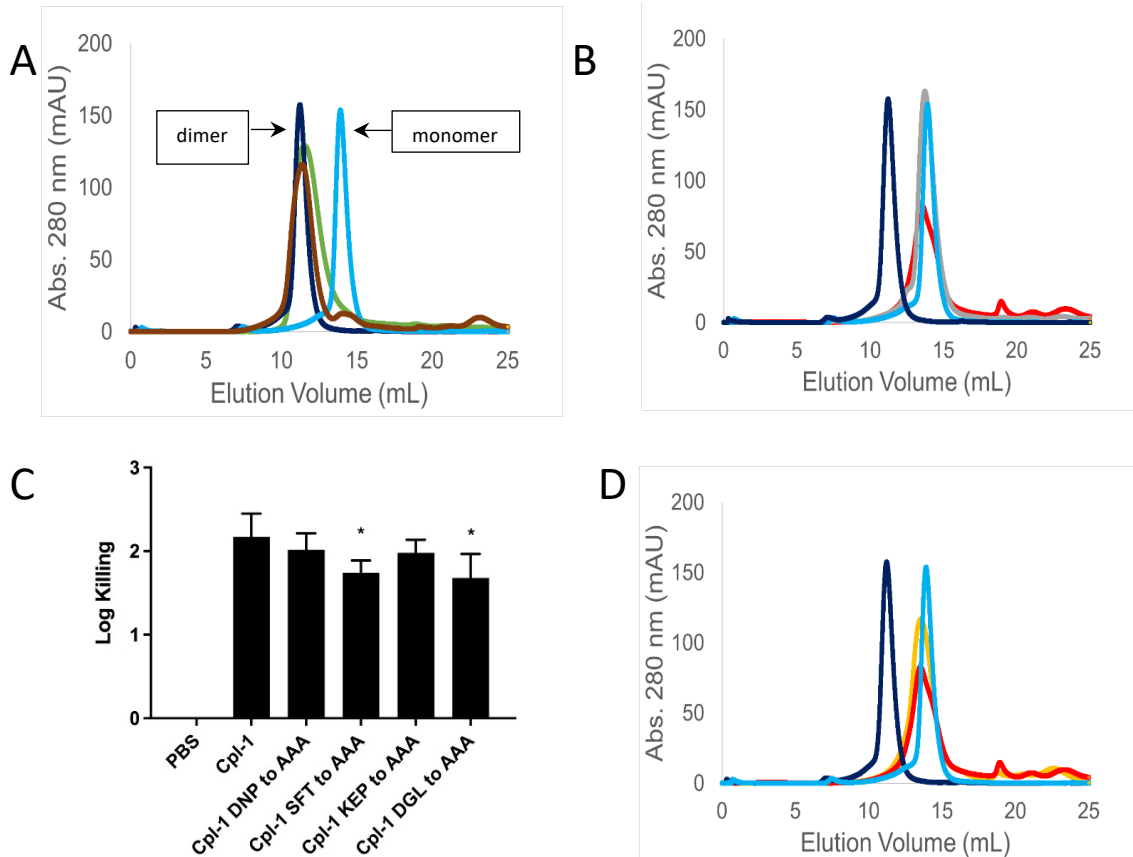


Figure 2-3. Size exclusion chromatography of Cpl-1 tail alanine scanning mutagenesis. (A) The Cpl-1 DNP→AAA mutant was analyzed by a Superose 12 column in the presence of 50 mM choline chloride. This mutant formed dimers in the presence of choline (green curve). The KEP→AAA mutant also formed dimers in the presence of 50 mM choline chloride (brown curve). Both mutants eluted at the same volume as wild-type Cpl-1 in choline (dark blue curve). The light blue curve represents wild-type Cpl-1 in PBS. (B) The Cpl-1 SFT→AAA mutant was analyzed by a Superose 12 column in the presence of 50 mM choline chloride and did not form dimers (red curve). Likewise, the DGL→AAA mutant also did not form dimers in the presence of 50 mM choline chloride (gray curve). Both these mutants eluted at the same volume as wild-type Cpl-1 in PBS (light blue curve). Wild-type Cpl-1 in PBS with 50 mM choline chloride is shown as the dark blue curve. (C) All four of the alanine scanning mutants display lytic activity as measured by a log-fold killing assay. The SFT→AAA mutant had significantly reduced lytic activity compared to wild-type Cpl-1 as did the DGL→AAA mutant. The DNP→AAA and KEP→AAA mutants displayed similar lytic activity to wild-type Cpl-1. Asterisks represent level of significance (*, $P < 0.05$) between mutant proteins and wild-type Cpl-1. (D) The Cpl-1 F328A mutant was analyzed on the Superose 12 column and did not form a dimer in the presence of 50 mM choline chloride (gold curve), similar to the SFT→AAA mutant (red curve). Wild-type Cpl-1 with (dark blue) and without (light blue) choline chloride are shown for comparison purposes.

The full-length crystal structure of LytA (PDB code: 4X36), containing the tail-to-tail association, reveals an overall “V” shaped dimer formed along a 2-fold crystallographic axis (Fig. 2-4A,C). Lacking a dimer structure for Cpl-1, the artificial intelligence (AI) program AlphaFold2 (Jumper et al., 2021) was used to model a Cpl-1 dimer. The prediction was performed with ColabFold-multimer (Mirdita et al., 2022) using MMseq2 (Steinegger & Soding, 2017) with the paired sequence default option and no template. The best ranked model is shown (Fig. 2-4B,D), which has a pLDDT=90.8, ptmscore=0.41, and iptm=0.389. Although the iptm value is low, the same interface for Cpl-1 was generated by AlphaFold2 as that of LytA. All five AlphaFold2 Cpl-1 models exhibit the same tail-to-tail dimer interface and differ mainly by the conformation of the CD-CWBD linker, yielding different orientations between the two domains. Nevertheless, all five models possess an overall “M” shape in contrast to the “V” shaped LytA. Notably, specific to the C-terminal tail regions, π -stacking is observed between the side chains of Phe307 of each LytA monomer (Fig. 2-4E). The corresponding Cpl-1 residue, Phe328, in the AlphaFold2 model exhibits the same π -stacking (Fig. 2-4F). The Cpl-1 SFT→AAA triple mutant containing Phe328 failed to form dimers in the presence of choline (Fig. 2-3B). To confirm that Phe328 plays a critical role in dimerization, a single site Cpl-1 mutant, F328A, was made. This mutant likewise was unable to form dimers in the presence of choline as monitored by analytical size exclusion chromatography (Fig. 2-3D). Moreover, the F328A mutant retained bacteriolytic activity against pneumococci, albeit diminished compared to wild-type Cpl-1 (Fig. 2-5). Nonetheless, this further supports our triple alanine scanning mutagenesis results indicating that Cpl-1 monomers are active, but formation of a dimer is necessary for full bacteriolytic activity. While other residues in the C-terminal β -hairpin, specifically those in the DGL→AAA mutation, may also contribute to the

dimeric state stability, the Phe328 mutant proves that the Cpl-1 dimerization is mediated by the C-terminus of the CWBD rather than the N-terminus. Thus, the sequence homology between LytA and Cpl-1 is also manifested in the common tail-to-tail mode of dimerization of the two proteins.

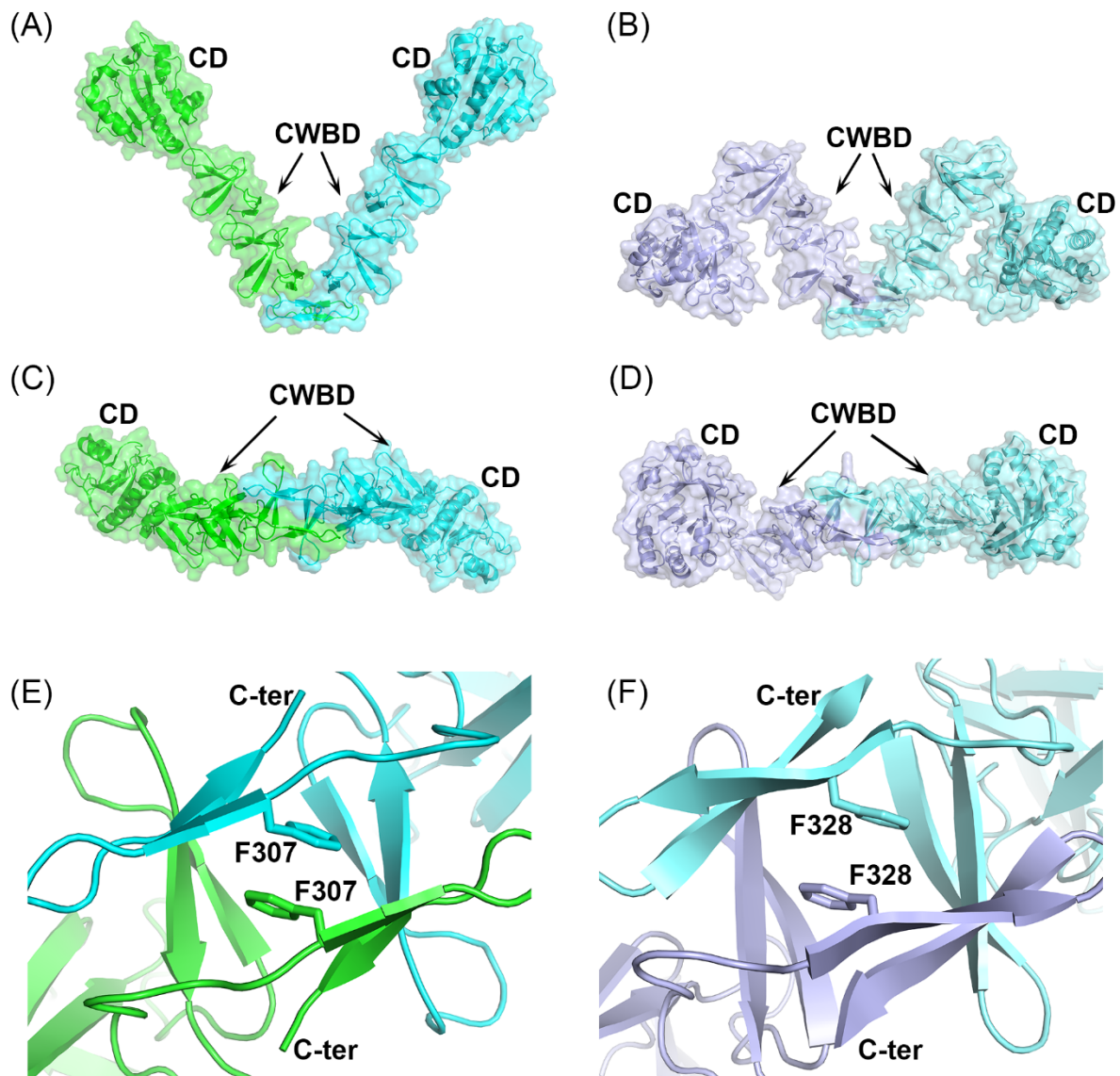


Figure 2-4. Comparison between the dimers of LytA crystal structure (PDB code: 4X36) and the model of Cpl-1 generated using the AI program AlphaFold2. LytA monomers are colored green and cyan (Figures A, C, E), and those of Cpl-1 are colored lavender and light blue (Figures

B, D, F). Domain labels are abbreviated as CD (the N-terminal catalytic domain) and CWBD (the C-terminal choline binding domain). A and B show the dimers along the 2-fold symmetry axis in cartoon and molecular surface representations. C and D show the dimer perpendicular to the 2-fold symmetry axis with the same representation as in A and B. E and F show a close-up of the dimer interface perpendicular to the 2-fold symmetry axis, highlighting the interacting Phe307 pair in LytA and the equivalent Cpl-1 Phe328 pair that was mutated in this study.

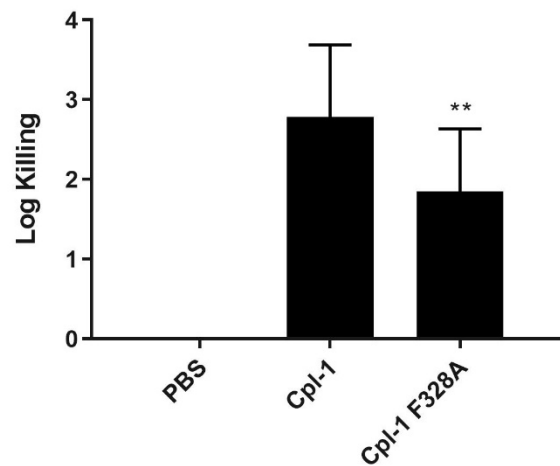


Figure 2-5. Bactericidal activity of Cpl-1 F328A against *S. pneumoniae* DCC1811. The single point F328A mutant displays significantly reduced lytic activity compared to wild-type Cpl-1. Experiments were done in triplicate, and the error bars represent the standard deviations. Asterisks represent level of significance ($P < 0.005$) between mutant proteins and wild-type Cpl-1.

Cpl-1 Dimerization Depends on Choline Concentration

The dependence of Cpl-1 dimerization on choline concentration was analyzed by analytical size exclusion chromatography. As seen in Fig. 2-6A, Cpl-1 remains a monomer in the presence of 1 mM choline (green curve). At 10 mM choline, Cpl-1 exists predominantly as a dimer (yellow curve). However, at 5 mM choline, Cpl-1 exists in an equilibrium state between a monomer and dimer (purple curve). These results suggest an apparent K_D value of ~5 mM, which is consistent with the previously reported value of 3.6 mM (Monterroso et al., 2008).

Choline chloride is considered and used as a proxy for the pneumococcal teichoic acid, however, choline is actually incorporated into the pneumococcal lipoteichoic and wall teichoic acids as phosphocholine (Fischer, 2000). Therefore, we measured the binding affinity for Cpl-1 in phosphocholine to compare to choline and were unable to detect a difference between the two. As with choline, Cpl-1 is a monomer in the presence of 1 mM phosphocholine (green curve), a dimer in 10 mM phosphocholine (yellow curve), and exists in equilibrium between a monomer and dimer in the presence of 5 mM phosphocholine (purple curve) (Fig. 2-6B).

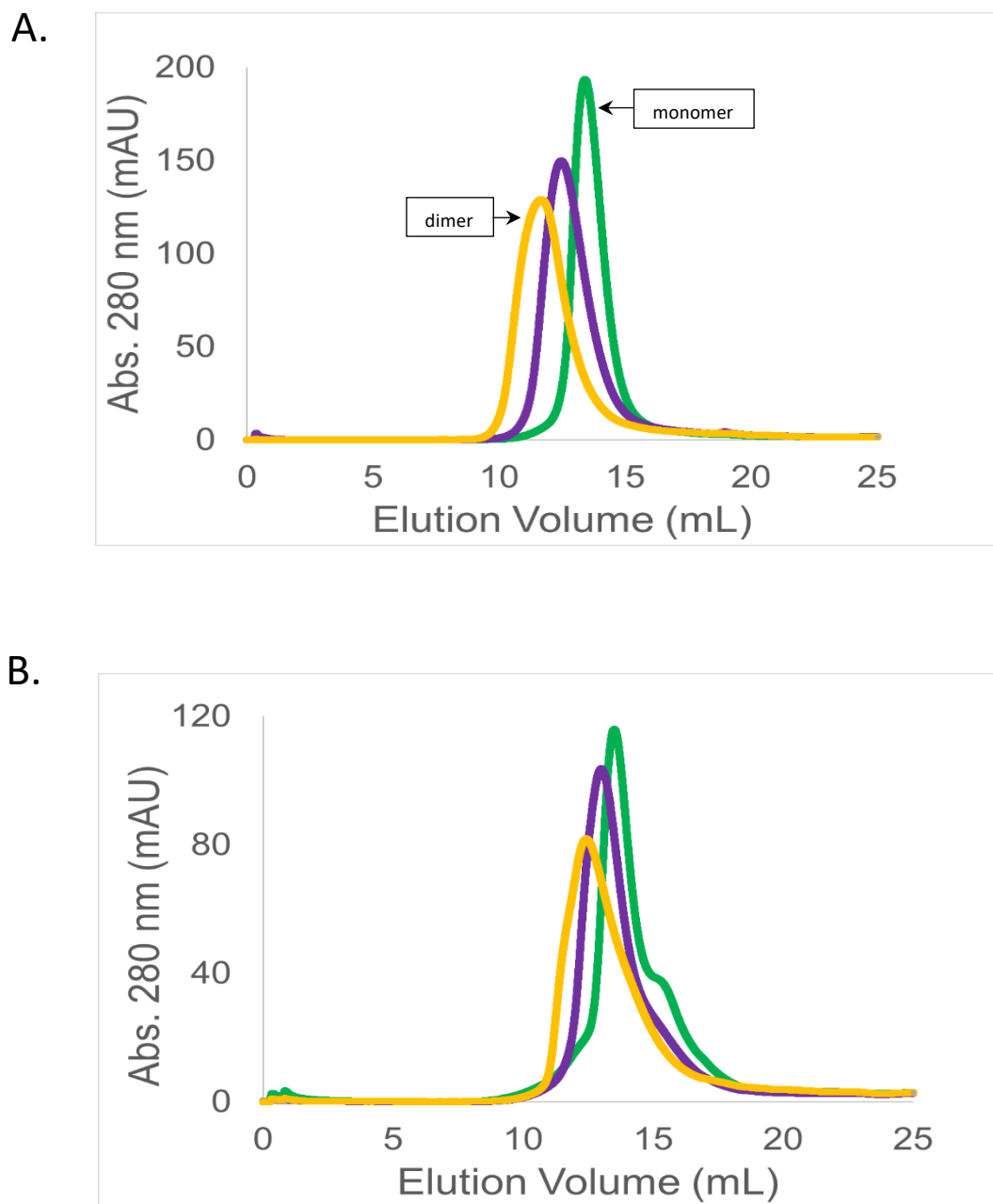


Figure 2-6. Choline concentration dependent formation of Cpl-1 dimers. (A) Cpl-1 (1.5 mg/ml) was analyzed by Superose 12 analytical size exclusion chromatography in the presence of 1 mM (green curve), 5 mM (purple curve), or 10 mM (yellow curve) choline chloride. (B) Similar parameters as in A, but Cpl-1 (1.3 mg/ml) in phosphocholine calcium chloride tetrahydrate was used in place of choline chloride at 1 mM (green curve), 5 mM (purple curve), and 10 mM (yellow curve) concentrations.

Conserved C-terminal β -hairpin Sequence Identifies Pneumococcal Endolysins that Likely Function as Dimers

The 16 C-terminal amino acid sequences of the pneumococcal CWBDs, which we refer to as the “tail”, are conserved between LytA, Cpl-1, and Pal (Fig. 2-2A) and all three enzymes form dimers in the presence of choline (Monterroso et al., 2008; Usobiaga et al., 1996; Varea et al., 2004). LytA and Pal have previously been shown to employ a tail-to-tail interaction between two monomers upon choline binding (Fernandez-Tornero et al., 2002; Gallego-Paramo et al., 2021) and here we confirm that Cpl-1 dimerization is also mediated by the tail. The C-terminal tail of these three enzymes contain a consensus sequence (DxPxFTxEPDGLIT), which can be used to identify putative pneumococcal cell wall hydrolases (i.e., autolysins and endolysins) and to predict their propensity to form dimers in the presence of choline. Recently, Fernandez-Ruiz et al. collated the sequences of 200,000 uncultured viruses and predicted that this collection contains 2,628 novel putative endolysins (Fernandez-Ruiz et al., 2018). We further screened these endolysins for the presence of the preliminary pneumococcal CWBD tail consensus sequence, DxPxFTxEPDGLIT, and identified 76 matches that had the exact consensus sequence of FxxEPDGLIT (Fig. 2-2A, bold sequences), all located 2 to 13 residues away from the C-termini (Appendix A). Sequence analysis using HMMSCAN, which is part of the HMMER suite, (Finn et al., 2011; Potter et al., 2018) and clustering based on Pfam families (Finn et al., 2014; Mistry et al., 2021) revealed that 73 of the 76 proteins contain typical autolysin and endolysin CDs including the Amidase Family 2 and Amidase Family 5 domains, and to a lesser extent, the cysteine, histidine-dependent amidohydrolases/peptidase (CHAP) and glycoside hydrolase family 25 (GH_25) catalytic domains (Fig. 2-7A). The remaining three protein sequences do not contain CDs or other secondary domains, but rather are stand-alone CWBDs.

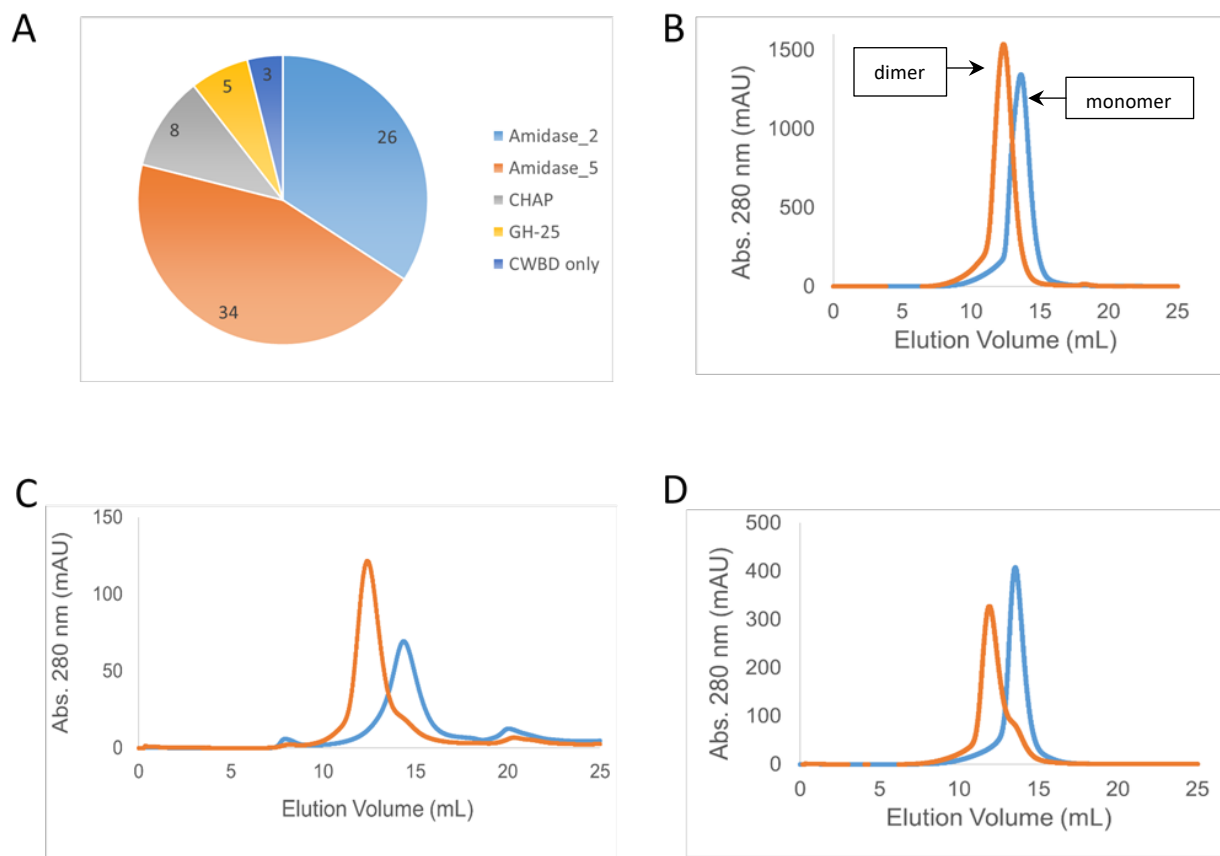


Figure 2-7. Novel endolysins identified by the consensus sequence form a dimer in the presence of choline. (A) Classes of CD domains of the 76 identified putative endolysins with the C-terminal β -hairpin consensus sequence. (B) Dimerization status of SP-CHAP on the Superose 12 column. The protein forms a dimer in the presence of 50 mM choline (orange curve). (C) SP-Amidase forms a dimer in the presence of 50 mM choline (orange curve). (D) SP-GH25 forms a dimer in the presence of 50 mM choline (orange curve). For Figures B to D, the blue curves represent the respective proteins analyzed in PBS without choline chloride.

To support the sequence analysis results, we randomly selected three representative CDs from the 76 putative endolysins bearing the conserved sequence motif (SRS numbers correspond to sequence IDs listed in Appendix A): SRS019125, containing a CHAP CD and renamed SP-CHAP here; SRS017713, containing an Amidase-2 CD and renamed SP-Amidase here; and SRS018591, containing a GH_25 CD and renamed SP-GH25 here. All three putative endolysins were cloned, expressed, purified, and evaluated for dimer formation in the presence of 50 mM choline chloride by analytical size exclusion chromatography. As seen in Fig. 2-7B, SP-CHAP is a monomer in the absence of choline chloride (blue curve) and forms a dimer in the presence of 50 mM choline chloride (orange curve). Similarly, in Fig. 2-7C and 2-7D, both SP-Amidase and SP-GH25, respectively, form dimers in the presence of 50 mM choline chloride. Taken together, the data suggests that most autolysins and endolysins that target the pneumococcal surface have evolved a widespread and conserved mechanism whereby these enzymes form dimers through their C-terminal tails upon binding a choline entity on the bacterial surface.

2.4 Discussion

It has been well established that LytA, Cpl-1, and Pal form dimers upon binding choline in solution, and they presumably also dimerize on the pneumococcal surface that displays choline-containing wall teichoic acids (Monterroso et al., 2008; Usobiaga et al., 1996; Varea et al., 2004). The LytA tail-to-tail mechanism of choline-dependent dimer formation is substantiated by several crystal structures. However, the head-to-head model proposed for Cpl-1 was based on fitting the crystal structure of a choline-free monomeric protein into a SAXS envelope generated in the presence of choline, and the subsequent identification of aromatic residues that potentially formed a non-canonical choline-binding site (Buey et al., 2007). However, the authors did not confirm that the three aromatic residues on the non-canonical N-terminal β -hairpin are required for dimer formation, nor did they pursue a tail-to-tail interaction hypothesis. Additionally, a crystal structure is a static image of one particular crystal packing arrangement, which may not represent a protein in solution, and the pitfalls of fitting the structure of multi-domain, flexible proteins into a rough SAXS envelope are well documented (Kikhney & Svergun, 2015). In contrast, we provide here biophysical evidence, supported by mutagenesis studies, that both disprove the head-to-head model and support the tail-to-tail dimer model of Cpl-1 in the presence of choline.

We have shown that Cpl-1 binds choline with millimolar affinity (Fig. 2-6A), consistent with earlier reports based on circular dichroism studies (Monterroso et al., 2008). We also show that Cpl-1 exhibits similar affinity towards phosphocholine, which more closely resembles the native choline in the pneumococcal wall teichoic acids (Fig. 2-6B). However, endolysin CWBDs are generally thought to bind the bacterial surface with much higher affinity than millimolar. *Bacillus* endolysins bind cell wall polysaccharides with micromolar affinity, (Ganguly et al., 2013; Mo et al., 2012) whereas *Listeria* endolysins bind cell wall polysaccharides with nano- to

picomolar affinities (Loessner et al., 2002; Schmelcher et al., 2010). The discrepancy in binding affinity may be attributable to several factors. First, given that Cpl-1 is predicted to possess five choline binding sites, an avidity effect may increase the apparent low affinity for the pneumococcal surface (Roig-Molina et al., 2020). A similar hypothesis has been developed for PlyCB, the CWBD of the PlyC endolysin that contains eight binding sites (Broendum et al., 2021). A second, and even more likely hypothesis, is that other components of the teichoic acid or moieties on the peptidoglycan itself may be necessary for high affinity binding, as suggested recently for other pneumococcal choline-binding proteins (Perez-Dorado et al., 2010).

There is growing evidence that endolysins and other cell wall hydrolases form multimeric structures more often than previously thought. PlyC, an endolysin from the streptococcal C1 phage, is comprised of two independently expressed subunits, PlyCA, which is the catalytic domain (i.e. CD), and PlyCB, the cell-wall binding domain (CWBD). In solution, PlyC multimerizes into eight PlyCB subunits associated with one PlyCA subunit (Nelson et al., 2006). Clostridial endolysins have also been shown to form multimers in solution. Recently, LysB, an endolysin from the prophage of *Clostridium botulinum* E3 strain Alaska E43 was shown to form a homodimer in solution (Morzywolek et al., 2021). Endolysin CTP1L that targets *Clostridium tyrobutyricum*, endolysin CS74L that targets *Clostridium sporogenes*, and endolysin CD27L that targets *Clostridium difficile*, also form multimers. However, these enzymes exist as heterotetramers composed of two full-length proteins (i.e., analogs of CD and CWBD) and two individual CWBDs generated by in-frame, internal translation start sites (iTSS) at the beginning of the CWBD coding sequence (Dunne et al., 2016). Endolysins obtained from phages that infect *Enterococcus* have also been shown to form multimers. Endolysin Lys170 that displays lytic activity on *Enterococcus faecalis* was shown to form a heterotetramer in solution, wherein a full-

length protein interacts with three individually expressed CWBDs, which are also generated by an iTSS (Proenca et al., 2015). Similarly, the enterococcal endolysin LysIME-EF1 forms a holoenzyme with a 1:3 stoichiometric ratio of full-length protein to CWBD (Zhou et al., 2020). Gp2 (LysA) from mycobacteriophage Ms6 lacks an identifiable CWBD, but a truncated version of its CD forms a heterodimer with the full-length protein (Catalao et al., 2011). Likewise, a truncated version of the staphylococcal phage 2638A endolysin ORF007 containing one of two CDs and a CWBD forms a heterodimer with the full-length protein (Abaev et al., 2013). Lastly, Pinto et al. recently performed a bioinformatic search of iTSS sites in putative endolysins and experimentally confirmed at least four new multimeric endolysins, including LysP7951, which adopts a 1:5 CD:CWBD stoichiometry (Pinto et al., 2022). All of the above enzymes fold into their multimeric state at the level of protein translation. Only the pneumococcal enzymes (i.e., Cpl-1, Pal, and LytA) are produced and purified as monomers and require binding to a ligand mimetic of epitopes on the bacterial surface to induce multimerization.

Given the modular nature of endolysin CD and CWBD domains, (Garcia et al., 1988) bioengineering and chimeragenesis studies have begun to dominate the endolysin field in recent years. Engineering is beneficial in that the natural bacteriolytic activity of endolysins can be improved by mixing and matching different domains. Domain addition, deletion, mutagenesis, and peptide fusion are just some of the various approaches being currently employed. A recent review by Sao-Jose (Sao-Jose, 2018) summarizes the generation of several chimeric endolysins along with their benefits compared to the parental enzyme. To this end, knowledge of the dimerization mechanism and the Cpl-1 consensus sequence can have important implications when considering engineering approaches focused on pneumococcal CWBDs. As an example, ClyJ is a chimera containing a CHAP CD derived from PlyCA and a CWBD derived from gp20, which shares a high

degree of similarity with that of LytA and is comprised of several CBRs (Yang et al., 2019). ClyJ was further engineered by modifying the linker sequence between the CD and CWBD to create ClyJ-3, which possesses higher antimicrobial activity than ClyJ, presumably due to differences in linker flexibility (Yang et al., 2020). As expected, both ClyJ and ClyJ-3 form dimers in the presence of choline. ClyJ-3 was subsequently engineered to truncate the C-terminal tail, yielding ClyJ-3m, which lacks the ability to form a dimer in the presence of choline (Luo et al., 2020). Surprisingly, this monomeric enzyme has higher lytic activity on *S. pneumoniae* *in vitro* and *in vivo* than either ClyJ-3 and Cpl-1. These data contrast with LytA deletion studies where activity was dramatically reduced with C-terminal deletions, (Garcia et al., 1994; Varea et al., 2000) as well as our own studies deleting the tail region of Cpl-1 (Fig. 2-2E). Also in contrast to the monomeric ClyJ-3m results, Resch et al. (Resch et al., 2011b) constructed a Cpl-1 pre-formed dimer by introducing cysteine residues at the start of the C-terminal β -hairpin of the protein. This pre-formed dimer displayed an increased bactericidal activity against *S. pneumoniae* and also an increased bioavailability by decreasing renal clearance when tested in an *in vivo* mouse model.

We have proven that Cpl-1 forms a homodimer through interactions with residues in the C-terminal tail in the presence of choline, and we highlighted the consensus motif in this region that is widespread and can be used to predict formation of a dimer in pneumococcal cell wall hydrolases, including both pneumococcal autolysins and bacteriophage endolysins. Nonetheless, the molecular mechanisms that potentiate a signal from the choline binding sites to the tail regions that induces dimer formation is yet to be determined. This knowledge will inform future engineering efforts.

2.5 Conclusion

S. pneumoniae infections are increasingly associated with antimicrobial resistance. Despite the availability of vaccines, the sheer capsular diversity among *S. pneumoniae* strains enable them to evade treatment. Endolysins have shown promise in treating Gram-positive bacterial infections. In the present study, we identify the dimerization mechanism for the pneumococcal endolysin Cpl-1 in the presence of choline. We disprove a previous hypothesis that suggested Cpl-1 dimer formation occurs at the N-terminus of its binding domain and subsequently show that its C-terminal tail is involved in formation of the dimer. Using AlphaFold2 modeling combined with experimental mutagenesis studies, we establish residue Phe328 as crucial to formation of the Cpl-1 dimer in the presence of choline. In addition, we identify a consensus sequence required for choline-dependent dimer formation, one that is frequently observed in pneumococcal endolysins and autolysins. Our findings illustrate the molecular mechanisms involved in dimer formation which can help guide future engineering approaches for endolysins targeting *S. pneumoniae*.

2.6 Materials and Methods

Bacterial Strains and Growth Conditions

Bacterial strains used in this work were routinely grown at 37°C unless otherwise noted. *S. pneumoniae* DCC1811, a serotype-11 capsulated strain (Sa-Leao et al., 2000), or *S. pneumoniae* R36A, a capsule-free strain (Avery et al., 1944), was grown with 5% CO₂ in Todd Hewitt Broth supplemented with 1% yeast extract (THY) without shaking, or on THY plates. *Escherichia coli* DH5 α and BL21 (DE3) strains were grown in Lysogeny Broth (LB) or on LB plates supplemented with ampicillin or kanamycin.

Cloning and Site-directed Mutagenesis

The *Cpl-I* used in this study was codon-optimized and synthesized (GeneArt). The primers used in this study are listed in Appendix B. A 6X-His tag was added to the C-terminal end of each protein. The Q5[®] Site-Directed Mutagenesis Kit from New England Biolabs was used to generate all point mutants and the manufacturer instructions were followed for all steps. Resulting mutants were confirmed by sequencing (Psomagen) before being transformed into *E. coli* BL21 (DE3) for protein expression. The pBAD24 and pET28a plasmids were used for protein expression. The ApE program (University of Utah) was used for DNA sequence analysis and manipulations.

Protein Expression and Purification

Overnight cultures of *E. coli* BL21 (DE3) containing the plasmid of interest were sub-cultured 1:100 into 1.5 L LB containing antibiotic for plasmid maintenance and grown at 37°C in baffled flasks to an OD₆₀₀ = 1. The LB was supplemented with either 100 μ g/ml ampicillin for

constructs using the pBAD24 vector or 50 µg/ml kanamycin for constructs using the pET28a vector. Expression was induced while cells were in mid-log phase with 0.25% arabinose overnight at 18°C for constructs in pBAD24 or 10 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) for those in the pET28a vector. Cells were harvested (5,000 $\times$ g, 15 min, 4°C), resuspended in 20 ml of lysis buffer (PBS, 10 mM imidazole, pH 7.4) and disrupted by freeze-thawing (-80°C to 4°C). This was followed by sonication on ice (power level 6 and duty cycle 30%). Insoluble cell debris was centrifuged (12,000 $\times$ g, 1 hour, 4°C) and the cell lysate containing the 6X-His tagged protein was loaded onto Ni-NTA resin (HisPur™ Ni-NTA Resin, Thermo Scientific) in a gravity flow column and eluted in 4 ml fractions using an imidazole step gradient in PBS buffer (20 mM, 50 mM, 100 mM, 250 mM, and 500 mM imidazole). Protein purity was analyzed on SDS-PAGE gel. Where further purification was necessary, samples were applied to a Sephacryl S-200 size exclusion column (Cytiva). Protein samples were pooled and concentrated using a protein concentrator with the appropriate molecular weight cut-off (Amicon® Ultra 15 ml centrifugal filters, Millipore Sigma).

Microbiology and Lytic Assays

Antimicrobial activity was quantified by means of colony forming unit (CFU) counting as previously described (Nelson et al., 2012). Briefly, sterile protein at a final concentration of 100 µg/ml was mixed with 100 µl of stationary-phase *S. pneumoniae* in PBS in a 96 well plate and incubated at 37°C for an hour. Then, five 10-fold dilutions in PBS were made and 10 µl from each dilution was plated on THY agar plates, air-dried and placed in a 37°C incubator overnight. The remaining CFUs were counted and the data was reported as the log-fold killing compared to the untreated control. Each CFU counting assay was performed in triplicate on three separate days.

Analytical Size Exclusion

Cpl-1 and its mutants were subjected to analytical size exclusion chromatography on a Superose 12 column (GE Healthcare) as previously described (Nelson et al., 2006). The column was calibrated with size exclusion standards (Bio-Rad) and equilibrated with either PBS or PBS supplemented with choline chloride at indicated concentrations for two column volumes (CV) before loading protein sample. Five-hundred μ l of a 1.5 mg/ml protein sample was loaded onto the column for each run and the sample was eluted for 1.5 CV. Alternatively, phosphocholine chloride calcium salt tetrahydrate was used in 10 mM Tris (pH 7.4) with 150 mM NaCl because it was insoluble in PBS.

Analytical Ultracentrifugation

Sedimentation velocity (SV) experiments were performed using a ProteomeLab Beckman XL-A with absorbance optical system and a 4-hole An60-Ti rotor (Beckman Coulter). For SV, 380 μ L protein in PBS, pH 7.4, and 400 μ L buffer were loaded into the sample and reference sectors of the dual-sector charcoal-filled epon centerpieces. The samples in PBS $\pm$ 50 mM choline chloride were equilibrated overnight at 20° C. The samples were centrifuged at 50 krpm and the absorbance data for 4.5-17.8 μ M proteins were collected at 280 nm. Absorbance signal was monitored in a continuous mode with a step size of 0.003 cm and a single reading per step. Sedimentation coefficients were analyzed from SV profiles using the program SEDFIT (Brown & Schuck, 2006). The continuous $c(s)$ distributions were calculated assuming a direct sedimentation boundary model with maximum entry regularization at a confidence level of 1 standard deviation.

Fluorescence Labeling of Proteins

Purified proteins were reacted with the carboxylic acid, succinimidyl ester of AlexaFluor[®]-555 (Molecular Probes) according to the manufacturer's instructions. Unreacted dye was removed from labeled protein by application to a 5 ml HiTrap desalting column (GE Healthcare) equilibrated with PBS.

Cell Wall Binding Assays and Fluorescence Microscopy

An overnight culture of *S. pneumoniae* DCC1811 or *S. pneumoniae* R36A was centrifuged and mixed with the AlexaFluor[®]-555 labeled Cpl-1 CWBD or the WFY Cpl-1 triple mutant CWBD and incubated at 37°C for 10 min. The cells were centrifuged, washed in PBS, and then visualized by fluorescence microscopy with an exposure time of 200 msec. An Eclipse 80i epifluorescent microscope workstation (Nikon) with X-Cite 120 illuminator (EXFO) and Infinity 3 camera (Teledyne Lumenera) camera was used. Infinity Analyze 7 software was used for image acquisition and analysis.

Bioinformatics and Discovery of Novel Enzymes

Recently, Fernandez-Ruiz et. al. conducted a bioinformatics pipeline to identify novel endolysins from nearly 200,000 uncultured viruses in which they discovered 2,628 putative endolysins (Fernandez-Ruiz et al., 2018). All putative endolysins were scanned and 76 were identified that contained the consensus sequence (FxxEPDGLIT), with almost all occurring at the C-terminus (Appendix A). Three proteins were picked at random for further study based on the type of catalytic domain they possessed and their percent identity to Cpl-1, Pal, and LytA. These three enzymes were synthesized (GeneArt), subcloned into pBAD24 vectors, expressed, and

purified. They were then tested for formation of dimers in the presence of choline by analytical size exclusion chromatography as described above.

Statistical Analysis

Experimental data are presented as means $\pm$ standard error of means. Student's *t* tests were performed for the statistical analysis. *P* values of <0.05 were considered to be statistically significant.

Chapter 3: SP-CHAP, an Endolysin with Enhanced Activity Against Biofilm Pneumococci and Nasopharyngeal Colonization

This chapter is submitted, in part, for publication as:

Alreja, A. B., Appel, A. E., Zhu, J. C., Riley, S. P., Gonzalez-Juarbe, N., & Nelson, D. C. (2023). SP-CHAP, an endolysin with enhanced activity against biofilm pneumococci and nasopharyngeal colonization.

Author Contributions

A.B.A. and D.C.N. conceived the study and designed the research plan. A.B.A. performed most experiments, analyzed the data, interpreted the results, wrote, reviewed and edited the original draft of the manuscript. A.E.A. and N.G.J. assisted with performing experiments. J.C.Z. assisted with performing confocal microscopy experiments. S.P.R. reviewed the manuscript. D.C.N. and N.G.J. edited and reviewed the manuscript.

Data Contributions

A.B.A. obtained all data except for the following:

A.E.A. obtained data for and provided Figs. 3-1 D, E.

N.G.J. obtained data for and provided Figs. 3-1 H and 3-2 E.

J.C.Z. provided Fig. 3-2 C.

3.1 Abstract

Streptococcus pneumoniae, a Gram-positive bacterium, is responsible for causing a wide variety of invasive infections. The emergence of multi-drug antibiotic resistance has prompted the search for antimicrobial alternatives. Phage-derived peptidoglycan hydrolases, known as endolysins, are an attractive alternative. In this study, an endolysin active against *S. pneumoniae*, designated SP-CHAP, was cloned, produced, purified, biochemically characterized and evaluated for its antimicrobial properties. Cysteine, histidine-dependent amidohydrolase/peptidase (CHAP) domains are widely represented in bacteriophage endolysins, but have never previously been reported for pneumococcal endolysins. Here we report the first pneumococcal endolysin with a CHAP catalytic domain. SP-CHAP was antimicrobial against all *S. pneumoniae* serovars tested, including capsular and capsule-free pneumococci, and it was found to be more active than the most widely studied pneumococcal endolysin, Cpl-1, while not affecting various oral/nasal commensal organisms tested. SP-CHAP was also effective in eradicating *S. pneumoniae* biofilms at concentrations as low as 1.56 µg/ml. In addition, a *S. pneumoniae* mouse nasopharyngeal colonization model was employed that showed SP-CHAP caused a significant reduction in *S. pneumoniae* CFUs, even more than Cpl-1. These results indicate that SP-CHAP may represent a promising alternative to combat *S. pneumoniae* infections.

3.2 Importance

In light of the high rates of pneumococcal resistance reported for several antibiotics, alternatives are urgently needed. In the present study, we report an endolysin targeting *S. pneumoniae* that is more active than Cpl-1, the most characterized pneumococcal endolysin to date. We used a combination of biochemical and microbiological assays to assess the stability and lytic potential of SP-CHAP and demonstrate efficacy on pneumococcal biofilms *in vitro* and in an *in vivo* mouse model of colonization. Our findings highlight the therapeutic potential of SP-CHAP as an antibiotic alternative to treat *S. pneumoniae* infections.

Keywords: endolysin, CHAP domain, bacteriophage, pneumococcus, antibiotic resistance

3.3 Observation

Endolysins offer an alternative to conventional antibiotics in addressing the emergence of multi-drug-resistant bacteria. These enzymes, encoded by bacteriophages, induce cell lysis upon synthesis at the end of the bacteriophage replication cycle (Fischetti, 2010; Nelson et al., 2006). Typically comprised of two domains, endolysins possess an N-terminal catalytic domain and a C-terminal cell binding domain (Fischetti, 2008). The catalytic domain is responsible for hydrolyzing various bonds within the bacterial peptidoglycan, while the cell binding domain is responsible for binding to a specific substrate within the cell wall, typically a carbohydrate (Fischetti, 2005, 2010). Binding to a unique cell wall ligand confers specificity to endolysins at the species, strain, or even serotype level, enabling exogenous application of endolysins to induce bacterial cell lysis in these strains. Notably, endolysins are not susceptible to efflux pumps, penicillin-binding proteins, or other mechanisms of resistance often observed with traditional antibiotics (Danis-Wlodarczyk et al., 2021; Linden et al., 2021). Moreover, their efficacy extends to eradicating bacterial biofilms. In recent years, the landscape of endolysin research has shifted toward engineering, domain shuffling, and chimeragenesis studies, strategies that optimize bacteriolytic activity and can prolong the efficacy of these enzymes should resistance eventually emerge.

Pneumococcal disease is an area where endolysins could have a significant impact. *S. pneumoniae*, a Gram-positive bacterium, resides within the human nasopharynx as part of the normal flora, yet it can cause a range of infections, including community-acquired pneumonia, invasive pneumococcal disease, otitis media, meningitis, and sepsis (Austrian, 1981; Jain et al., 2015; Shenoy & Orihuela, 2016). As the approval of novel antibiotics by the U.S. Food and Drug Administration has dwindled and antibiotic-resistant strains of *S. pneumoniae* have surged,

addressing this pathogen has become imperative (CDC, 2019; Song et al., 2013). While pneumococcal vaccines based on a set of capsular polysaccharides exist, the over 100 identified serotypes of this bacterium allow its persistence (Ganaie et al., 2020). Several potent endolysins that target *S. pneumoniae* have been examined and have demonstrated efficacy both *in vitro* and *in vivo*. Noteworthy examples include Cpl-1 (Loeffler et al., 2003) and Pal (Loeffler et al., 2001), endolysins with bacteriolytic efficacy either individually or synergistically against *S. pneumoniae* (Jado et al., 2003). An array of models, including mouse nasopharyngeal colonization models (Loeffler et al., 2001), otitis media models (McCullers et al., 2007), pneumococcal bacteremia models (Loeffler et al., 2003), rat endocarditis models (Entenza et al., 2005) and rat pneumococcal meningitis models (Grandgirard et al., 2008), have shown these enzymes are excellent bacteriolytic agents *in vivo*. Additionally, LytA, a pneumococcal autolysin, exhibits lytic activity by degrading the pneumococcal cell wall at the culmination of the stationary bacterial growth phase.

Among endolysins, one of the more potent catalytic domains is known as the cysteine, histidine-dependent amidohydrolase/peptidase (CHAP) domain (Broendum et al., 2018). CHAP domains are widespread among endolysins derived from bacteriophages that target Gram-positive hosts. Notably, both N-Rephasin® (SAL200) and Exebacase (CF-301), endolysins with antimicrobial activity against *Staphylococcus aureus* that have undergone evaluation in Phase 2a and Phase 3 clinical trials, respectively, utilize a CHAP catalytic domain. However, no naturally occurring pneumococcal endolysin with a CHAP domain has been identified to date. Recently, we identified 76 putative pneumococcal endolysins from uncultured bacteriophage genomes by searching for a consensus sequence found in the cell binding domain region of Cpl-1, Pal, and LytA, which has been determined to be important for dimer formation of the endolysin in the presence of choline (Alreja et al., 2023). Among these endolysins, one named SP-CHAP due to

the presence of a CHAP domain, was confirmed to form a dimer in the presence of choline, suggesting increased binding to the pneumococcal cell wall. Here, we characterize the biochemical and antimicrobial properties of SP-CHAP and benchmark it against Cpl-1, which is currently the most studied pneumococcal endolysin.

3.4 Results

Biochemical Characterization of SP-CHAP

A turbidity reduction assay was employed to characterize lytic activity of SP-CHAP at various pH conditions. The lytic activity of SP-CHAP was tested over a pH range from 3.0 to 10.0 in order to determine the optimum pH for SP-CHAP enzymatic activity on *S. pneumoniae*. The optimum pH was determined to be 7.0 while enzymatic activity was also seen at pH 5.0 and 6.0 (Fig. 3-1A). At pH extremes however, SP-CHAP activity was dramatically reduced. This neutral pH profile range for SP-CHAP is similar to that seen for Cpl-1, Pal and LytA (Domenech et al., 2011). Subsequent turbidity reduction assays to further characterize SP-CHAP were conducted at pH 7.0 based on these pH optimum results.

Next, the thermodynamic stability of SP-CHAP was investigated by incubating the protein at a range of temperatures from 4°C to 55°C for 30 minutes and then placing on ice for 5 minutes. SP-CHAP displayed lytic activity on *S. pneumoniae* at all the temperatures tested from 4°C to 37°C while the highest activity was recorded at 25°C (Fig. 3-1B). There was a marked defect in activity at temperatures 40, 45 and 55°C, suggesting a melting temperature (T_m) between 37°C and 40°C. This T_m is consistent with melting temperatures reported for pneumococcal endolysins Cpl-1 of 43.5°C (Sanz et al., 1993), 37°C for Pal (Varea et al., 2004) and 43.9°C for Cpl-7 (Bustamante et al., 2012). It has been previously shown that the antibacterial activity of several endolysins can be augmented by the addition of NaCl (Becker et al., 2008; Binte et al., 2020; Garcia et al., 2010). In order to test whether SP-CHAP displayed similar salt-enhanced activity, we investigated the effect of addition of NaCl in the range of 0 to 500 mM concentration. SP-CHAP was used at a starting concentration of 25 µg/ml. As seen in Figure 3-1C, there was a dose-dependent decrease in enzymatic activity with an increase in NaCl concentration. At the 500 mM NaCl concentration,

SP-CHAP displayed significantly reduced activity while the activity was only slightly attenuated at lower NaCl concentrations. SP-CHAP did however have the highest activity when no salt was added. These results indicate that contrary to previous reports for other endolysins, NaCl actually has an inhibitory effect on SP-CHAP lytic activity, even at low concentrations.

SP-CHAP has Greater Lytic Activity than Cpl-1 and can be Partially Inhibited by Pneumococcal Capsule

To benchmark SP-CHAP against Cpl-1, we evaluated their bacteriolytic activity against several pneumococcal serovars as well as against several oral/nasal commensal organisms. While both SP-CHAP and Cpl-1 kill all pneumococcal strains tested, at the same concentration SP-CHAP appeared to be more effective against three of the five strains (Fig. 3-1D). Neither of the two endolysins had any killing activity on the oral/nasal commensal organisms tested (Fig. 3-1E). To further quantify the lytic potential of SP-CHAP, a colony forming unit (CFU) assay was conducted. A dose-dependent effect of SP-CHAP lytic activity on *S. pneumoniae* Lyt4.4 was observed at all concentrations tested from 1.56 µg/ml to 100 µg/ml with one hour of treatment (Fig. 3-1F). SP-CHAP caused a reduction of 7 logs of *S. pneumoniae* Lyt4.4 at 100 µg/ml while Cpl-1 caused only a 2-log reduction at the same concentration. At 6.25 µg/ml, SP-CHAP and Cpl-1 displayed identical activity while both enzymes appear to have no lytic activity at 3.12 µg/ml and below. On a minimum inhibitory concentration (MIC) assay, Cpl-1 displayed higher antimicrobial activity with a lower MIC than SP-CHAP against several *S. pneumoniae* strains tested with the difference being more prominent on the non-autolytic and capsule free Lyt4.4 strain (Fig 3-1G). While Cpl-1 displays a lower MIC than SP-CHAP, we attribute this difference to an inherent difference in the assay methodology. All other lytic assays conducted involved incubating the enzyme with bacteria for one hour whereas the MIC assay involved incubating enzymes with

bacteria overnight. The MIC assay is a bacteriostatic assay and suppresses the growth of bacteria keeping them in the stationary phase whereas the other assays conducted being bactericidal, kill the bacteria instantly. From a therapeutic standpoint, endolysins are being developed as bactericidal agents as opposed to bacteriostatic ones, therefore their instant lytic activity is of greater significance and more clinically relevant. In fact, this differing degree of activity from method to method has been previously observed by Kusuma et al., who reported varying degrees of lysostaphin susceptibility for *S. aureus* strains when comparing the MIC assay to the disk diffusion assay (Kusuma & Kokai-Kun, 2005). Nevertheless, we do note that the MIC values for both SP-CHAP and Cpl-1 are within the range that has been previously reported (Corsini et al., 2018). It should be noted that due to mass differences between endolysins and antibiotics (e.g. penicillin is ~0.34 kDa and SP-CHAP is 33.5 kDa), the molarity of antibiotics is significantly higher than endolysins at comparable MIC values. The MIC results also highlighted variations in SP-CHAP activity against the capsulated D39 strain and its capsule-free derivatives, R6 and Lyt4.4. Consequently, we hypothesized that the presence of capsular polysaccharide may modulate SP-CHAP activity. This hypothesis stems from previous findings demonstrating that capsule shedding is a known pneumococcal response to antimicrobial peptides (Kietzman et al., 2016) and that the presence of a capsule in *Klebsiella pneumoniae* is linked to reduced endolysin activity (Hong et al., 2022). Pre-incubation of the purified capsular polysaccharide with the endolysins, 30 min prior to and throughout the duration of a lytic assay, resulted in a partially yet significant reduction in the lytic activity of SP-CHAP, but this effect was not observed with Cpl-1 (Fig. 3-1H).

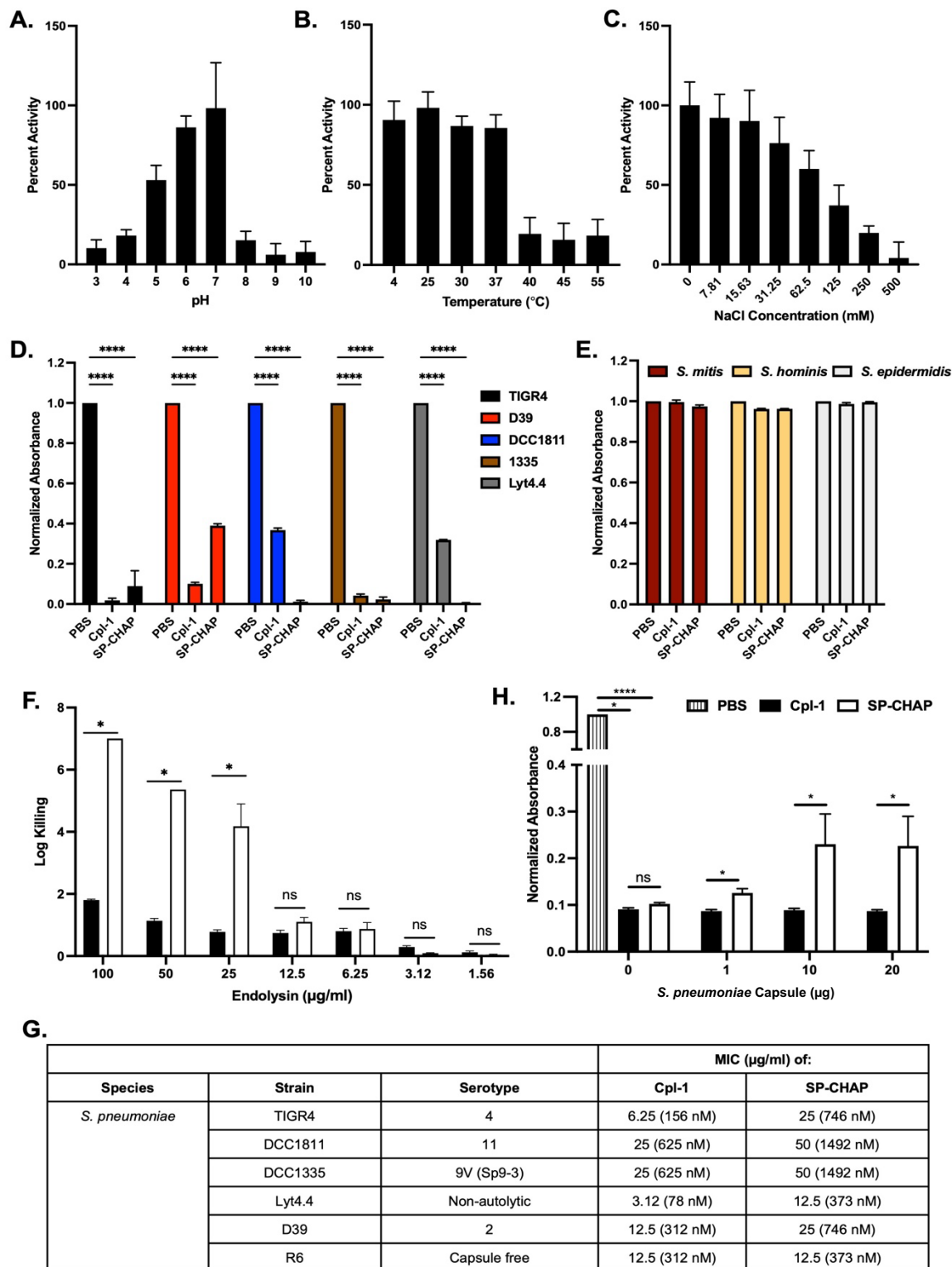


Figure 3-1: Endolysin SP-CHAP is highly active against planktonic-grown pneumococci and its activity is modulated by capsule polysaccharide availability. Biochemical characterization of optimal conditions for SP-CHAP activity. The role of pH (A), temperature (B), and salt (C) on

SP-CHAP lytic activity against *S. pneumoniae* DCC1811. Lytic activity of Cpl-1 and SP-CHAP (50 µg/ml) on **(D)** *S. pneumoniae* strains TIGR4, D39, DCC1811, 1335, and Lyt4.4 or on **(E)** commensal bacteria *S. mitis*, *S. hominis* and *S. epidermidis* measured by absorbance OD₆₀₀. **(F)** Lytic activity of SP-CHAP and Cpl-1 at various concentrations on *S. pneumoniae* Lyt4.4 as measured by the CFU assay. Data is reported as the log-fold killing compared to untreated control. **(G)** MIC (µg/ml of endolysin) of SP-CHAP and Cpl-1 on *S. pneumoniae* strains TIGR4, DCC1811, DCC1335, Lyt4.4, D39, and R6. **(H)** Lytic activity of Cpl-1 and SP-CHAP on *S. pneumoniae* strain Lyt4.4 after incubation with purified pneumococcal capsular polysaccharide. Experiments were done in triplicate, and the error bars represent the standard deviations. For one-way ANOVA and two-way ANOVA, asterisks denote the level of significance observed: ns = not significant; * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$; **** = $p \leq 0.0001$.

SP-CHAP has Greater Activity than Cpl-1 Against Biofilm Pneumococci *in vitro*

S. pneumoniae can form biofilms to promote their colonization of the nasopharynx and to improve their resistance to antibiotics (Blanchette-Cain et al., 2013). Using a static *in vitro* biofilm model, we observed that SP-CHAP was highly efficient at eradicating pneumococcal biofilms when compared to Cpl-1 (Fig. 3-2A). Furthermore, when 24-hr mature biofilms were treated with SP-CHAP for 1-hr, it caused a dose-dependent decrease in the amount of biofilm formed. SP-CHAP eradicated biofilms at all concentrations tested. Cpl-1 only significantly dispersed biofilms only at concentrations at or above 12.5 µg/ml (Fig. 3-2B). Supporting these observations, confocal imaging of biofilms demonstrated that biofilms treated with endolysins exhibited thinner and damaged structures compared to the control group. Further, SP-CHAP caused a greater reduction in biofilm thickness than Cpl-1 (Fig. 3-2C). Taken together, this data shows that SP-CHAP is the only known *S. pneumoniae* endolysin to have high efficacy against *S. pneumoniae* biofilms as both Cpl-1 and Pal (Domenech et al., 2011) require higher doses to achieve decolonization.

SP-CHAP Treatment Does Not Cause Host Toxicity and Decolonizes the Nasopharynx *in vivo*

Using a model of toxicity, we evaluated the systemic responses to SP-CHAP following injection into mice. Using a total injection of 600 µg SP-CHAP, or 10X the dose used in the decolonization studies below, we observed that none of the test group animals exhibited any biological reactivity at any of the specified time points during the three-day observation period. All animals from the test group gained weight at the end of the test period. No abnormalities were noted at gross necropsy (Fig. 3-2D). Finally, to assess the effectiveness of SP-CHAP as a decolonizing agent *in vivo*, we used a mouse *S. pneumoniae* nasopharyngeal colonization model.

We observed that administration of 60 µg SP-CHAP per mouse significantly reduced *S. pneumoniae* CFUs in the nasal tissue, and this was slightly more pronounced than the effect of Cpl-1 (Fig. 3-2E). Of note, previous studies have demonstrated intranasal delivery of 100 µg of penicillin was inadequate to decrease nasopharyngeal colonization by *S. pneumoniae* (Marks et al., 2012), highlighting the advantages of endolysins over antibiotics.

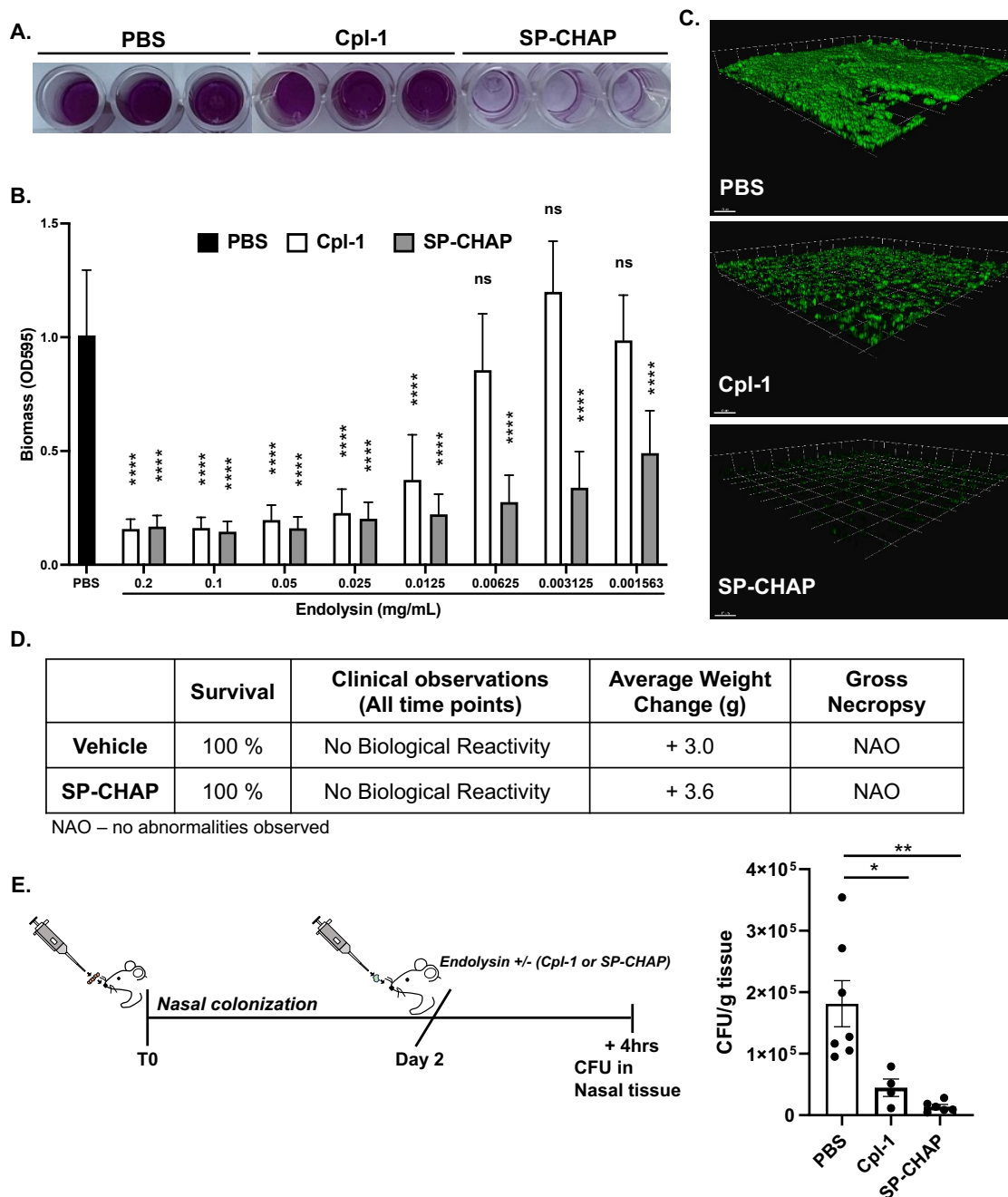


Figure 3-2: SP-CHAP reduces pneumococcal biofilm biomass *in vitro* and nasopharyngeal colonization *in vivo*. (A) Disruption of *S. pneumoniae* Lyt4.4 biofilms visualized by crystal violet staining. (B) Quantification of biomass measured by OD₅₉₅ of crystal violet stained biofilms. SP-CHAP displayed biofilm eradication ability at all concentrations tested while Cpl-1 removed biofilms at concentrations as low as 12.5 μ g/ml. Statistical analysis denotes differences between PBS and endolysin-treated wells. (C) 3D confocal microscopy of *S. pneumoniae* biofilms treated with SP-CHAP and Cpl-1. Scale bar = 10 μ m. (D) Acute systemic toxicity in mice challenged with vehicle or SP-CHAP intraperitoneally. (E) Sketch of mouse colonization model and CFU per gram

of nasopharyngeal tissue from mice infected with *S. pneumoniae* strain TIGR4 and treated with 60 µg of endolysin 24 hr post-colonization. Samples tested with two-way ANOVA or one-way ANOVA Kruskal-Wallis test with Dunn's multiple-comparison post-test. Asterisks denote the level of significance observed: ns = not significant; * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$, **** = $p \leq 0.0001$.

3.5 Conclusion

In summary, our results indicate that SP-CHAP is a non-host toxic, highly efficient endolysin with the ability to lyse *S. pneumoniae* and decolonize biofilm-grown bacteria both *in vitro* and *in vivo*. The fact that capsule can reduce the activity of SP-CHAP during planktonic growth also establishes a novel observation as a possible mechanism for pneumococci to protect themselves against this class of antimicrobial. Taken together, our data show new evidence as to the therapeutic capabilities of SP-CHAP and the mechanism that *S. pneumoniae* may use to counter it when in planktonic form.

3.6 Materials and Methods:

Bacterial Strains and Growth Conditions

S. pneumoniae was grown at 37°C with 5% CO₂ in Todd Hewitt Broth supplemented with 1% yeast extract (THY) without shaking, or on THY plates. *Escherichia coli* DH5 α and BL21 (DE3) strains were grown in Luria Broth (LB) or on LB plates. All strains were stored at -80°C.

Cloning

Cpl-1 (GenBank NP_044837.1) and SP-CHAP (GenBank OR644363) used in this study were codon-optimized and synthesized (GeneArt). A 6X-His tag was added to the C-terminal end. The constructs were sub-cloned into the pBAD24 expression vector and the sequence was confirmed (Psmagen, Rockville, MD, USA) before being transformed into *E. coli* BL21 (DE3) for protein expression. The ApE program (University of Utah) was used for DNA sequence analysis and manipulations.

Protein Expression and Purification

Overnight cultures of *E. coli* BL21 (DE3) containing the plasmid of interest were sub-cultured 1:100 into 1.5 L baffled flasks containing LB supplemented with 100 μ g/ml antibiotic and grown at 37°C. Expression was induced while cells were in mid-log phase with 0.25% arabinose and incubated overnight at 18°C. Cells were harvested (5,000 $\times$ g, 15 min, 4°C), resuspended in 20 ml of lysis buffer (PBS, 10 mM imidazole, pH 7.4), and disrupted by freeze-thawing (-80°C to 4°C). This was followed by sonication on ice (power level 6 and duty cycle 30%). Insoluble cell debris was centrifuged (12,000 $\times$ g, 1 hr, 4°C) and the cell lysate containing the 6X-His tagged protein was loaded onto Ni-NTA resin (HisPur™ Ni-NTA Resin, Thermo

Scientific) in a gravity flow column and eluted in 4 ml fractions using an imidazole step gradient in PBS buffer (20 mM, 50 mM, 100 mM, 250 mM, and 500 mM imidazole). Protein purity was analyzed on an SDS-PAGE gel. The Sephacryl S-200 column (Cytiva) was used to further purify samples. Samples were pooled and concentrated using a protein concentrator with the appropriate molecular weight cut-off (Amicon® Ultra 15 ml centrifugal filters, Millipore Sigma).

Turbidity Reduction Assay

Bacteriolytic (lytic) activity was measured by means of a turbidity reduction assay as previously described (Nelson et al., 2012). Overnight cultures of *S. pneumoniae* were centrifuged and the pellets were washed twice and resuspended in PBS to a final optical density at 600 nm (OD₆₀₀) of 0.9-1.2. Cells were mixed with SP-CHAP or Cpl-1 with a final protein concentration of 100 µg/ml. The OD₆₀₀ was monitored by use of a microplate spectrophotometer (SpectraMax ABS; Molecular Devices, USA) every 15 s for 30 min at 37°C. Lytic activity was quantified as the reduction in turbidity and measured as the difference in OD₆₀₀ between PBS control treated cells and cells that were treated with SP-CHAP or Cpl-1 over 30 min.

Colony Forming Unit Assay

Colony forming unit (CFU) counting was also conducted to quantify antimicrobial activity. Sterile protein at a final concentration of 100 µg/ml was used for this assay. Briefly, 100 µL of protein was mixed with 100 µL of *S. pneumoniae* Lyt4.4 in PBS in a 96-well plate and incubated at 37°C for an hr. Then, five 10-fold dilutions in PBS were made and 10 µL from each dilution was plated on THY agar plates, air-dried and placed in a 37°C incubator overnight. The CFUs were counted and the data was reported as the log-fold killing compared to the untreated control.

Due to autolytic activity pneumococci normally display in overnight cultures, the mutant Lyt 4.4 strain was chosen because it lacks the major autolysin LytA and does not display autolysis overnight. Each turbidity reduction assay and CFU counting assay was performed in triplicate on three separate days.

Minimum Inhibitory Concentration Assay (MIC)

A 10 ml overnight culture of each *S. pneumoniae* strain was grown for 12-16 hours. Subsequently, 100 μ L of the overnight *S. pneumoniae* culture was diluted in 10 ml of THY media (1:100), to which 50 μ M catalase was added. Ten two-fold dilutions of endolysins were prepared, starting at 100 μ g/ml. Following this, 100 μ L of the diluted culture was mixed with 100 μ L of each endolysin treatment at every concentration in triplicate, using a 96-well plate. The plate was then incubated overnight in a 37°C non-shaking incubator and observed the following day to identify the lowest concentration at which there was a clear well with no bacterial growth. The OD₆₀₀ was measured for accuracy using a SpectraMax ABS plate reader.

Biofilm Assay

Biofilms were produced in Costar 96-well flat-bottom polystyrene (PST) plates (Corning). *S. pneumoniae* Lyt4.4 were grown in THY media to an OD₆₀₀ of 0.5 to 0.6, sedimented by centrifugation, and resuspended in an equal volume of fresh THY media. The cells were then diluted 1:2 into a fresh 15 ml tube and 200 μ L of this was dispensed into each well of the 96-well plate. The plate was then incubated at 37°C overnight. The next day, the media was carefully pipetted out from each well and the biofilm formed was gently washed twice with PBS to remove unattached cells. Two-fold serial dilutions of SP-CHAP or Cpl-1 were added to treatment wells

and the plate was incubated at 37°C for 1 hr. The liquid was then pipetted out and the wells were washed with distilled water and left to air-dry. The formed biofilms were stained with 50 µL of a 0.2% crystal violet solution and incubated at room temperature for 15 min. After pipetting out excess crystal violet, the wells were washed three times with distilled water and air-dried. The biofilm was solubilized in 10% SDS to extract crystal violet from the biomass, transferred to a fresh 96-well plate, and then quantified by determining the OD₅₉₅ using a SpectraMax M5 plate reader.

Alternately, biofilms were grown in chamber slides (Nunc Lab-Tek chamber slides, Thermo Fisher Scientific) following previously established procedures (Shenoy et al., 2017). To each chamber, 100 µL of a 100 µg/ml solution of each endolysin was added and allowed to incubate at 37°C for 1 hr. Subsequently, the chambers wells were washed twice with PBS, and the biofilms were stained using the Live/Dead BacLight stain (Thermo Fisher Scientific) before being examined by confocal microscopy.

Confocal Microscopy

Mounted samples were visualized using Zeiss LSM800 confocal microscope at 63x with 1x digital zoom. Z-stacks of each biofilm samples were acquired using the 488 nm laser and 500-619 nm detection filter with a constant 0.5µm Z-axis interval for 19 slices. 3D images of the biofilms were rendered using the normal shading mode of Imaris 9.0.1 software (Oxford Instruments). All acquisition and image settings were set the same across all samples.

Acute Systemic Toxicity Testing

Acute systemic toxicity testing was done by Pacific BioLabs. Briefly, SP-CHAP was prepared according to ISO 10993-12 and Pacific BioLabs internal SOPs (Pacific BioLabs, Hercules, CA). Saline without SP-CHAP was used as the vehicle (negative) control. Ten Swiss-Webster albino mice were randomly assigned to SP-CHAP or vehicle groups. Each animal was injected intraperitoneally with 200 μ L of saline or SP-CHAP at a concentration of 3 mg/ml. The animals were observed for biological reactivity immediately after and at 4-hr $\pm$ 15-min, 24 $\pm$ 2-hr, 48 $\pm$ 2-hr, and 72 $\pm$ 2-hr following injection. Animals were weighed at 24 $\pm$ 2 hr, 48 $\pm$ 2-hr, and 72 $\pm$ 2-hr following injection. At the end of the study, animals were euthanized, and gross necropsy was performed.

Mouse Colonization Studies

The mouse colonization experiments were similar to those previously described (Blanchette-Cain et al., 2013; Riegler et al., 2019). Briefly, mice were infected with 1×10^5 CFU *S. pneumoniae* TIGR4 in each nostril. Forty-eight hrs after inoculation, mice received PBS (vehicle), Cpl-1, or SP-CHAP intranasally (25 μ L per nostril). After four hrs, mice were sacrificed, and their nasopharyngeal tissue was excised and weighed. Homogenates were serially diluted and plated on blood agar plates to determine CFU per gram tissue. All endolysins used for mouse experiments were sterile and passed through an endotoxin removal column prior to administration. Both Cpl-1 and SP-CHAP were used at a concentration of 1.2 mg/ml.

Statistical Analysis

Unless otherwise noted, all *in vitro* experiments had a minimum of 3 biological replicates, with ≥ 3 technical replicates. GraphPad Prism 8 (La Jolla, CA) was used for statistical analyses. Comparisons between two cohorts at a single time point were calculated by Mann-Whitney U test. Comparisons between groups of >2 cohorts or groups given multiple treatments were calculated by ANOVA with Tukey's (one-way) or Sidak's (two-way) post-test or by Kruskal-Wallis H test with Dunn's multiple comparison post-test, as determined by the normality of data groups. Repeated measures are accounted for whenever applicable.

Chapter 4: Discussion and Future Directions

4.1 Summary of the Dissertation

Streptococcus pneumoniae is a significant pathogen, with a wide variety of virulence factors at its disposal, including the ability to form biofilms. It causes infections ranging from simple otitis media to more serious disease such as bacteremia, invasive pneumococcal disease and meningitis, while having the biggest impact on children and the elderly. The emergence of antibiotic resistant strains represents a major challenge for pharmaceutical companies and healthcare networks worldwide, prompting the search for alternatives. This prompted us to research the Cpl-1 endolysin, which displays antimicrobial activity against *S. pneumoniae*.

The goal of this dissertation was to identify the molecular mechanism for dimer formation of the pneumococcal endolysin Cpl-1, in the presence of choline. While there has been a lot of speculation on the mechanism, to date no one has conclusively established it. In Chapter 2 we sought to determine this. A prior study by Buey et al. proposed a model for Cpl-1 dimer formation by fitting the monomeric Cpl-1 crystal structure obtained in the absence of choline into a SAXS envelope generated in the presence of choline (Buey et al., 2007). From this, they identified and hypothesized three aromatic residues involved in the formation of the Cpl-1 dimer in the presence of choline, all located at the N-terminus of the cell binding domain. However, they did not actually test whether these residues were involved in dimerization. In contrast, we went about testing their hypothesis, disproved it using various mutagenesis studies, and analytical approaches, and instead showed that Cpl-1 dimer formation occurs at the C-terminus of its cell binding domain rather than at the N-terminus. We experimentally identified residue F328 as being critical to formation of Cpl-

1 dimers in choline as a single point F328A mutant was unable to form dimers in the presence of choline as monitored by analytical size exclusion chromatography. Structural modeling using AlphaFold2 showed the F328 residue being involved in π -stacking, confirming our experimental findings.

Consistent with previous studies, we have shown that Cpl-1 binds choline with millimolar affinity (Monterroso et al., 2008) (Fig. 2-6A). We have also shown, for the first time, that Cpl-1 exhibits similar affinity towards phosphocholine, which more closely resembles the native form of choline present in the pneumococcal wall teichoic acids (Fig. 2-6B).

Through a bioinformatics approach, we identified a consensus sequence present in the C-terminus of the cell binding domain that is widespread and can be used to predict dimer formation in pneumococcal cell wall hydrolases, including both pneumococcal autolysins and bacteriophage endolysins. Our search led to the identification of a novel endolysin we named SP-CHAP. Not only does SP-CHAP contain the C-terminus dimerization consensus sequence (and we show it does dimerize), but it is also the first pneumococcal endolysin identified to date possessing a cysteine, histidine-dependent amidohydrolase/peptidase (CHAP) domain.

In Chapter 3 we took our findings from Chapter 2 further and sought to biochemically characterize this new endolysin and also evaluate its antimicrobial properties. We show that SP-CHAP is active in all physiologically relevant conditions (pH, temperature) and displays bactericidal activity against various *S. pneumoniae* strains, both encapsulated and capsule free, while not affecting oral/nasal commensal organisms. Our results showed SP-CHAP was highly efficient at eradicating static pneumococcal biofilms as seen by quantification of crystal violet staining and subsequently visualized by confocal microscopy. We show SP-CHAP does not elicit host toxicity and is a successful decolonizing agent *in vivo*. Our preclinical evaluation of SP-CHAP

in a mouse *S. pneumoniae* nasopharyngeal colonization model did not result in any adverse effects and protected mice from infection by killing *S. pneumoniae* when 60 µg was administered intranasally. Of importance, the bacteriolytic activity of SP-CHAP was greater than Cpl-1 against planktonic cultures, pneumococcal biofilms and in a nasopharyngeal colonization model. Our findings warrant SP-CHAP be pursued further for translational development as a highly active pneumococcal endolysin.

4.2 Discussion & Future Directions

The scourge of global mortality due to ravaging bacterial infection now attends the decades-long decline of the golden age of antibiotic discovery which big pharma has since exited. Of 13.7 million infection-related deaths in 2019, lower respiratory and bloodstream infections each caused more than 2 million deaths. Along with *Staphylococcus aureus*, *Streptococcus pneumoniae* leads in causing this crisis; it happens to be the only organism in the spotlight through global surveillance measures and public health initiatives (GBD Antimicrobial Resistance Collaborators, 2022). Combating the situation in breadth and depth demands deploying horizontal and vertical global organizational power. Just as mighty oaks from little acorns grow, creative solutions can germinate from innovation-driven scientific labs, but their application to restoring global public health needs organizational leadership from public-private-policy partnerships and the navigational might of big pharma. A parallel can be drawn to Alexander Fleming's laboratory curiosity of penicillin turning into a life-saving drug through the Oxford team of Howard Florey and Ernst Chain, who unearthed its chemotherapeutic properties, and the consortium of five US companies who followed through with its large-scale manufacture (Aldridge, 1999).

Amidst the widespread recognition, even alarm, about how devastating the antimicrobial resistance situation can become, there is the Combating Antibiotic-Resistant Bacteria Biopharmaceutical Accelerator (CARB-X), set up as a global nonprofit public-private partnership initiative in 2016. Such enterprises dedicated to accelerating antibacterial-related research to tackle the surging development of drug-resistant pathogens afford hope (Alm & Gallant, 2020).

CARB-X's composition of partners is loaded with domain expertise in infectious disease, information technology resources, and the wherewithal to fund and raise funds from philanthropic agencies across the world committed to bettering the universe. These include the Biomedical

Advanced Research and Development Authority (BARDA) of the US HHS, the Wellcome Trust of the UK, Germany's Federal Ministry of Education and Research (BMBF), the UK Government's Global Antimicrobial Resistance Innovation Fund (GAMRIF), and the Bill & Melinda Gates Foundation. With about half a billion dollars available to invest, the mission of CARB-X is to identify, select, and fund topnotch science from around the world that aligns with its goal.

CARB-X finds the current antibacterial ecosystem fragile, primarily because big pharma is circumspect on financial returns from this therapeutic area, and smaller companies while leading in innovation may not have the bandwidth to chaperone and champion research programs all the way into the clinic. As of 2020, CARB-X had funded 60 projects in 7 countries and had an active portfolio of 39 projects (Alm & Gallant, 2020). It is encouraging to see that commitment to innovation and the demand for funding and support to this critical therapeutic area of dire need is still strong.

The CDC estimates over 2.8 million infections and 35,000 deaths each year are caused by antibiotic-resistant bacteria in the United States alone (CDC, 2019). Resistance being a global issue, the misuse and overuse of antibiotics everywhere keeps making the problem worse. While antibiotic overuse has led to increased resistance development, it has also contributed to illnesses and deaths associated with various opportunistic pathogens. For example, the nonspecific killing of beneficial commensal bacteria that colonize the gut microbiome due to the broad-spectrum action of antibiotics has helped in the pathogenicity of *Clostridium difficile*. The CDC reports that over 223,000 illnesses and 13,000 deaths each year are associated with *C. difficile* infections and has classified it as an urgent threat to human health. Now that the gut-lung axis is recognized in mediating respiratory outcomes (Lane et al., 2023), the deleterious effect of antibiotic resistance

encountered while treating infections in other organ systems could worsen lung health due to disruption of the intestinal flora. The economic burden caused by antibiotic resistant infections is significant enough to warrant addressing the threat through alternative antimicrobial therapies.

Endolysins, are one such attractive alternative to antibiotics, to address the surge of antimicrobial resistance. Endolysins have been evaluated for their ability to lyse bacterial cell walls *in vitro* for several years. However, it was not until the first *in vivo* mouse study conducted by Vincent Fischetti's group in 2001 to treat a streptococcal infection, that endolysins started gaining attention to be used as therapeutics (Nelson et al., 2001). While their application extends to various fields other than medicine, such as food safety, agriculture and animal safety, one of their main advantages over antibiotics is that resistance to these antimicrobials has not been documented to date, despite studies conducted to generate resistant mutants. The high species specificity displayed by endolysins, modulated by its cell binding domain, enables them to be used as targeted therapies to target only a particular pathogen of interest while leaving others unaffected, which is especially important for the gut microbiota and to reduce the chance of developing secondary infections as a result of antimicrobial treatment.

In this context to take our work further, while we assessed the bacteriolytic activity of SP-CHAP in Chapter 3 and benchmarked it against Cpl-1, we would like to further investigate its antimicrobial activity when used synergistically in combination with other endolysins and also in combination with various antibiotics. SP-CHAP can be used in combination with Cpl-1, Pal, and LytA to evaluate how a combination of endolysin treatment can achieve bacteriolytic killing, both *in vitro* and *in vivo*. Cpl-1, Pal, and LytA all possess catalytic domains that are different from the CHAP domain present in SP-CHAP. Synergy here could be manifested as the simultaneous cleavage of different bonds in the peptidoglycan, leading to enhanced cell lysis. In addition, it is

probable that one endolysin can break a bond in the peptidoglycan enabling the second endolysin to have better access to the bond it cleaves, allowing for the targeting of multiple different sites that ultimately leads to cell lysis. SP-CHAP can also be used in combination with antibiotics such as penicillin, daptomycin, and gentamycin. When used synergistically with other endolysins and antibiotics, we would expect the dose requirements to be lower as compared to when SP-CHAP is used independently, and thereby display a lower MIC. The synergy results can be used to help translate into a therapeutic cocktail formulation that could be used clinically. A cocktail containing endolysin combination would also be associated with a lower likelihood of development of resistance since the effective dose of each endolysin would be lower when used in combination. While the regulatory hurdles associated with a protein cocktail are certainly significant, the benefits associated with it would warrant going through the process.

A future experiment would involve testing the activity of SP-CHAP in different mouse pneumonia models. In Chapter 3, we showed efficacy in an *in vivo* mouse nasopharyngeal model. Given that *S. pneumoniae* causes a wide variety of infections ranging from otitis media to bacteremia and meningitis, we propose SP-CHAP be tested in mouse otitis media, meningitis, endocarditis, and bacteremia models. Demonstrating the efficacy of this endolysin in different models with varying severity of pneumococcal disease is critical for the advancement of the endolysin to clinical trials. A better understanding of the pathogenesis of how *S. pneumoniae* can radiate and migrate into the bloodstream and cross the blood-brain barrier to cause meningitis would enable more targeted therapies.

In Chapter 2 we identified the F328 residue in Cpl-1 to be critical for formation of a dimer in the presence of choline, and we showed that SP-CHAP also forms a dimer in choline. As expected, SP-CHAP also contains the corresponding phenylalanine residue as part of the C-

terminal consensus sequence in its cell binding domain. Given that Resch et al. have shown construction of a pre-formed dimer of Cpl-1 had greater activity and a higher bioavailability *in vivo* (Resch et al., 2011b), a future experiment would be to construct a pre-formed dimer of SP-CHAP and evaluate its activity and bioavailability. While we would expect a pre-formed SP-CHAP dimer to display enhanced activity similar to the results obtained by Resch et al., we cannot be 100% certain. The ClyJ-3m endolysin, which was engineered to be a monomer in the presence of choline, showed higher lytic activity on *S. pneumoniae* *in vitro* and *in vivo* compared to its dimeric parent enzyme ClyJ-3 and Cpl-1 (Luo et al., 2020). In contrast, our own activity studies on the monomeric Cpl-1 F328A mutant showed a decrease in bacteriolytic activity compared to wild-type Cpl-1 suggesting dimerization is essential for full lytic activity. Thus, it would be interesting to investigate whether a SP-CHAP monomeric mutant still displays lytic activity, and if this activity is less than or greater than wild-type SP-CHAP. Construction of a pre-formed dimer will also have important implications when considering the pharmacokinetics of SP-CHAP.

The Cpl-7 endolysin encoded by the pneumococcal Cp-7 bacteriophage comprises a catalytic domain with muramidase activity and a cell wall binding domain made up of three almost identical repeats (CW_7 repeats). Cpl-7 is the only pneumococcal endolysin that does not bind choline on the surface of pneumococci, yet it successfully lyses *S. pneumoniae*. Published results suggest that the CW_7 repeats recognize the peptidoglycan network as the binding target of Cpl-7 (Bustamante et al., 2012). A study by Bustamante et al. sought to identify the binding epitope of Cpl-7 on the bacterial cell wall (Bustamante et al., 2017). Through saturation transfer difference NMR spectroscopy, they found that Cpl-7 binds to GlcNAc-MurNAc-L-Ala-D-isoGln (GMDP) on the bacterial cell wall. GMDP is a highly conserved region of the bacterial peptidoglycan and is found in most Gram-positive bacteria. Because Cpl-7 binds to a conserved region of bacterial

peptidoglycan, future experiments could exploit this and expand the host-range and lytic activity of Cpl-7 by evaluating different catalytic domains fused to the Cpl-7 CBD containing the CW_7 motifs. This approach can help inform a long-term goal of evaluating constructs with multiple catalytic domains that may act synergistically. To test whether the Cpl-7 CBD does indeed bind to multiple different bacteria, we have obtained unpublished preliminary data on the host-range of Cpl-7. As seen in Table 4-1, the Cpl-7 CBD bound to all *Streptococcus* and *Enterococcus* strains tested while no binding was observed with *Staphylococcus* or *Bacillus* species. Notably, the GMDP motif is absent in *Bacillus* spp. This preliminary host-range analysis offers hope that novel chimeric enzymes could be generated with broad specificity and lytic activity. While our preliminary host-range data is not developed enough for a manuscript, future researchers could build on our findings in the ways outlined below.

Table 4-1. Cpl-7 CBD host range.

Bacteria	Strain	Cpl-7 CBD binding
<i>Streptococcus pneumoniae</i>	R6	++
<i>Streptococcus pneumoniae</i>	D39	++
<i>Enterococcus faecalis</i>	EF24	+
<i>Enterococcus faecalis</i>	EF25	++
<i>Enterococcus faecalis</i>	JH2-2	+
<i>Bacillus thuringiensis</i>	BAA10792	+
<i>Bacillus anthracis</i>	UM23	-
<i>Bacillus cereus</i>	11778	-
<i>Staphylococcus aureus</i>	COL	-
<i>Staphylococcus aureus</i>	RN4220	-
<i>Streptococcus mutans</i>	B14	++
<i>Streptococcus mutans</i>	25175	-
<i>Streptococcus pyogenes</i>	D471	++
<i>Streptococcus dysgalactiae</i>	26RP66	++
<i>Streptococcus dysgalactiae</i>	21597	++
<i>Streptococcus mitis</i>	J22	++
<i>Streptococcus uberis</i>	27598	++
<i>Streptococcus uberis</i>	700107	++
<i>Streptococcus equi</i>	9528	++
<i>Streptococcus equi</i>	700400	++

Note: “++” denotes strong binding, “+” denotes weaker binding, “-” denotes no binding observed

As a future experiment, the Cpl-7 catalytic domain could be replaced with different catalytic domains such as amidases, endopeptidases, glucosaminidases, and muramidases and activity studies could be conducted. The best (most active) constructs could be chosen and the number of CW_7 repeats be modulated. Wild-type Cpl-7 has three CW_7 repeats and a construct with two CW_7 repeats had higher lytic activity as compared to wild-type Cpl-7 (Bustamante et al., 2017). Since this construct with two repeats had higher lytic activity, it is reasonable to

hypothesize that modulation of the number of CW_7 repeats for the best construct chosen may generate an enzyme with even better activity. In order to further our efforts to come up with a highly active and broad spectrum endolysin, the CBD of the most active construct could be modulated. The CW-7 repeats could be replaced with the CBD from Cpl-7S, an engineered endolysin in which the net charge of the CBD has been inverted (Diez-Martinez et al., 2013). Cpl-7S has fifteen amino acid substitutions which results in a change in net charge of the CBD from -14.93 to +3.0 at neutral pH. Cpl-7S has been shown to have improved lytic activity against *S. pneumoniae*, as well as *S. pyogenes*. It is likely that inversion of the net charge of the CBD of our construct may result in a more active enzyme with higher lytic activity. This is a reasonable strategy to improve the lytic activity of endolysins since it facilitates their pass through the negatively charged bacterial wall enabling interaction with the cell wall target. Engineering and domain swapping of Cpl-7 with the goal of developing a “holy grail” endolysin that binds to and lyses multiple different bacteria could be very useful.

The COVID-19 pandemic has led to the discovery and use of mRNA vaccines that enabled pharmaceutical companies to produce vaccines against SARS-Cov-2 in record time. This mRNA approach instructs cells to make the spike protein found on the surface of the COVID-19 virus, resulting in antibody production thereby providing protection against future infection. As a future experiment, we propose using the mRNA-based approach and applying it to endolysins as an mRNA based-antibacterial therapy. Administration of synthetic mRNA that encodes for an endolysin would enable endogenous synthesis and production of endolysin. This would enable organ targeting, overcoming a challenge associated with endolysin site-specific delivery. To this end, we propose applying SP-CHAP mRNA to human cell lines that would result in the expression and production of SP-CHAP in the cytosol. SP-CHAP mRNA could be transfected into HEK293

or A549 cells, the cells lysed, and the lysate be tested for bacteriolytic activity against *S. pneumoniae*. If activity is seen, this would suggest that SP-CHAP encoding mRNA was successfully translated into a biologically active form of the protein. This can prove to be very useful for treating intracellular *S. pneumoniae* infections. To take this further, we could attach a signal peptide (SP) to the N-terminus of the SP-CHAP sequence so as to enable the secretion of this endolysin. Secretory proteins are equipped with various signal peptides that direct proteins to the secretory pathway via the endoplasmic reticulum. The supernatant of cells transfected with SP fused SP-CHAP mRNA could be tested for activity against *S. pneumoniae*. If the supernatant displays killing activity against *S. pneumoniae*, this would suggest that addition of the SP led to secretion of the endolysin, thus having implications for extracellular targeting of bacteria. Given that endolysins would be produced in mammalian cells, they would be subject to post translational modifications such as glycosylation. Glycosylation is either N-linked or O-linked and in the case of N-linked glycosylation occurs when oligosaccharides attach to the nitrogen atom of asparagine residues (Khoury et al., 2011). This is typically associated with a consensus sequence of Asparagine-X-Serine/Threonine, where X can be any amino acid except proline, in the case of N-linked glycosylation (Kelleher et al., 1992). Individual amino acid modification to prevent glycosylation can help retain the biological activity of the endolysin. An innovative approach such as this, if successful, could be expanded to target other pathogenic bacteria by utilizing mRNA from endolysins that target their respective bacteria. At the very least, this would serve as a proof-of-concept study to validate the capability of human cells to produce endolysins from their mRNA. This mRNA based endolysin approach can be especially useful to treat intracellular pathogens such as *S. pyogenes*. Taking it further, we propose application of endolysin mRNA on an accessible or targeted site, such as a wound infection or bladder to treat urinary tract infections. For example,

MRSA is commonly associated with wound infections. Application of a patch containing endolysin mRNA that targets *S. aureus* to a wound would lead to site-specific endolysin production that can cause targeting killing of *S. aureus*. Finally, we have shown that the Cpl-1 endolysin can be spray-dried and converted into a powder for inhaled delivery to treat pulmonary infections, while retaining activity (Wang et al., 2023). Application of Cpl-1 mRNA to the lungs through the use of a nebulizer will enable site-specific targeting.

As mentioned in Chapter 1, endolysins do generate an immune response, however, this does not seem to significantly affect their activity. For *in vivo* applications, a reduction in the immune response would nevertheless be beneficial. A recent study has demonstrated immunogenic epitope scanning to identify immunogenic amino acids that were altered to modulate the immune response using the EndoScan technology (Harhala et al., 2023). In this technology, a phage display library of all epitopes in the endolysins was constructed with a nucleotide printed library of oligopeptides representing overlapping sequences of the Pal and Cpl-1 endolysins. The library was then incubated with mouse serum containing high levels of anti-Pal and anti-Cpl-1 IgG. The investigated sera were immunoprecipitated with IgG specific magnetic beads and a fraction of the phage display library that exposed endolysin epitopes reactive to IgG present in sera was isolated. Since the library contains whole phage particles, each phage that displays a specific epitope contains a sequence coding for this epitope in its genome. Then, next-generation sequencing of the relevant region in the library genomes was used to reveal which epitopes were effectively recognized by specific IgG. Key amino acids residues encompassing the various epitopes that generated an immune response were subsequently identified. One can envision applying the EndoScan technology to further the translational development of SP-CHAP. Upon identification of key amino acid residues that contribute to an immune response, these residues could be mutated

and the enzyme assessed for a change in immunogenicity and evaluated for its antibacterial killing ability. Because antibodies bind specific sites within proteins that are antigenic, such a targeted approach to disrupt endolysin antigenic activity can be useful when designing an engineered endolysin for therapeutic applications, as long as its biological activity is retained.

When one looks into the future for directions to steer the present study, the attributes of endolysins lend amenability for multiple translational developments. Engineering enhances the intrinsic characteristics to make endolysins even more attractive. For instance, low resistance potential (if at all) and specificity of antimicrobial targeting are vital characteristics that make an endolysin a suitable research candidate for enhancing its druggability (i.e. pharmaceutical characteristics like absorption, distribution etc.) in various ways. Targeted delivery is an area where progress is being made in applications such as inhalation (Wang et al., 2023), hydrogels (Li et al., 2023) and nanoparticles, which confer higher potency and thermal stability (Dzuovor et al., 2022).

There will undoubtedly be developments in technologies for rapid detection of bacterial infection in field locations rather than restricted to sophisticated central clinical laboratories as at present. With rapid growth in information technology and broadband connectivity across the globe, one hopes and expects it will become easier to detect and contain the geographic spread of infection outbreaks closer to where they fulminate.

Driven by the global medical call for action, and stimulated by incentives and policies through public-private-philanthropic initiatives, there is reason to expect that large pharmaceutical companies too may reenter the field of antimicrobial therapeutics. This could happen through standalone pharma majors, consortia, or mergers and acquisitions with small innovative research

boutiques. A variety of research approaches strung together with bioinformatics involving combinatorial and newer technologies could enhance the success rate.

As an example, the mechanism of dimerization of Cpl-1 having been conclusively elucidated, its conformation having been studied now with the rigor of the 2023 Lasker-award winner AlphaFold2, and the novel endolysin SP-CHAP having been characterized, are factors that can open a unique opportunity in the infectious disease space for investigation through a medicinal chemistry approach. Protein mapping of G-protein coupled receptors (GPCR's) for intra-receptor distances to optimize ligands for specificity of pharmacologic profile, exploration of homomer and heteromeric possibilities, factoring out the addiction potential from newer opioid drugs used for intractable pain, etc. are examples of research that have capitalized on state-of-the-art science for better probes and drug candidates. Conceivably, the information available from the present endolysin investigation could be similarly pursued to engineer endolysin molecules by a combination of molecular biology and chemical techniques, just as fermented penicillin technology spawned semi-synthetic, highly successful analogs like amoxicillin with finer-tuned profiles. This could be possible since we now have information not only on the cell wall of the pathogenic organism, but also details of the endolysin molecules that bind in a specific manner at specific site(s) to the cell wall to selectively exert their hydrolytic action. Just as more than a single pharmacophore moiety can be crafted into a drug candidate via linkers to simultaneously modulate responses at multiple opioid GPCR's or biochemical pathways, designer infectious disease treatments could emerge on the horizon. As an example, to illustrate the concept, we could conceivably combine more than one endolysin of interest into the same molecule, to achieve therapeutic goals such as desired narrow or wide spectrum of activity, avoid development of

resistance, target the therapy to a desired site, etc. and stay clear of disrupting the commensal microflora.

In conclusion, this study investigates the molecular basis for Cpl-1 endolysin dimer formation and challenges a long accepted mechanism, while identifying a novel C-terminal consensus sequence shared by pneumococcal endolysins. The discovery of a new endolysin, SP-CHAP, exhibiting antimicrobial activity against *S. pneumoniae* and possessing a unique cysteine, histidine-dependent amidohydrolase/peptidase (CHAP) domain, adds to the arsenal of potential therapeutics. SP-CHAP demonstrates robust antimicrobial properties, including efficacy against *S. pneumoniae* biofilm eradication. Validation in a mouse model underscores its therapeutic potential, marking SP-CHAP as a promising candidate for further translational development as an effective endolysin against *S. pneumoniae* infections.

Appendix A: Sequence of 76 novel endolysins identified containing C-terminal consensus sequence

Of the 2,628 putative novel endolysins identified by Fernandez-Ruiz et al., in Chapter 2 we identified 76 that contained a C-terminal tail consensus sequence (FxxEPDGLIT). The amino acid sequences of these 76 enzymes are listed in this appendix. The “SRS” numbers corresponds to the sequence IDs from the (Fernandez-Ruiz et al., 2018) paper. Text highlighted in yellow corresponds to the tail consensus sequence established in Chapter 2 of this dissertation.

>7000000001.a:SRS022530_LANL_scaffold_28930

MVEIINTNVFNGIAGARPTDKPKYYILHNDAGSMTPESYIGWLQERYNNGQSALGFAHYIINRDTIARVENTYS
GTWSVANYDANMNSISYEVCQQYNASDAEFIEENMVLRLQMAEDMTYYGDTPNYSNIKFHNEFSSTSCPKRSLA
LHGGYNDSLRDYVIAKIKYYQSLGSTVQEMLGEAKAEAGWHNNSTGWWDVNEEDGTYPNKGQVIDDVWYYFDSN
GYMKANTWHKHSDTNWYYLLPDGSMATGWVLI DNSWYFDNQAMVTGWVKYKDTWYYLDYQKGSMSVNAFVKS
ADGKGWYYLKDDGSLADKPEFEVEPDGLITLAHTKEETTK*

>7000000001.a:SRS022530_LANL_scaffold_62421

MTINIEAIAWMRARQGQVSYSMEDRDGPDSYDCSSSVYYALRSAGASSAGWAVNTEYEHDLVKNGYELITEN
SPWDAQRGDIIFIWGRGASAGAFGHTGMFIDEDNIIHCNYAYNGISVNDHDERWYYAGQPYFYIYRLTNPDAQP
EEPKKGWQKDDQGHWYARANGSYPKAEFEYIEENKSWFYFDAEGYCVKSDWVLHTDGKWYFDDKGYMATSWKK
VADKWYFNRDGAMQTGWVKYDKWYLDATNGDMKSDCFIKYNDGWYLLLPDGRGLADKPEFTIEPDGLITTK*

>7000000002.a:SRS022532_LANL_scaffold_9343

MYANRWLKHDTGHWYFWDSSGYMATSWKKIAGVWYFNRDGAMQTGWVKYKWKWYLDVANGDMKSNTFVPYNG
GYLLLPDGRGLADKESFKIEPDGLITTK*

>7000000005.a:SRS022536_LANL_scaffold_30505

MDIDKSRLRSGPLQVGVQPYRQVHAHSTGNRNSTAQNEADYHYRKDPELGFFSHVVGNGRVMQVGPVNNGSWDV
GGGWNAAESYAAVELVESHGSKEEFMRDYKLYVELLRNLADEAGLPKTLDTGSLAGIKTHEYCTNNQPNHSDHV
DPYPYLAKWGISREQFKKDIEDGLSESGWKRNETGWWWEESDGSYPTDRWKRVNNEWFYFDNRYCLINRWFND
GQDWFYLDKRGAMVTGWMYINNRYWYFKSDGRMAKGWVKYRETWYYLDEKDGDMKSDQFVKYGNGWYYLKPDGS
MADKPEFTVEPDGLITTK*

>7000000005.a:SRS022536_LANL_scaffold_52876

MGLNLETAIAWMRARKGQVSYSMDDRNGPDSYDCSSSIYYALLSGGAVSAGWAVNTEYEHDLKKNKYELIAEN
TPWDAKRGDIIFIWGRRGYSAGAFGHTGMFIDGDNIIHCNYAYDGISVNDHDERWAYAGKPYYYIYRLTNPDPATP
APVKKGWQNDGTGYWFRANGSYPKDQFEYIEENKSWFYFNSEGYMVAEDWVKHTDGKWYFDDKGYMATSWKK
INGKWYFNRDGSMQTGWVKYKWKWYLNSENGDMVSNFVPPYNGGYLLLPDGRGLADKESFNIEPDGLITTK*

>7000000007.a:SRS017139_Baylor_scaffold_85819

MMRMSKKQDMINDLMHADAGTGVDYDGMYGQCADVTCYGIYEYFGTRLWGNAIDLLRSAAEAGLQVYGAQY
PKAGWFFVKNFVAGDGVNYGHTGLVYEDSDGSTIKTIEQNIDGNWDYLEVGGPCRYNERSVNSIVGYIVPPEED
QSGWKRNDTGWWWSRKDGSYPTSKFEKIADTWYFDSRGYMYAERWLKHTDGYWYFDTSGAMVTGWKKIGGLW
YYFNRDGAMQTGWVKYKWKWYLDVANGDMKSNTFVPYNGGYLLLPDGRGLADKAAFTVEPDGLITTK*

>7000000009.a:SRS018300_Baylor_scaffold_55979

MTINIEAIAWMRARQGQVSYSMEDRDGPDSYDCSSSVYYALRSAGASSAGWAVNTEYMHDLIKNGYELIAEN
TECTAQRGDIIFIWGRKGASAGAFGHTGMFIDSDNIIHCNYAYNGISINNHDERWYYAGQPYFYIYRLTNPDAQP
EEVKKGWQKDSTGYWYARANCSPKAEFEYIEENKSWFYFDESGYMAEKWLHHTDGNWYYFDKDGYMATSWKK
IDGKWWYFNRDGSMQTGWVKYYDKWYYLDTTNGDMKSDCFVKYNNGWYLLLPDGRMAEKEAFTVEPDGLITAK*

>7000000018.a:SRS022725_LANL_scaffold_49418

MERRVSMGLNLETAIAWMQARKGQVSYSMDYRNGPDSYDCSSSIYYALRSGGASDAGWAVNTEYEHKWLIDNGY
TLISENTPFDAQRGDVFIWGRKGTSSGADGHTGIFIDSDNIIHCNRYDGTITVNDHDDIWVYAGKPYYYIYRLT
NQNAAPKVIKWKTDAGDQWYQRANGTWPDKGFYIEENKSWFYFNSEGYCVKSDWILHTDGEWYYFDADGYM
VTSWRRIDGKWWYFNRDGSMKKGWVKYYDDWYFLDEKHGDMLSDRFVPYKNGYYMLLPDGRMADKEAFNIEPDG
LITTK*

>7000000025.a:SRS023617_Baylor_scaffold_11604

MDIDTSRLRTDLPQVGIQPYRQVHAHSTGNRNSTAQNEADYHMRPAPESGFFSHVVGNGRVMQTLTDRGAWDV
GGGWNVEGYAQVELIESHESKEEFMRDYRLYVQLLRELADEAGIPKTLDSDSLAKIKTHQYCTYNQPNNASDHV
DPYPYLAKEGWSREQFKKDIEGGISTEAGWRQDAYGWWWEESDGSYPTNDWKQINGEWFRRDDKGYCLINKWFF
DGKYWFYLDKRGATVTGWVFINHRWYFDNDGGMVKGWVKYRETWYYLDDKDGMYLSKQFVKSGDGWYYLKANG
ELHTEPAFTIEPDGLITVSE*

>7000000027.a:SRS056622_LANL_scaffold_11215

MEIDKSRLRTNLPQIGEOPYRQVHAHSTGNPNSTAQNEADYHMRPVESEGFFSHVVGNGRIMQTLVDRGAYDV
GGGWNVEGYAQVELIESHESKEEFMRDYRLYVQLLRELADEAGIPKTLDSSSLAKIKTHQFCTYNQPNNGSDHV
DPYPYLAKEGWSREQFKKDIESGLTEGNWKRNEVGWWEESDGSYPKSQWKNIKGEWYFDNRGYCFINKWFND
GKDWYFDKRGAMVTGWMHIDHRWYFYSKSDGRMAKGWVKYRETWYYLDEKDGDMKSKQFIKSGNGWYYLKADGS
LSVKPEFTIEPDGLITTN*

>7000000027.a:SRS056622_LANL_scaffold_52179

MKGKKKMTINIEDAIAWMREREGQVYYSMEYRDGPDSYDCSSSVYYALRSAGAVSAGWAVNTEYEHDLVNNGY
ELIAENQEWNAQRGDIIFIYGRRGYSAGAGGHTGMFVDSNIIHCNYARNGITVNNHDAIWNAAGQPYFYAYRLT
NPDAQPEEVKKGWQNNNDNGYWYVRSNGSPKQFEKIDGTWYYFDENGYMLSDKWKQRPDGTWYYFDNSGEMAT
DWNKIDGKWWYFNRDGAMATGWVKYYDKWYYLDVVNGDMKSDCFIKYNDGWYLLLPDGRMADKPEFTVEPDGLI
TTK*

>7000000028.a:SRS024138_Baylor_scaffold_48614

MKKNLDFIDVSSHNGYDISGLETMTNTTIKISESTSYINPCLSAQIEQSTPVGFYHFAWFGGDVEEAEREA
RYFLDNVPTQVPYLVLDYEDHASGDRQANTNACLRFMQIIADAGYKPIYYSYKPFITLNNIDYHQILEQFPNSLW
VAGYGLNDGNADFEYFSPMDGIRWWQYSSNPFDKNIVLLDDEPEQEPKTAGTWKKDSTGWWFRRANGSFYPYKWK
EKIADVWYYFDSKGYCLTSQWVLYKDKWYYLKDNGAMVTGWALVGSEWYMDDSGAMVTGWVKYKNNWYYMTNE
RGNMISNEFIKSGKGYFMNANGELADNPSFTKEPDGLITVA*

>7000000039.a:SRS013164_Baylor_scaffold_37024

MAVNIDNAIAWMQARKGKVYYSMEYRDGEESYDCSSSVYYALRSAGASSAGWAVNTEYQHNWLIDNGYTLISEN
TPWSAQRGDIIFIWGRKGASAGAGGHTGMFIDEDNIIHCNYAYDGISINDHDERWLAAGRPPYYIYRLTNPDPVTP
KEVKKGWQKDNTGYWYARANGTYPKAQFEYIEENKSWFYFDDKGYAYSQKWLKYKDGKWWYFDKDGYMATSWKK
INGKWWYFNRDGAMQNGWVKWYDKWYYLDAQNGEMKSNFTVPYNGGYLLLPDGRMADKEAFKIEPDGLITTL*

>7000000073.a:SRS021496_Baylor_scaffold_33623

MDIDKSRLRTGLPQVGVPYRQVHAHSTGNRNSTAQNEADYHYRKDPPELGFFSHVVGNGRVMQVGPVNNGSWDV
GGGWNAAESYAAVELIESHSTKEEFMTDYRLYIELLRNLAAEAGLPVTLTDDLAGIKTHEYCTNNQPNNNSDHV
DPYPYLAKEGWSRSQFKKDIEHGLVVEPGWKKNDTGYWYVRSNGSYPKQFEKIDGTWYYFDGSGYMLADKWKK
RPDGAWYYFNKSGEMATGWKLISNKWYFKEDGEMVTGWVKYKDTWYYLDSKEGNMVSNAFVKSEKGYWYLLKED
GSMADKPEFTVEPDGLITTTIK*

>7000000073.a:SRS021496_Baylor_scaffold_54760

MVVNIETAIAMQARKGNVTYSMDHRDGPYSYDCSSSVYYALREAGAVSAGWAVNTEYQHDWLLKNGYTLITEN
QPWDAKRGDIFIWGRRGYSSGAGGHTGMFVDSNIIHCNYARNGITVNDHDDIWYSAGQPYFYAYRLTNPNAKP
EEVKKGWQSDDKGDWYARANGTYPKSDFEYIEENKSWFYFDAKGVCVKSVDWVFHTDGKWWYFEEDGYMVTSWRK
INNKKWYFNRDGSMTGWIKYYDKWYYCDSVNGDMKSDAFIKYNDGWYLLLPDGRLEKPEFTVEPDGLITTK*

>7000000086.a:SRS049147_LANL_scaffold_44091

MAVNIDNAIAWMQARKGKVTYSMEYRDGEESYDCSSSVYYALRSAGASSAGWAVNTEYQHKWLIDNGYTLISEN
TPWSAQRGDIFIWGRKGASAGAGGHTGMFIDSDNIIHCNYAYNGISINDHDERWLYAGRPPYYIYRLTNPNAQP
EKPKKGWQSDDKGDWYARANGSYPKSEFEYIEENKSWFYFNAEGYCVKSVDWVFHADGKWWYFEEDGYMVTSWRK
INNKKWYFNRDGSMTGWVKYYDKWYYLDATNGDMKSDCFVKYNNGWYLLLPDGRMAEKEAFNVEPDGLITAK*

>7000000109.a:SRS011306_Baylor_scaffold_74819

MTINIEAIAWMKARQGRVSYSMEERDGDSDYDCSSSVYYALRSAGAVSAGWAVNTEYEHNLINNGYTLISEN
TPWDAQRGDVFIWGRKGASAGAGGHTGIFIDSDNIIHCNYRYDGISVNDHDDIWLYAGRPPYYIYRLTNPNAQP
EAPKKGWQKDDKGYWYARANGSYPKGFQFEYIDENKSWFYFDELGYMSDRWLKHTDGHWWYFDKDGYMATSWKK
INGKWWYFNRDGAMATGWVKWYERWYYLDSGDGMKSNTFVPYNGGYLLLPDGRMADKESFTVEPDGLITAK*

>7000000130.a:SRS062540_LANL_scaffold_21066

MTINIEAIAWMKARQQVYYSMYERDGPDSYDCSSSVYYALRSAGASSAGWAVNTEYEHDLIKNGYELIAEN
TECNAKRGDIFIWGRKGASAGAFGHTGMFIDSVNIIHCNYAYNGISVNEHDERWYAGQPYFYIYRLTNPDAQP
EEPKKGWQSDDKGDWYARANGTYPKSEFEYIEDNKSWFYFDESGYAYTEKWLHHTDGQWYFDDKDGYMATSWKK
IADKWWYFNRDGSMTGWVKYYDKWYYLDATNGEMKSDCFVKYNDGWYLLLPDGRMADKPEFTVEPDGLITTK*

>7000000130.a:SRS062540_LANL_scaffold_36374

MEIDTSRYREGLPQIGYEPYRQVHAHSTGNPNSTAQNEADYHLRREVTSGFFSHVVGNGRVMQVGPVNQGAJDV
GGGWNAAEGYGQVELIESHSTMEEFMTDYRLYVELLRNLADEAGIPKTLDSDDLEGIKTHYYCTYNQPNYS DHV
DPYPYLAKWGISREQFKHDIENGLEEIKEGWHSNSKGWYQNSDGSYPTNKWQKINEKWWYFFNEDGYCLTNKWI
KRGGVWYWLNDNGTGMATGWKKINNEWYFQKQDGMITGWVKYYDKWYYLNTNNGFMESNAFVKAGKWWYVVED
GTLDEKPEFTVEPDGLITTQKENK*

>7000000140.a:SRS044373_WUGC_scaffold_34052

MGINIENSIAMQARKGKVYYSMYERDGPDGYDCSSSIYYALRSAGGAVSAGWAVNTEYEHDLINNGFELISEN
TPWDAKRGDVFIFWGRRGYSAGAGGHTGMFVDSNIIHCNYAYNGITVNDHDERWLAAGRPFYFYVYRLTNANAKP
VEPKRGWQKDDDGFWYARTNGTYPKGQFEYIEENKSWFYFDESGYMYADKWLHHTDGNWYWFDDKDGYMATSWKK
IADKWWYFNRDGSMTGWIKYYDNWYYCDSVNGDMKSDAFIKYNDGWYLLLPDGRLEKPEFTVEPDGLITTTI
K*

>7000000155.a:SRS020862_Baylor_scaffold_18793

MSKKQDMINDLMAHADAGTGVYDKMYGYQCTDVTCTYGIYEFYFGTRLWGNAIDLLRSAESAGLQVVYGAQYPKA
GWFFVKNFVAGDGVNYGHTGLVYEDSDGSTIKTIEQNIDGNADFLEVGGPCRYNERSVDSIVGYIVPPEEDES
WKHDDTGWWWRRKEGSYPTAKFEAVDGNWFYFNENGYMYANQWLLHTDGKWWYFDDKDGYMANSWKVNGKWWY
FNADGAMQGTGWVKYYEKWYYLNSNGDMVSNTFVPYNGGYLMLLEDGRLAEKESFKIEPDGLITTK*

>7000000156.a:SRS064423_LANL_scaffold_60595

MVVNIETAIAMQARKGNVTYSMDHRDGPYSYDCSSSVYYALREAGAVSAGWAVNTEYEHDLIKNGYELISEN
QPWDAKRGDIFIWGRRGYSSGAGGHTGMFVDSNIIHCNYARNGITVNDHDDIWYSAGSPYFYAYRLTNSNSNP
EEIKRGWQKNDKGYWVFRPNTLPKDRFEYIEENKSWFYFDADGYMYADRWLKHNDGFWYWFDDKDGYMATAWK
IADKWWYFNRDGSMTGWIKYYDNWYYCDSVNGDMKSDTFIKYNDGWYLLLPDGRLEKPAFTVEPDGLITITD
K*

>7000000180.a:SRS017013_Baylor_scaffold_3390

MSRIKDMINDLMHADAGTGVDRMYGYQCADVTCTYGIYFYFGIRLWGNALDLLRSAAAGLQVVYGAQYPKA
GWFFVKNFVAGDGVNYGHTGLVYEDSDGSTIKTIEQNIDGNWDYLDVGGPCRYNERSVNSIVGYIVPPQEDQSG
WKHDDTGWWWSRKDGSYPTSKFEAVDGNWFYFNDNGMYANQWLHHTDGCWYWFDNNGYMANSGWKKINGKWYY
FNADGAMQTGWVKYIEKWYYLDSQNGDMKSDCFVPYNGGYLMLLEDGRLADKESFTVEPDGLITTK*

>7000000185.a:SRS017120_Baylor_scaffold_81774

MEIDTSRYREGLPQIGWPKFHQVHAHSTGNPNSTAQNEADYHMRPVDSGFFTHVVGNGRVMQVGPVNNGAYDV
GGGWNFETYAAVELIESHQTKKEEFLKDYRLYIELLRALADEGDI PKTLDSDSIEGIKSHEYCTYNQPNYSDHV
DPYPYLAKWGISREQFKHDIENGLSELKAGWQKNSTGWWRNVDSYFPVNKWQNIAGKFWFWDGNGYCYINRWL
KDKDKWYWLDDGDGMATGWKKISGSWYFYPKPDGSMATGWVKYDKWYYLNTNNGFMESNAFVKGKDGWYYISDD
GTLSDKPDFTVEPDGLITVKKEKETTEQESK*

>7000000194.a:SRS022077_Baylor_scaffold_49298

MVEIDKSRLRTNLPQVGVQPYRQIHAHSTGNRNSTAQNEADYHYRKDAELGFFSHVVGNGRILQTLWLTDRGAWD
VGGGWNVEGYAHVELIESHATKEEFMRDYKLYVKLLRLADEAGI PKTLDSDSLAKIKTHEFCTYNQPRNASDH
VDPYPYLAKWGISREQFKKDIEGGVNTEAGWRQNEYGWWWEADGSFPKKSWKINNEWFRFDERGYCLINRWF
YDGKYWFYLDKRGATVIGWNFINHRWYFDPKDGMYITGWVKYRETWYYLSEQSGEMLSKQFVKHGDGWYYLKAN
GELHEQPAFKVEPDGLITFIDPAKVEETTK*

>7000000197.a:SRS017808_Baylor_scaffold_49868

MDIDTSRYREGLPQIGNAPYRQIHAHSTGNKNSTAQNEADYHMRPVESEGFFSHVVGNGRVMQVGPVNNGAYDV
GGGWNYYETAAVELIESHSTEEEFLEDYRLYIELLRNLADEAGLPKTLDSDALEGIKSHEYCTNNQPNFSDHV
DPYPYLASWGISREQFKHDIENGLEVKKGWQKNDKGYWYVRSDSYPKQFEQIDGTWYFDFGSGYMLADKWKK
QPDGAWYFDFKSGEMATGWKLI SNKWYFDPDDGEMVTGWVKYDKWYYLDAVNGDMKSDCFVKYNDGWYLLLS
GRMADKPEFTVEPDGLITAK*

>7000000221.a:SRS043676_WUGC_scaffold_16785

MIMDIDTSRLRTDLPVGFEPFRQVHAHSTGNRNSTAQNEADYHYRKDPGLGFFSHVVGNGRVMQVGPVNKGW
DVGGGWNAAETAAVELIESHSTKEEFMTDYRLYIELLRNLADEAGLPKTLDTDDLAKIKTHEYCTNNQPDNSSD
HVDYPYLAKWGVSRQFKRDIENGLGSETGWQKNDTGWYVVKEDGSYPKDKFEKIDGTWYFDFGSGYMLADRW
KKHSDGNWYFDFQSGEMATGWKKIADKWYFDFVEGAMKTGWVKYKDTWYYLDSKDGNMVSNEFVRAGQGWYIK
ADGTMDADKPEFTVEPDGLITTK*

>7000000230.a:SRS015574_WUGC_scaffold_67953

MALNIETAIAMQARKGNVSYSMTSRDGDSDSYDCSSSVYYALRSAGAVSAGWAVNTEYEHQWLINNGYELISEN
TPWDAQRGDVFIFWGRKGYSAGGHTGIFVDSDNIIHCNRYRDTGTVNDHDDIWLYAGRPPYYVYRLTNPNAP
EKPKKGWQKDDTGWYFARANGTYPKAQFEYIEENQSWFYFDESGYMYSEKWLKHTDGKWYWFDSSGYMATSWKK
IAGVWYFNRDGSMTGWVKYDWNYYLDAVNGDMKSDTFVRYNDGWYLLLPDGRMADKAAFTVEPDGLITTT
K*

>7000000264.a:SRS052668_LANL_scaffold_29687

MTVNIETAIAMQARKGQVTYSMDYRNGPDSYDCSSSIYYALLSGGAVSAGWAVNTEYEHGWLEKNGYELISEN
QPWDAQRGDIFIFWGRKGYSAGGHTGIFVDSDNIIHCNRYRDTGTVNNHDDIWLYAGRPPYYVYRLTNPNAP
EKPKKGWQKDDTGWYFARANGTYPKSQFEYIEENQSWFYFDESGYMYSEKWLKHTDGKWYWFDSSGYMATSWKK
IAGAWYFNRDGSMTGWIKYDWNYYCDATNGDMKSDTFVRYNDGWYLLLPDGRMADKAAFTVEPDGLITTK*

>7000000295.a:SRS019219_WUGC_scaffold_16895

MSKIDDMVQYFIEKADAGAGVDYDGMGYQCADLTGCMGVYKFFGARLWGNALDLLRSAAAGLQVVYGAQYPKA
GWFFVKNFVAGDGVNYGHTGLVYEDSDGSTIKTIEQNIDGNWDYLDVGGPCRYNERSVNSIVGYIVPPQEDQSG
WKHDDTGWWWSRKDGSYPTSKFEAINGNWFYFNDNGMYANQWLHHTDGNWYWFDKDGMYMANTGWKKINGKWYY
FNADGAMQTGWVKYIEKWYYLDAQNGNMVSNTFVPYNGGYLMLLEDGRLAEQAEFTIEPDGLITTK*

>7000000305.a:SRS062761_LANL_scaffold_41047

MEIDTSRYREGLPQIGYAPYRQIHAHSTGNKNSTAQNEADYHMRPVESGFFSHVVGNGRVMQVGPVNNGSYDV
SGGWNYYETAAVELIESHSTKEEFMEDYRLYIELLRNLADEAGLPKTLDSDDLEGIKSHEYCTNNQPNFSDHV
DPYPYLASWGISRSQFKHDIEHGLVVEAGWKNDTGFYVVKEDGSSYPKEKFEKINDVWYYFDNSGYMLAERWKK
HTDGRWYWFDDKSGAMATSWKKIAEKWYYFNDRDGMQTGWVKYYDKWYYLDATNGDMKSDCFVKYNDGWYLLLPD
GRMADKPEFTVEPDGLITTK*

>7000000312.a:SRS014684_WUGC_scaffold_44171

MALNIETAIAWMQARKGQVSYSMDDRNGPDSYDCSSSVYYALRSAGASSAGWAVNTEYEHDWLIKNGYELIAEN
VECNAQRGDVFIWGRRGASAGAFGHTGVFIDENNI IHCNYAYNGISINNHDERWLYAGQPYFYIYRLTNPNAEP
EASKKGWQKDDKGYWYARANGSYPKDQFEKIDGTWYYFDSQGYMLADKWQKQPDGTWYYFDKSGEMATDWNKIN
GKWYYFSRDGAMVTGWVKYYDNWYFLDAANGDMKSDCFVRYNDGWYLLLPDGRMEDKPEFTIEPDGLITTK*

>7000000321.a:SRS013234_Baylor_scaffold_27571

MSKINEMVQFFIDKANEGAGVDYDGMGYQCADLTGCMGVYKFFGARLWGNADLLRSAAEASLQVVYGAQYPKA
GWFFVKNFVAGDGVNYGHTGLVYEDSDGSTIKTIEQNIDGNWDYLDVGGPCRYYERSVNSIVGYIVPPQEDQTG
WKHDDTGWWWRSKDGSYPTSKFEAINGNWFYFDESGYMYANRWLKHDTGYWYWFDDKGYMANSGWKKINGKWYY
FNADGAMQTGWVKYYEKWYYLDGQNGDMKSDCFVPYNGGYLLMLEDGRLAEQAEFTIEPDGLITTK*

>7000000382.a:SRS020858_Baylor_scaffold_2668

MAVNIETAIAWMRARQGQVSYSMDDRNGPDSYDCSSSIYYALLSAGAVSAGWAVNTEYEHGWLEKNGYELISEN
QPWDAQRGDI FIWGRKGYSSGAGGHTGMFIDSDNI IHCNWAYDGISVNDHDERWLAAGRPPYYVYRLTNAGKQA
AEPKRGWQKDDTGFWYVRANGTYPKGQFEYIEDNKAWFYFNESGYMYAEKWLHHTDGNWYWFDDKGYMATSWKR
IGGVWYFFNDRDGSMTGWIKYYDNWYYCDSVNGDMKSDTFIKYNDGWYLLLPDGRMADKPEFTIEPDGLITTK*

>7000000389.a:SRS017439_Baylor_scaffold_53002

MSVNIIEKGIAWMEARKGRVYYSMEERDGDSDSYDCSSSMYYALRSAGASSAGWAVNTEYMHDLVKNKYELITEN
TPWDAQRGDI FIWGRRGYSSGAGGHTGMFVDSNI IHCNRYRNGITVNDHDDIWLYAGRPPYYVYRLTNPDQAP
VEAKRGWQKDDTGFWYARANGTYPKAEFEYIEENKSWFYFDNQGYMLAEKWLKHDTDGNWYWFDDSGYMATSWKK
IGGSWYFFNDRDGSMTGWVKWYDNWYYLDAANGDMKSSTFVPYNGGYLLLPDGRMADKEAFTIEPDGLITAKA
*

>7000000394.a:SRS019607_WUGC_scaffold_42379

MTINIENAIAWMRARQGLVSYSDYRDGDSDSYDCSSSVYYALRSAGASSAGWAVNTEYEHDWLIKNGYELIAEN
TECEAQRGDVFIWGRKGASAGAGGHTGIFVDSENI IHCNRYRNGITVNSHDYIWALAGQPYFYIYRLTNPNAEP
EAPKKGWQEDNEGYWYARANGSYPKDEFKIDGTWYYFDENGYILSDKWKQRPDGTWYYFDKSGEMATDWNKIA
DKWYYFSRDGAMVTGWVKYYDKWYYLDAQNGDMKSDCFIKYNDGWYLLLPDGRMADKPEFTVEPDGLITAK*

>7000000402.a:SRS024580_LANL_scaffold_22243

MTINIENAIAWMRARQGNVTYSMDHRDGPYSYDCSSSVYYALREAGAVSAGWAVNTEYEHDWLIKNGYELISEN
QPWDAKRGDI FIWGRRGYSSGAGGHTGMFVDSNI IHCNYAYNGISINNHDERWYAGQPYFYIYRLTNPNAQP
EEVKKGWQKDEQGHYARANGSYPKAEFEYIEENKSWFYFDESGYAYTDKWLHHTDQWYFFDKDGYMATSWKK
IADKWYYFNDRDGSMTGWVKYYDKWYYLDAQNGDMKSDCFVKYNDGWYLLLPDGRMADKPEFTVEPDGLITTK*

>7000000402.a:SRS024580_LANL_scaffold_45241

MTLGKETMQMSKKQDMINDLIAHADAGTGVDYDGMGYQCADVTCYGIYEYFGIRLWGNADLLRSAAEAGLQV
VYGAQYPKAGWFFVKNFVAGDGVNYGHTGLVYEDSDGSTIKTIEQNIDGNWDFLEVGGPCRYNERSVNSIVGYI
VPPQEDQSGWKHDDTGWWWRSKDGSYPTSKFEAVDGNWFYFDESGYMYADQWLHHTDGYWYWFNKDGYMANTGW
KKINGKWYYFNADGAMQTGWVKYYEKWYYLDAENGDMKSDCFVPYNGGYLLMLEDGRLADKESFTVEPDGLIT
K*

>7000000405.a:SRS024277_LANL_scaffold_14367

MTINIEAIAWMRARQGQVSYSMEDRDGPDSYDCSSSVYYALRSAGASSAGWAVNTEYEHDLVKNGYELITEN
SPWDAQRGDIFIWGRKGASAGAFGHTGMFIDENNI IHCNYAYNGISVNDHDERWYYAGQPYYYIYRLTNPNAQP
EEVKKGWQKNDKGYWYVRPNGTYPTAKFEYIEENKSWFYFDESGYMYSEKWLKHTDGKWWFDDKGYMATSWQK
IGGAWYYFNDRDGSMTGWIKWYDKWYYCDETNGNMKSDCFVKYNNGWYLLLPDGRMAEKEA**FNVEPDGLIT**AK*

>7000000422.a:SRS057791_LANL_scaffold_25670

MTINIEAIAWMKAREGKVSYSMDYRDGEDSYDCSSSVYYALRSAGAVSAGWAVNTEYEHDLIKNGYELIAEN
TECEAQRGQDVFIWGRKGSSAGAGGHTGIFVDSENI IHCNRYRNGITVNSHDYIWALAGQPYFYIYRLTNPNAQP
EAPKKGWQEDDEGYWYARANGSYPKDEFEKIDGTWYYFDESGYMLADKWKRPDGAWYYFDKSGEMATDWNKIN
DKWYYFNDRDGMVTGWVKYYDKWYYLDAQNGDMKSDCFIKYNNGWYLLLPDGRMADKPE**FTVEPDGLIT**AK*

>7000000435.a:SRS015646_WUGC_scaffold_6265

MGVNIDEGIAWMQARKGQVTYSMDYRNGPDSYDCSSSIYYALLSAGAVSAGWAVNTEYEHDLIKNGYELIAEN
TPWDAQRGDIFIWGRKGYSAGAGGHTGMFIDSDNI IHCNWPYDGISVNDHDERWLYAGRPPYYVYRLTDEGKQA
VEPKRGWQKDDTGFWYVRANGTYPKGQFEYIEDNKAWFYFDDLGYMLAEKWLHHTDGNWYWFDDKGYMATSWKR
IGGAWYFFNDRDGSMTGWIKYYDNWYYCDATNGDMKSSTFVPYNGGYMLLEDGRLADKES**FKVEPDGLIT**TKQ
GGQDAESK*

>7000000436.a:SRS015644_WUGC_scaffold_44342

MTINIEAIAWMRARQGQVSYSMEDRDGPDSYDCSSSVYYALRSAGASSAGWAVNTEYEHQWLINNGYTLISEN
TPWDAQRGDVFIFWGRKGASAGAGGHTGIFIDEDNI IHCNRYRNGISVNDHDDIWLYAGRPPYYIYRLTNPNAQP
EKPKKGWQSDDKGDWYARANGSYPKSEFEYIEENKSWFYFNAEGYCVKSDWVFHTDGKWWYFEEDGYMVTSWRK
INNKWYYFNDRDGSMTGWVKYYDKWYYLDGQNGDMKSDCFVKYNNGWYLLLPDGRMAEKEA**FNVEPDGLIT**AK*

>7000000442.a:SRS051791_LANL_scaffold_44502

MVEIIERNIFQGIAGARPTESPKYFILHNDAGSMSAKSYLSWLQGRYDNGQAELGFAHYIIDRNTIVRVEDTYN
GSWSTANYDGNMNSIGYEVCCQFSASEAEFRENEEMCLRQMAEDMKYYGVEPSMETIRLHHEFSSTSCPARTLE
LHGQSNKSVREYIIKRIKYYMSIGNTVQEMLNAEIERVEGWIKNSVGWYQYNSGSPHDEINVNGETFRFDG
NGYCLINKWYYDGERGVYRYDTRGAKISGKWSQIDGWYFKEKNGDMATGWVKYKEHWYFLDNKNGDMKSNQY
IKHGDGWYFVKENGEMATEEA**FKIEPDGLIT**TVK*

>7000000445.a:SRS045262_WUGC_scaffold_1172

MGVNIDEGVAWIQARKGQVTYSMDYRNGPDSYDCSSSIYYALMSAGAVSAGWAVNTEYEHGWLEKNGYELISEN
QPWDAQRGDIFIWGRKGYSAGAGGHTGMFIDSDNI IHCNWAYDGISVNDHDERWLAAGRPFYFYVYRLTDAGKQA
AEPKRGWQKDDTGFWYVRPNGTYPKGQFEYIEDNKAWFYFNESGYMYAEKWLHHTDGNWYWFDDKGYMATSWKR
IGGAWYFFNDRDGSMTGWIKYYDNWYYCDATNGDMKSSTFVPYNGGYMLLEDGRLADKEA**FNVEPDGLIT**TKQ
GGQDA*

>7000000446.a:SRS013948_WUGC_scaffold_18410

MAVNIGTVIAWMRAKQGNVTYSMEHRDGPYSYDCSSSVYYALREAGAVSAGWAVNTEYEHDLIKNGYELISEN
QPWDAKRGDIFIWGRRGYSAGAGGHTGMFVDSNI IHCNRYRNGITVNDHDDIWYSAGTPYFYAYRLTNPNAQP
EEVKKGWQKDDKGHWYARANGSYPKDEFEKIDGTWYYFDESGYMLADKWKRPDGTWYYFDKSGEMATDWNKIN
DKWYYFSRDGAMVTGWIKYYDNWYYCDSVNGDMKSDTFIKYNDGWYLLLPDGRLEKPA**FTVEPDGLIT**ITDK*

>7000000461.a:SRS045127_LANL_scaffold_9842

MDIDKSRLRTGLPQVGVRPYRQVHAHSTGNRNSTAQNEADYHYRKDPPELGFFSHVVGNGRVMQVGPVNNGSWDV
GGGWNAAESYAAVELIESHKTQEEFNDRDYRLYVPLLRNLADEAGLPKTLDTDDLGIKTHEYCTNNQPDNHS DHV
DPYPYLAKWGISREQFKHDIEHGLDVEPGWKQNDTGYYVHKDGSYPKEKFEKIKGVWYYFNNLGMYLAECWQK
HKDGKWWYFDDKSGAMATGWKSI SNKWWYFDADGAMVTGWVAYKDKIYHLKETSGEMTSKEFEVKTEKGWYYVND
GTRADKPE**FTVEPDGLIT**TVVK*

>7000000501.a:SRS047265_WUGC_scaffold_6825

MGLSLETAIAWMQARKGQVSYSDYRNGPDSYDCSSSVYYALLSAGASSAGWAVNTEYEHKWLEDNGYTLISEN
TPFDAQRGDIIFIWGRKGYSAGAGGHTGIFIDSDNIIHCNWPYDGISVNDHDERWLYAGKPYYYIYRLTNPEAAP
KVIEKGWKSDDSGDWYQRANGTYPKSEFEYIEENKSWFYFNYEGYCVKSDWILHTDGEWYFFDADGYMVTSWRR
IDGYWYFFNRDGSMDRGWVKYYSWYLLDEKDGMLSSRFVYPNDGYMMLLPDGRADKAA**FVVEPDGLIT**TK*

>7000000511.a:SRS018591_WUGC_scaffold_45191

[Renamed SP-GH25 in this dissertation]

MKKNDLFIDVSSHNGYDITGILYSMGTNTTIKISESTSYINPCLSSQAEQSTPVGFYHFAWFGGDVSQAEAEA
QFFLENVPQSVKYLVLVDYEDHASGDAQANTDACIRFMDVIKNAGYEPIIYYSYKPTLSNVYYEQIIAKYPNSLW
IAGYGLNDGNADFEYFPSMEGIRWWQYSSNPFDKNIVLLDDEETEIDSGWKKNDTGFWYVRSDDGYTPKDQFEKI
DGTWYFFDENGMYLSDKWKQRPDGTWYFFDKSGEMATDWKKIADKWYFFNRDGAMATGWVKWYEWYLLDSENG
GMKSDCFIKYNDGWYLLLPDGRMADKPE**FTVEPDGLIT**TK*

>7000000558.a:SRS019125_WUGC_scaffold_8684

[Renamed SP-CHAP in this dissertation]

MSKKQEMIQQFIDKANSGDGVNDGAYGFQCADVPCYGLRHWYGVTLWGNAYDLLESARSQGLKVVDYAEYPKA
GWFFVKSYSVAGDGVNYGHTGLVYEDSDGSTIKTIEQNIDGNWDYLEVGGPCRYNERSVSEIVGYIVPPEEVEIG
WQQNQYGWVVREDGSYPTKEWEKINDVWYFFDDKGFMRSTWLNKYDAWYWFDDSGSMAIGWARINNAWYFFD
EEGKMVTGWIKHKLTYWYLLDGKEGTMVSNAFIQSADKKGWYLLKPDGTMEDKPE**FEIEPDGLIT**TK*

>7000000563.a:SRS014124_WUGC_scaffold_3907

MEIDTSRYREGLPQIGYAPYHQIHAHSTGNRNSTAQNEADYHMRPVESGFFSHVVGNGRVMQVGPVNQAYDV
GGGWNYYETAAVELIESHSTKEEFMEDYRLYIQLLRDLADEAGLPKTLDSDALEGIKSHDYCTNNQPNFSDHV
DPYPYLAKWGISREQFKRDIENGLEFKAGWQKSSGWYQNAAGSYPTNKWQQINGKWWFFNDEGYCLTNKWIK
RDGVWYWLNDGTMATGWKKINNEWYFFKQDGTMTGWVKYDKWYLLNTNNGFMESNQFVKGKDGWYLLSEDG
TMAEKPE**FTVEPDGLIT**AKEMNK*

>7000000571.a:SRS017713_Baylor_scaffold_29730

[Renamed SP-Amidase in this dissertation]

MVEIINKDIFQGIAGARPTEAPKYFILHNDAGSMSAKDYIGWLQSRYNNGQSELGFAHYIIDRNAIVRVENTYN
GSWSTANYEGNMNSIGYEVCCQFSTNDVDFRENEEMCLRQMAEDMKFYGVEPNYDTIKLHNEFSSTSCPARTLE
LHGNSNNSVRDFIIRRIKHYMSLGNTVQEMIDAQVQTSQGWVKDEKGWYRRKDGTPVDINELIGNEWQFQDK
DGYCLINKWYFDGTFYHYHDARGASIRSRWAQIDGWYFFNLSGEMQGTGWVKYKDHWFYCDAKNGDMKSDQYIK
SGNGWYLLKSNEMATKEA**FTIEPDGLIT**VK*

>7000000575.a:SRS017227_Baylor_scaffold_77430

MTVNIETAIAMRRARQGVSYSDDRNGPDSYDCSSSIYYALTSAGAASAGWAVNTEYEHDLKKNGYELISEN
TPWDAQRGDIIFIWGRRGYSSAGAGGHTGIFVDSNIIHCNRYRFDGITVNDHDDIWLYAGRPYYVYRLTNPSATA
EEVKTKGWQKNDTGYWVVRANGSYPKDQFEYIEENKSWFYFDSRGYMYSERWLKHTDGKWWYFDDKGYMATSWKK
INGKWWYFFNRDGSMTGWVKYEWYLLDAVNGDMKSNTFVYPYNGGYMMLLPDGRMADKEA**FNIEPDGLIT**TK*

>7000000576.a:SRS022621_Baylor_scaffold_75380

MAIKIENAIAMWEAREGKVSYNMEYRDGDDSYDCSSAMYSLRSAGASSAGWAVNTEYMHDLKKNGYELIAEN
TECNAQRGDIIFIWGRKGASAGAFGHTGMFIDSDNIIHCNAYNGISINNHDERWYAGQPYFYIYRLTNPNAPQ
EEVKKGWQKDDTGYWVVRANGSYPKSEFEYIEENKSWFYFDESGYMSDKWLHHTDGHWHYFDDKGYMATSWKK
INEKWWYFFNRDGAMQIGWVKWYEWYLLDQNGDMKSSTFVYPYNGGYLLLPDGRADKES**FTVEPDGLIT**AK*

>7000000618.a:SRS062544_LANL_scaffold_11414

MEIDTSRYREGLPQIGYAPYRQVHAHSTGNRNSTAQNEADYHMRPVESGFFSHVVGNGRVMQVGPVNQAYDV
GGGWNAAEGYGQVELIESHSTMEEFMTDYRLYVELLRNLADEAGIPKTLDSDDLEGIKTHYYCTYNQPNYS
DPYPYLAKWGISREQFKRDIHGLGEIKEGWQKNATGWYQSKDGSYPKDKWYINGVWYLLFDASGYCVLNKVV
KRADAWYWLDSGAMATGWKKINNEWYFFRADGAMVTGWVKYADEWYHLNTSNGFMESNAFVKGKDGWYLLNED
GTMAEKPE**FTVEPDGLIT**AKEVHR*

>7000000622.a:SRS077736_LANL_scaffold_33972

MAVNIETAIAMMRARQGNVTYSMDHRDGPYSYDCSSSVYYALREAGASSAGWAVNTEYEHDLIKNGYELIAEN
TECTAQRGDIIFIWGKRGASAGAFGHTGMFIDSUNIHCNYAYNGISVNEHDERWLYAGQPYFYIYRLANPNAQP
EEPKKGWQKDDNGYWYARVNGSYPKKEFEYIEENKSWFYFNAEGYMAEDWVKYTDGKWWFDDKGYMATSWKK
IGGAWYYFNRDGSMQTGWVKYYDKWYYLDATNGDMKSDCFVKYNDGWYLLLPDGRMADKPEFTVEPDGLITTK*

>7000000640.a:SRS015278_WUGC_scaffold_5514

MSKKQDMINDIMSHADAGTGVDYDGMGYQCADVTCYGIYEFYFIRLWGNADLLRSAESAGLQVVYGAQYPKA
GWFFVKNFVAGDGVNYGHTGLVYEDSDGSTIKTIEQNIDGNWDYLEVGGPCRYNERSVDSIVGYIVPPEEDES
WKHDDTGWWWRRKDGSYPTAKFEAIEGNWFYFNDNGYMYANQWLHHTDGYWYWFDDKGYMATSWKKINGKWWYF
NRDGSMTGTGWVKYYEKWYYLDAVNGDMKSNFTVPYNGGYMLLPDGRMADKEAFNIEPDGLITTK*

>7000000653.a:SRS016740_Baylor_scaffold_34762

MKKNDLFIDVSSHNGYDITGILYSMTTNTTIKISESTSYINPCLSSQTEQSTPVGFYHFAWFGGDSDEAEREA
RFFIDNVPQKVYLVLDYEDHASGDAQANTDACIRFMDVIKDAGYEPIYYSYKPFTLNVEYEQIIAKYPNSLW
IAGYGLNDGNADFEYFPSMEGIRWWQYSSNPFDKNIVLLDDEETEIDSGWKKNDKGYWYVRSDDGSYPKEQFEKI
DGTWYYFDGSGYMLADKWKRPDGTWYYFDKSGEMATGWKKIADKWWYFKEDGAMVTGWVKYKDTWYYLDGKEG
NMVSNSFVKSEKGWYYLKEDGTLSDKPEFTVEPDGLITAK*

>7000000653.a:SRS016740_Baylor_scaffold_71277

MTINIEAIAWMRDHEGQVYYSMEYRDGPDYDCSSSVYYALRNAGAVSAGWAVNTEYEHDLVKNGYELIAEN
QEWNAQRGDIIFIYGRRGYSAGAGGHTGMFVDSUNIHCNYARNGITVNNHDAIWNAAGQPYFYAYRLTNPNAQP
EEPKKGWQEDDNGYWYARSNGSYPKDQFEKIDGTWYYFDENGYMLSDKWKQRPDGTWYYFDKSGEMATDWNKID
GKWWYSSRDGAMIKGWVKYYDKWYYLDAQNGDMKSDCFVKYNDGWYLLLPDGRMADKPEFTVEPDGLITTK*

>7000000653.a:SRS016740_Baylor_scaffold_35094

MDIDKSRLRRGLPQVGVRPYRQVHAHSTGNRNSTAQNEADYHYRKDPELGFFSHVVGNGRVMQVGPVNNGSWDV
GGGWNAAESYAAVELIESHKTQEEFNRDYRLYVPLLRNLAEAGLPVTLDTDDLAIKTHEYCTNNQPNNSDHV
DPYPYLAKWGISRSQFKHDIEHGLVVEPGWKKNDTGWYVKEDGSYPKEQFEKIDGTWYYFDGSGYMLSDKWK
RPDGTWYYFDKSGEMATGWKLISNKWYFKEDGEMVTGWVKYQDTWYYLDGKEGNMVSNAFIQSADGTGWYYLK
EDGSMADKPEFTVEPDGLITTTIK*

>7000000665.a:SRS016037_WUGC_scaffold_22403

MTINIEAIAWMKARQGQVTYSMEYRDGDSSYDCSSSVYYALRSAGAVSAGWAVNTEYEHDLIKNGYELIAEN
TECNAQRGDIIFIWGKRGASAGAFGHTGMFIDSUNIHCNYAYNGISINNHERWYAGQPYFYIYRLTNPDAQP
EEVKKGWQKDDQGHYARANGSYPKAEFEYIEENKSWFYFDESGYAHTDKWLHHTDQWYYFDKDGYSMATSWKK
INEKWWYFNRDGAMQTGWVKYYDKWYYLDAVNGEMKSDCFVKYNDGWYLLLPDGRMADKPEFTVEPDGLITTK*

>7000000666.a:SRS022719_LANL_scaffold_9514

MEIDTSRYREGLPQIGYAPYHQIHAHSTGNRNSTAQNEADYHMRRPVESGFFSHVVGNGRVMQVGPVNQAYDV
GGGWNYYETAAVELIESHSTKEEFMEDYRLYVQLLRDLADEAGIPKTLTDLSEALEGIKSHDYCTNNQPNNFS
DPYPYLAKWGISREQFKHDVEHGLEFKEGWQKNSNGWYQNDAGSYPVNKWQKIGNVWYWFDDNSGYCYIERWLK
YNNQWYWLDDGSGAMVTGWKKISGSWYFYSKSDGSMATGWVKYYEKWYYLNTNNGFMESNAFVKGTGWYYISDDG
TMADKPDFTVEPDGLITVKNEKEKPTEESK*

>7000000666.a:SRS022719_LANL_scaffold_12357

MEIDTSRYREGLPQIGYAPYHQIHAHSTGNRNSTAQNEADYHMRRPVESGFFSHVVGNGRVMQVGPINQAYDV
GGGWNYYETAAVELIESHSTKEEFMEDYRLYIQLLRDLADEAGLPKTLTDLSSALEGIKSHEYCTYNQPNNFS
DPYPYLAKWGISREQFKHDIEHGLGEIKKGWQKDDTGWLYQNSDGSYPTNKWQQINGKWWYFFNEEGYCLTNKI
KRGTAWYWLDDGVMATGWKKINNEWYFKADEGEMATGWVKYYDKWYYLNTNNGFMESNAFVKGDGWYYISED
GTMAEKPEFTVEPDGLITAK*

>7000000666.a:SRS022719_LANL_scaffold_17772

MKKNDLFIDVSSHNGYDITGILYSMGTNTTIKISESTSYINPCLSSQTEQSTPVGFIYHFAWFGGDSDEAEREA
RFFIDNVFPQKVYLVLDYEDHASGDAQANTDACIRFMDVIKNAGYEPYIYYSYKPFTLSNVYYEQIIAKYPNSLW
IAGYGLNDGNADFEYFPSMEGIRWWQYSSNPFDKNIVLLDDEETEIDSGWKKNDTGFVHVRSDGSYPKEQFEKI
DGTWYYFDGSGYMLADKWKRPDGTWYYFDKSGEMATGWKKIADKWDYFDKDGAMVKGWVKYKDVWYFLDYQQG
AMVSNAFIQSADKKGWYYLKEDGTLADKLEFTVEPDGLITTK*

>7000000666.a:SRS022719_LANL_scaffold_32108

MGVNIESAIAWMQARKSRVSYSMYRDGEVSYDCSSAIYYALRSGGAVSAGWAVNTEYEHDLIKNGYELISEN
TPWDAKRGDIFIWGRRGYSSGAGGHTGMFVDSNIIHCNYAYDGITINDHDERWLAAGRPYFYVYRLTNANAQP
EEVKKGWQSDDKGDWYARANGTYPKSDFEYIEENKSWFYFDAKGYCVKSDWVFHTDGKWWYFEEDGYMVTSWRK
INNKKWYFNRDGSMTGWIKYYDKWYCDVNGDMKSDAFIKYNDGWYLLLPDGRLEKPAFTVEPDGLITITD
K*

>7000000671.a:SRS022143_LANL_scaffold_94835

MVKIINNTIFSGIAGARPTKPKYYIMHNDAGSMTAESYVDWLQSRDNGQSELGFAHYIITRDAIARVEDTYN
GTWSAANYDANMNSLSYEVCCQYNSTDAEFIEENMVLRLQMAEDMTYYGDTPNYSNIKFHNEFSSTSCPARSLK
LHGGYNDSLRDYVIAKIKHYQSLGSTVQEMLDNEGNQEGWKKNATGWVYVNSDGTYPKNKWQKINGVYYYFDQN
GYMKANTWYDHTDGYWYLLPSGGMATGWALIANKWYFKEDGAMATGWVKYKDHWYLLDAKDGMKSKQFIKS
ADGSGWYLLKEDGSMADKPEFTVEPDGLITTK*

>7000000681.a:SRS018439_Baylor_scaffold_21476

MVEIINTNVFNIGIAGTRPTHKPKYYILHNDAGSMTPEYVSWLRSRYNNGQSALGFAHYIINRDTIARVENTYS
GTWSAANYDANMNSISYEVCCQYNSSDAEFIEENMVLRLQMAEDMTYYGDTPNYSNIKFHNEFSSTSCPARSLA
LHGGYNDSLRDYVIAKIKHYQSLGSTVQEMLGAEAGWHNNSTGWVYVNEEDGTYPNKKWQVIDDVWYFFDSN
GYMKANTWHKHSNTWYLLPNGAMATGWVLDNSWYFDNQGAMVTGWVKYKETWYFLDYQKGSMSVNAFVKS
ADGTGWYLLKDDGSLADKPEFEVEPDGLITLAQAKAEK*

>7000000698.a:SRS024318_LANL_scaffold_2667

MKKNDLFIDVSSHNGYDITGILYSMGTNTTIKISESTSYINPCLSSQTEQSTPVGFIYHFAWFGGDSDEAEREA
RFFIDNVFPQKVYLVLDYEDHASGDAQANTDACIRFMDVIKNAGYEPYIYYSYKPFTLSNVYYEQIIAKYPNSLW
IAGYGLNDGNADFEYFPSMEGIRWWQYSSNPFDKNIVLLDDEETEIDSGWKKNDTGFVYVRSDGSYPKEQFEQI
DGTWYYFDGSGYMLADKWKRPDGTWYYFDKSGEMATGWKKIAEKWYFDKDGAMVKGWVKYKDKWYLLDAING
DMKSDCFVKYNDGWYLLLPDGTADKPEFTVEPDGLITTK*

>7000000713.a:SRS024355_LANL_scaffold_97309

MDIDTSRYREGLPQVGVQPYRQVHAHSTGNRNSTAQNEADYHYRKDPKLGFFSHVVGNGRVMQVGPVNNGSWDV
GGGWNAAESYAAVELIESHGSKEEFMRDYKLYIELLRNLADEADIPKTLDSGSLAGIKTHEYCTNNQPNHSDHV
DPYPYLAKWGISREQFKKDIEGGLSEGGWKRNDTGWWWEESDGSYPTKRWKKINNEWFYFDDRGYCLINRWFDN
GQDWFYLDKRGAMVTGWMYINNRRWYFFKSDGRMAKGWVKYRETWYLLDEKDGDMKSDQFVKYGNWYLLKPDGS
MADKPEFTVEPDGLITTK*

>7000000716.a:SRS044662_LANL_scaffold_3441

MVEIINTNVFNIGIAGARPTHKPKYYILHNDSGSMTPKSYIGWLRERYNNGQSELGFAHYIINRDTIARVENTYS
GTWSAANYDANMNSISYEVCCQYNASDAEFIEENMVLRLQMAEDMTYYGDTPNYSNIKFHNEFSSTSCPKRSLA
LHGGYNDSLRDYVIAKIKHYQSLGSTVQEMLGAEAGWHNNSTGWVYVNEEDGTYPNKKWQVIDDVWYFFDSN
GYMKANTWHKHSNTWYLLPNGAMAIGWVLDNSWYFDNQGAMVTGWVKYKDTWYLLDYQKGSMSVNAFIQS
ADGKGWYLLKDDGSLSDKPEFEVEPDGLITLAKPKEEK*

>7000000716.a:SRS044662_LANL_scaffold_35559

MVVNIETAIAMRAKQGNVTYSMDHRDGPYSYDCSSSVYYALREAGAVSAGWAVNTEYQHDWLLKNGYTLITEN
QPWDAKRGDIFIWGRRGYSSGAGGHTGMFVDSNIIHCNRYARNGITVNDHDDIWYSAGQPYFYAYRLTNPNQA
EEVKTGWQNNDKGYWYVRSNGSYPKDKFEKIDGTWYYFDESGYMLSDKWKRPDGTWYYFDKSGEMASDWKKIN
GKWWYFNRDGMATGWIKYYDNWYCDVNGDMKSDTFIKYNDGWYLLLPDGRLEKPAFTVEPDGLITITDK*

>7000000742.a:SRS015044_WUGC_scaffold_56267

MAVNIETAIAMMRARQGQVSYSMDDRNGPDSYDCSSSIYYALTSAGAVSAGWAVNTEYEHDLIKNGYELITEN
QPWDAERGDIFIWGRRGYSSGAGGHTGIFVDSNIIHCNYRFDGITVNDHDDIWLYAGRPIYYVYRLTNPSAAA
EEVKTGWQKDDTGYWFRANGTYPKDQFEYVEENKSWFYFDSRGYMYSERWLKHTDGKWWFDDKGYMATSWKK
INGKWWYFNDRGSMQTGWIKYYDNWYYCDATNGDMKSNTFVPYNGGYMLLPDGRMADKEA**FNIEPDGLIT**TK*

>Bacilli_gi_448879146-1659265-1699117|Streptococcus_sp._F0441

MSKKQDMINDLIAHADAGAGVDYDGMGYQCADVTCTYGIYKYFGIRLWGNADLLRSAEAAGLQVVYGAQHPKA
GWFFVKNFVAGDGVNYGHTGLVYEDSDGSTIKTIEQNIDGNWDYLEVGGPCRYNERSVDSIVGYIVPPEEDES
WKHDDTGWWWRRKDGSYPTAKFEAVDWNWFYFDENGYMYANQWLHHTDGYWYWFDDKGYMANTGWKKINGKWWY
FNADGAMQTGWVKYIEKWWYLNSENGDMVSNTFVPYNGGYLMLLEDGRLAEKES**FNIEPDGLIT**TK*

>Bacilli_gi_335029684-0-26362|Streptococcus_infantis_SK1076

MTINIEDAIAWMREREGQVYYSMEDRDGPDSYDCSSSVYYALRSAGASSAGWAVNTEYEHDLVNNGYELIAEN
QEWNARQGDIFIYGRRGYSAGAGGHTGMFVDSNIIHCNRYARNGITVNNHDAIWNAAGQPYFYAYRLTNPDQA
EEVKKGWQKDDDGWYARANGSYPKDQFEKIDGTWYFYFDENGYMLSDKWKQRLDGTWYFYFDESSEMATDWNKIN
EKWYFFNRDGAMYTGWVKYIDNWYFDAQNGEMKSDTFVRYNDGWYLLLPDGRMADKPE**FTVEPDGLIT**TK*

>Bacilli_gi_383939350|Streptococcus_pseudopneumoniae_SK674

MDIDTSRLRTGLPQVGYAPYRQVHAHSTGNRNSTVQNEADYHYRKNPELGFFSHVVGNGRVLQVGPVNNGAWDV
GGDWNAESYAAVELIESHSTKEEFMTDYRLYIELLRNLADEAGLPKTLDTSSLEGIKTHEYCTNNQPDNHS
DPYPYLAKWGISREQFKQDIESGLTIEVGWKKNDTGWYVHSDGSYPKDKFEKVNGTWYFYFDGSGYMLADR
HTDGNWYWFDDSGELATGWKKIAEKWYFDTEGAMKTGWVKYKDTWYLLDKEGAMVSNAFIQSADGTGWYLLK
PDGTMADKPE**FTVEPDGLIT**TK*

>Bacilli_gi_565858025-8013-50395|Streptococcus_pseudopneumoniae_5247

MDIDTSRLRTGLPQVGYAPYRQVHAHSTGNRNSTVQNEADYHYRKNPELGFFSHVVGNGRVMQVGPVNNGAWDV
GGGWNAESYAAVELIESHSTTEEFMADYRLYIELLRNLADEAGLPKTLDTSSLEGIKTHEYCTNNQPDNHS
DPYPYLAKWGISREQFKQDIESGLTIETGWQKNDTGWYVHSDGSYPKDKFEKINGTWYFYFDGSGYMLVNRWKK
HTDGNWYWFDDSGEMATGWKKIADKWYFDTEGAMKTGWVKYKDTWYLLDSKEGAMVSNAFVQSADGTGWYLLK
ADGTMADKPE**FAVEPDGLIT**TK*

>Bacilli_gi_565866318|Streptococcus_mitis_27/7

MDIDTSRLRTDLPIVGFEPFRQVHAHSTGNRNSTVQNEADYHYRKDPPELGFFSHVVGNGRVMQVGPVNNGSYDV
GGGWNAETYAAVELIESHSTKEEFMTDYRLYIELLRNLADEAGLPKTLDTDLVLAGIKTHEYCTNNQPDNSS
DPYPYLAKWGSREQFKRDIENGLGAETGWQKNDTGWYVYSDGSYPKNKFEKINGTWYFYFDGSGYMLADR
HTDGNWYFDRSGEMATGWKKIAEKWYFDVEGAMKTGWVKYKDTWYLLDSKDGNMVSNEFVRAGQGWYIKPD
GTMADKPE**FEIEPDGLIT**TK*

Appendix B: Supplementary data published in manuscripts

Table B-1. Primers used in Chapter 2

Primer Name	Sequence
Cpl1 W210A Fwd	5'- CAAGGGGTGGGCGTTCAGACGAA -3'
Cpl1 W210A Rev	5'- CTGTCTTGTTCACGTTCCAGC -3'
Cpl1 F218A Fwd	5'- CAATGGCAGTGCGCCTTATAATA -3'
Cpl1 F218A Rev	5'- TTTCGTCTGAACCACCACCCCTT -3'
Cpl1 Y238A Fwd	5'- TAGTAAAGGAGCGTGCTTAACGA -3'
Cpl1 Y238A Rev	5'- TCGAAGTAGTACCACACACCACC -3'
Cpl1 W210A:F218A Fwd	5'- CAATGGCAGTGCGCCTTATAATAAATGGGAAAAAATC -3'
Cpl1 W210A:F218A Rev	5'- TTTCGTCTGAACGCCACCCCTT -3'
Cpl1 W210A:Y238A Fwd	5'- CAAGGGGTGGGCGTTCAGACGAAAC -3'
Cpl1 W210A:Y238A Rev	5'- CTGTCTTGTTCACGTTTC -3'
Cpl1 F218A:Y238A Fwd	5'- TAGTAAAGGAGCGTGCTTAACGAGCGAATG -3'
Cpl1 F218A:Y238A Rev	5'- TCGAAGTAGTACCACACAC -3'
Cpl1 WFY Fwd	5'- TAGTAAAGGAGCGTGCTTAACGAGCGAATG -3'
Cpl1 WFY Rev	5'- TCGAAGTAGTACCACACAC -3'
Cpl1 WFY CBD Fwd	5'- TAAGAAGGAGATATACCATGGGAACGTGGAAACAAGACAG -3'
Cpl1 WFY CBD Rev	5'- CAGCTTCCTTTCGGGCTTTGTTAATGGTGATGATGATGATGTG -3'
Cpl1 DNP→AAA Fwd	5'- GGCGAGTTTCACGAAAGAACCAGAC -3'
Cpl1 DNP to AAA Rev	5'- GCCGCTGCAAGCTCTCCGTTTGT -3'
Cpl1 SFT→AAA Fwd	5'- GGCGAAAGAACCAGACGGGCTT-3'
Cpl1 SFT→AAA Rev	5'- GCCGCTGGATTGTCTGCAAGCTC -3'
Cpl1 KEP→AAA Fwd	5'- GGCGGACGGGCTTATAACCGTAG-3'
Cpl1 KEP→AAA Rev	5'- GCCGCCGTGAAACTTGGATTGTCTG -3'
Cpl1 DGL→AAA Fwd	5'- GGCGATAACCGTAGCACATCATC -3'
Cpl1 DGL→AAA Rev	5'- GCCGCTGGTTCTTTCGTGAAACTTG -3'
Cpl1 ITV→AAA Fwd	5'- GGCGGCACATCATCATCACC -3'
Cpl1 ITV→AAA Rev	5'- GCCGCAAGCCCGTCTGGTTCTTT-3'
Cpl1 F328A Fwd	5'- CAATCCAAGTGCGACGAAAGAACCAGAC -3'
Cpl1 F328A Rev	5'- TCTGCAAGCTCTCCGTTT -3'

Appendix C: Sequence Information for Constructs used in Dissertation-Related Studies

pBAD24::*Cpl-1*

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGCACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTTTCACGAAAGAACCAGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 40.01 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTNIIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPGDAQANTNACLRFMQMIADAGYKPIIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFKNIVLLDDEE
DDKPKTAGTWKQDSKGWWFRRNNGSFPYNKWEKIGGVWYFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSFTKEPDGLITVAHHHHHH*

pBAD24::*Cpl-1 W210A*

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGGCGTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTTTCACGAAAGAACCAGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.89 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTNIIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPSGDAQANTNACLRFMQMIADAGYKPIIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFDKNIVLLDDEE
DDKPKTAGTWKQDSKGWAFRRNNGSFPYNKWEKIGGVWYFYFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSFTKEPDGLITVAHHHHHH*

pBAD24::*Cpl-1 F218A*

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTGC GCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTTTCACGAAAGAACCAGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.93 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTTIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDP SGDAQANTNACLRFMQMIADAGYKPIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPF DKNIVLLDDEE
DDKPKTAGTWKQDSKGWWFRRNNGSAPYNKWEKIGGVWYFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSFTKEPDGLITVAHHHHHH*

pET28a::Cpl-1 Y238A

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGAGCGTGCTTAAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTTTCACGAAAGAACCAGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.92 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTNIIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPSGDAQANTNACLRFMQMIADAGYKPIIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFDKNIVLLDDEE
DDPKKTAGTWKQDSKGWWFRNNGSFPYNKWEKIGGVWYFDSKGACLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSFTKEPDGLITVAHHHHHH*

pET28a::Cpl-1 W210A F218A

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGGCGTTCAGACGAA
ACAATGGCAGTGC GCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTTTCACGAAAGAACCAGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.82 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTTIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPSGDAQANTNACLRFMQMIADAGYKPIYYSYK
PFTHDNVQYQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFDPKNIVLLDDEE
DDKPKTAGTWKQDSKGWAFRRNNGSAPYNKWEKIGGVWYFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSFTKEPDGLITVAHHHHHH*

pET28a::Cpl-1 F218A Y238A

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTGC GCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGAGCGTGCTTAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTTTCACGAAAGAACCAGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.84 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTTI IKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPSGDAQANTNACLRFMQMIADAGYKPIIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFDPKNIVLLDDEE
DDPKPTAGTWKQDSKGWWFRRNNGSAPYNKWEKIGGVWYFYFDSKGACLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSFTKEPDGLITVAHHHHHH*

pET28a::Cpl-1 W210A Y238A

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACATAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGGCGTTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGAGCGTGCTTAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTTTCACGAAAGAACCAGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.80 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTTI IKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPSGDAQANTNACLRFMQMIADAGYKPIIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFDPKNIVLLDDEE
DDPKPTAGTWKQDSKGWAFRRNNGSFPYNKWEKIGGVWYFYFDSKGACLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSFTKEPDGLITVAHHHHHH*

pET28a::Cpl-1 W210A F218A Y238A

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGGCGTTTCAGACGAA
ACAATGGCAGTGC GCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGAGCGTGCTTAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTTTCACGAAAGAACCAGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.73 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTNTIIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPSGDAQANTNACLRFMQMIADAGYKPIIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFDDKNIVLLDDEE
DDPKKTAGTWKQDSKGWAFRRNNGSAPYNKWEKIGGVWYFYFDSKGACLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSFTKEPDGLITVAHHHHHH*

pET28a::*Cpl-1 W210A F218A Y238A CBD*

Insert Sequence (444 nucleotides)

ATGGGAACGTGGAAACAAGACAGCAAGGGGTGGgcgTTCAGACGAAACAATGGCAGTgcgCCTT
ATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAGTAAAGGAgcgTGCTTAAC
GAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGACAACGGCGCAATGGCGACT
GGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTCAGGCGCTATGGTTACTGGTT
GGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGGTAACATGGTTTCTAATGA
ATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGAGAGCTTGCAGACAATCCA
AGTTTCACGAAAGAACCAGACGGGCTTATAACCGTAGCACATCATCATCACCATTAA

Protein Sequence (147 aa, 17.21 kDa)

MGTWKQDSKGWAFRRNNGSAPYNKWEKIGGVWYYFDSKGACLTSEWLKDNEKWYYLKDNGAMAT
GWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNGELADNP
SFTKEPDGLITVAHHHHHH*

pET28a::*Cpl-1 CBD*

Insert Sequence (459 nucleotides)

ATGAAACCGAAAACCGCAGGCACCTGGAAACAGGATAGCAAAGGTTGGTGGTTTCGTCGTAATA
ATGGTAGCTTTCCGTATAACAAATGGGAGAAAATTGGTGGCGTGTGGTATTATTTGACAGCAA
AGGCTATTGTCTGACCAGCGAATGGCTGAAAGATAACGAGAAATGGTATTACCTGAAAGACAAT
GGTGCAATGGCCACCGGTTGGGTTTTAGTTGGTAGCGAATGGTACTATATGGATGATAGCGGTG
CCATGGTTACCGGCTGGGTGAAATATAAGAATAACTGGTATTACATGACCAACGAGCGTGGCAA
TATGGTTAGCAACGAATTTATCAAAAGCGGCAAAGGCTGGTACTTTATGAATACCAATGGTGAA
CTGGCAGATAATCCGAGCTTTACCAAAGAACCGGATGGTCTGATTACCGTTGCACATCATCATC
ATCATCATTAATAA

Protein Sequence (151 aa, 17.89 kDa)

MKPKTAGTWKQDSKGWWFRRNNGSFPYNKWEKIGGVWYYFDSKGYCLTSEWLKDNEKWYYLKDN
GAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNGE
LADNPSFTKEPDGLITVAHHHHHH**

pBAD24::*Cpl-1* $\Delta C2$

Insert Sequence (870 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTAATAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTCATCATCATCACCATTAATAA

Protein Sequence (288 aa, 33.41 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTTNTIIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPGDAQANTNACLRFMQMIADAGYKPIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFDKNIVLLDDEE
DDKPKTAGTWKQDSKGWWFRNNGSFYPYKWEKIGGVWYFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTHHHHHH**

pBAD24::*Cpl-1* *Δtail*

Insert Sequence (990 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCACATCATCATCACCATTAA

Protein Sequence (329 aa, 38.32 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTTNTIIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPSGDAQANTNACLRFMQMIADAGYKPIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFKNIVLLDDEE
DDKPKTAGTWKQDSKGWWFRNNGSFYPYKWEKIGGVWYFFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELAHHHHHH*

pBAD24::*Cpl-1* DNP → AAA

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTGAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAgcggcgggcgAGTTTCACGAAAGAACCAGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.90 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTNIIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPGDAQANTNACLRFMQMIADAGYKPIYYSYK
PFTHDNVQYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFKNIVLLDDEE
DDKPKTAGTWKQDSKGWWFRRNNGSFYPNKWEKIGGVWYYFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELAAAASFKEPDGLITVAHHHHHH*

pBAD24::*Cpl-I SFT* → AAA

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTACG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAgcggcggcgAAAGAACCAGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.89 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTNIIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPSGDAQANTNACLRFMQMIADAGYKPIIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFDDKNIVLLDDEE
DDKPKTAGTWKQDSKGWWFRNNGSFPYNKWEKIGGVWYFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPAAAKEPDGLITVAHHHHHH*

pBAD24::*Cpl-1* *KEP* → *AAA*

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTGAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTTTCACGgcggcggcgcgGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.87 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTNIIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPGDAQANTNACLRFMQMIADAGYKPIIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFDKNIVLLDDEE
DDKPKTAGTWKQDSKGWWFRRNNGSFYPNKWEKIGGVWYFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSFTAAADGLITVAHHHHHH*

pBAD24::*Cpl-1 DGL* → *AAA*

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTTTCACGAAAGAACCAGcggcggcgcgATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.94 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTNIIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPSGDAQANTNACLRFMQMIADAGYKPIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFDKNIVLLDDEE
DDKPKTAGTWKQDSKGWWFRNNGSFPYNKWEKIGGVWYYFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSFTKEPAAAITVAHHHHHH*

pBAD24::*Cpl-1* ITV → AAA

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTTAAAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTTTCACGAAAGAACCAGACGGGCTTgcggcgggcgGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.91 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTNIIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPGDAQANTNACLRFMQMIADAGYKPIYYSYK
PFTHDNVDYQQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFDDKNIVLLDDEE
DDKPKTAGTWKQDSKGWWFRNNGSFPYNKWEKIGGVWYYFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSFTKEPDGLAAAAHHHHHH*

pBAD24::*Cpl-I F328A*

Insert Sequence (1038 nucleotides)

ATGGTTAAAAAGAATGATTTATTTGTAGATGTTTCAAGTCACAACGGTTACGATATAACAGGTA
TCTTGGAGCAAATGGGAACAACCTAACACCATCATTAATAATTTCTGAAAGTACGACCTATTTAAA
CCCTTGCTTGTCTGCTCAAGTGGAGCAGTCAAACCCTATTGGCTTTTATCACTTCGCACGCTTT
GGCGGAGACGTAGCAGAAGCCGAAAGAGAAGCGCAGTTTTTTCCTTGACAACGTGCCTATGCAAG
TTAAATACCTTGTATTGGACTACGAGGACGACCCAAGCGGAGACGCACAAGCGAACACTAACGC
ATGCTTACGCTTTATGCAGATGATTGCTGACGCTGGATATAAACCTATTTATTATAGTTATAAA
CCGTTTACACATGATAATGTGGACTATCAGCAAATCCTTGACACAGTTCCTAATTCTCTATGGA
TTGCAGGCTATGGCTTAAACGATGGTACAGCTAACTTTGAATACTTCCCAAGCATGGACGGGAT
AAGATGGTGGCAGTATTCTAGTAACCCGTTTGACAAGAATATTGTACTGTTAGACGATGAAGAA
GACGACAAGCCAAAGACCGCTGGAACGTGGAAACAAGACAGCAAGGGGTGGTGGTTCAGACGAA
ACAATGGCAGTTTCCCTTATAATAAATGGGAAAAAATCGGTGGTGTGTGGTACTACTTCGATAG
TAAAGGATATTGCTTAACGAGCGAATGGCTCAAAGATAATGAAAAATGGTACTACCTCAAGGAC
AACGGCGCAATGGCGACTGGTTGGGTGCTAGTCGGGTCAGAGTGGTATTATATGGACGATTTCAG
GCGCTATGGTTACTGGTTGGGTCAAGTATAAGAATAACTGGTACTATATGACAAATGAACGTGG
TAACATGGTTTTCTAATGAATTTATTAAGTCTGGAAAAGGTTGGTATTTTCATGAACACAAACGGA
GAGCTTGCAGACAATCCAAGTgcgACGAAAGAACCAGACGGGCTTATAACCGTAGCACATCATC
ATCATCACCATTAA

Protein Sequence (345 aa, 39.93 kDa)

MVKKNDLFVDVSSHNGYDITGILEQMGTNTTIKISESTTYLNPCLSAQVEQSNPIGFYHFARF
GGDVAEAEREAQFFLDNVPMQVKYLVLDYEDDPGDAQANTNACLRFMQMIADAGYKPIYYSYK
PFTHDNVQYQILAQFPNSLWIAGYGLNDGTANFEYFPSMDGIRWWQYSSNPFKNIVLLDDEE
DDKPKTAGTWKQDSKGWWFRRNNGSFPYNKWEKIGGVWYFDSKGYCLTSEWLKDNEKWYYLKD
NGAMATGWVLVGSEWYYMDDSGAMVTGWVKYKNNWYYMTNERGNMVSNEFIKSGKGWYFMNTNG
ELADNPSATKEPDGLITVAHHHHHH*

pBAD24::*Pal*

Insert Sequence (909 nucleotides)

ATGGGAGTCGATATTGAAAAAGGCGTTGCGTGGATGCAGGCCCGAAAGGGTCGAGTATCTTATA
GCATGGACTTTTCGAGACGGTCCTGATAGCTATGACTGCTCAAGTTCTATGTACTATGCTCTCCG
CTCAGCCGGAGCTTCAAGTGCTGGATGGGCAGTCAATACTGAGTACATGCACGCATGGCTTATT
GAAAACGGTTATGAACTAATTAGTGAAAATGCTCCGTGGGATGCTAAACGAGGCGACATCTTCA
TCTGGGGACGCAAAGGTGCTAGCGCAGGCGCTGGAGGTCATACAGGGATGTTTATTGACAGTGA
TAACATCATTTCACTGCAACTACGCCTACGACGGAATTTCCGTCAACGACCACGATGAGCGTTGG
TACTATGCAGGTCAACCTTACTACTACGTCTATCGCTTGACTAACGCAAATGCTCAACCGGCTG
AGAAGAACTTGGCTGGCAGAAAGATGCTACTGGTTTCTGGTACGCTCGAGCAAACGGAACCTTA
TCCAAAAGATGAGTTCGAGTATATCGAAGAAAACAAGTCTTGGTTCTACTTTGACGACCAAGGC
TACATGCTCGCTGAGAAATGGTTGAAACATACTGATGGAAATTGGTATTGGTTCGACCGTGACG
GATACATGGCTACGTCATGGAAACGGATTGGCGAGTCATGGTACTACTTCAATCGCGATGGTTC
AATGGTAACCGGTTGGATTAAGTATTACGATAATTGGTATTATTGTGATGCTACCAACGGCGAC
ATGAAATCGAATGCGTTTATCCGTTATAACGACGGCTGGTATCTACTATTACCGGACGGACGTC
TGGCAGATAAACCTCAATTCACCGTAGAGCCGGACGGGCTCATTACTGCTAAAGTTCATCATCA
TCATCACCATTAA

Protein Sequence (302 aa, 35.27 kDa)

MGVDIEKGVAWMQARKGRVSYSMDFRDGPDSYDCSSSMYYALRSAGASSAGWAVNTEYMHAWLI
ENGYELISENAPWDAKRGDIFIWGRKGASAGAGGHTGMFIDSDNI IHCNYAYDGI SVNDHDERW
YYAGQPYYYVYRLTNANAQPAEKKLGWQKDATGFWYARANGTYPKDEFEYIEENKSWFYFDDQG
YMLAEKWLKHTDGNWYWFDRDGYMATSWKRIGESWYYFNRDGSMVTGWIKYYDNWYYCDATNGD
MKSNAFIRYNDGWYLLLPDGRLADKPQFTVEPDGLITAKVHHHHHH*

pET28a::Pal F284A

Insert Sequence (909 nucleotides)

ATGGGAGTCGATATTGAAAAAGGCGTTGCGTGGATGCAGGCCCGAAAGGGTCGAGTATCTTATA
GCATGGACTTTTCGAGACGGTCCTGATAGCTATGACTGCTCAAGTTCTATGTACTATGCTCTCCG
CTCAGCCGGAGCTTCAAGTGCTGGATGGGCAGTCAATACTGAGTACATGCACGCATGGCTTATT
GAAAACGGTTATGAACTAATTAGTGAAAATGCTCCGTGGGATGCTAAACGAGGCGACATCTTCA
TCTGGGGACGCAAAGGTGCTAGCGCAGGCGCTGGAGGTCATACAGGGATGTTTCATTGACAGTGA
TAACATCATTCACTGCAACTACGCCTACGACGGAATTTCCGTCAACGACCACGATGAGCGTTGG
TACTATGCAGGTCAACCTTACTACTACGTCTATCGCTTGACTAACGCAAATGCTCAACCGGCTG
AGAAGAACTTGGCTGGCAGAAAGATGCTACTGGTTTCTGGTACGCTCGAGCAAACGGAACCTTA
TCCAAAAGATGAGTTTCGAGTATATCGAAGAAAACAAGTCTTGGTTCTACTTTGACGACCAAGGC
TACATGCTCGCTGAGAAATGGTTGAAACATACTGATGGAATTTGGTATTGGTTCGACCGTGACG
GATACATGGCTACGTCATGGAAACGGATTGGCGAGTCATGGTACTACTTCAATCGCGATGGTTC
AATGGTAACCGGTTGGATTAAGTATTACGATAATTGGTATTATTGTGATGCTACCAACGGCGAC
ATGAAATCGAATGCGTTTATCCGTTATAACGACGGCTGGTATCTACTATTACCGGACGGACGTC
TGGCAGATAAACCTCAA_{gcg}ACCGTAGAGCCGGACGGGCTCATTAAGTTTCATCATCA
TCATCACCATTAA

Protein Sequence (302 aa, 35.20 kDa)

MGVDIEKGVAWMQARKGRVSYSMDFRDGPDSYDCSSSMYYALRSAGASSAGWAVNTEYMHAWLI
ENGYELISENAPWDAKRGDIFIWGRKGASAGAGGHTGMFIDSDNIIHCNYAYDGISVNDHDERW
YYAGQPYYYVYRLTNANAQPAEKKLGWQKDATGFWYARANGTYPKDEFEYIEENKSWFYFDDQG
YMLAEKWLKHTDGNWYWFDRDGYMATSWKRIGESWYYFNRDGSMVTGWIKYYDNWYYCDATNGD
MKSNAFIRYNDGWYLLLPDGR_{LADKPQATVEPDGLITAKV}HHHHHH*

Appendix D: List of Published, Submitted and Planned Co-authored Manuscripts

Understanding the Molecular Basis for Homodimer Formation of the Pneumococcal Endolysin Cpl-1

Adit B. Alreja, Sara B. Linden, Harrison R. Lee, Kinlin L. Chao, Osnat Herzberg, and Daniel C. Nelson*



Cite This: <https://doi.org/10.1021/acsinfecdis.2c00627>



Read Online

ACCESS |



Metrics & More



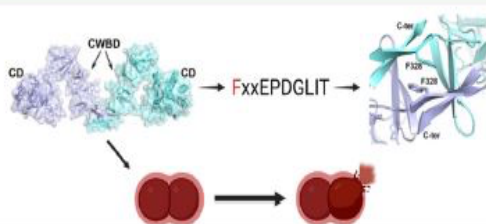
Article Recommendations



Supporting Information

ABSTRACT: The rise of multi-drug-resistant bacteria that cannot be treated with traditional antibiotics has prompted the search for alternatives to combat bacterial infections. Endolysins, which are bacteriophage-derived peptidoglycan hydrolases, are attractive tools in this fight. Several studies have already demonstrated the efficacy of endolysins in targeting bacterial infections. Endolysins encoded by bacteriophages that infect Gram-positive bacteria typically possess an N-terminal catalytic domain and a C-terminal cell-wall binding domain (CWBD). In this study, we have uncovered the molecular mechanisms that underlie formation of a homodimer of Cpl-1, an endolysin that targets *Streptococcus pneumoniae*. Here, we use site-directed mutagenesis, analytical size exclusion chromatography, and analytical ultracentrifugation to disprove a previous suggestion that three residues at the N-terminus of the CWBD are involved in the formation of a Cpl-1 dimer in the presence of choline in solution. We conclusively show that the C-terminal tail region of Cpl-1 is involved in formation of the dimer. Alanine scanning mutagenesis generated various tail mutant constructs that allowed identification of key residues that mediate Cpl-1 dimer formation. Finally, our results allowed identification of a consensus sequence (FxxEPDGLIT) required for choline-dependent dimer formation—a sequence that occurs frequently in pneumococcal autolysins and endolysins. These findings shed light on the mechanisms of Cpl-1 and related enzymes and can be used to inform future engineering efforts for their therapeutic development against *S. pneumoniae*.

KEYWORDS: endolysin, Cpl-1, pneumococcus, bacteriophage, antibiotic resistance, cell wall hydrolase





Application of bacteriophage-derived endolysins to combat streptococcal disease: current State and perspectives

Sara B Linden¹, Adit B Alreja¹ and Daniel C Nelson^{1,2}



The decline in new antibiotic candidates combined with an increase in antibiotic-resistance necessitates development of alternative antimicrobials. Bacteriophage-encoded endolysins (lysins) are a class of peptidoglycan hydrolases that have been proposed to fill this antimicrobial void. The past 20 years has seen a dramatic expansion of studies on endolysin discovery, structure/function, engineering, immunogenicity, toxicity/safety, and efficacy in animal models. These collective efforts have led to current human clinical trials on at least three different endolysins that are antimicrobial toward staphylococcal species. It can be anticipated that endolysins targeting streptococcal species may be next in line for translational development. Notably, streptococcal diseases largely manifest at accessible mucous membranes, which should be beneficial for protein therapeutics. Additionally, there are a number of well-identified streptococcal diseases in both humans and animals that are associated with a single species, further favoring a targeted endolysin therapeutic.

holins form a lesion on the bacterial membrane through which phage-encoded peptidoglycan hydrolases, known as endolysins, access and lyse the peptidoglycan (i.e. cell wall), resulting in immediate cellular death and release of viral particles [2].

In the past 20 years, scientists have started to evaluate the antimicrobial properties of endolysins *vis-à-vis* their ability to rapidly degrade the bacterial peptidoglycan. Significantly, when applied externally to susceptible Gram-positive bacterial cells in the absence of holins or phage, the endolysins alone are sufficient to produce rapid bacterial lysis. Although the endolysins do not intrinsically affect the subjacent bacterial membrane, the high internal turgor pressure of the bacterial cytoplasm combined with the loss of structural integrity of the protective peptidoglycan layer is enough to cause osmotic lysis of the membrane resulting in cellular death. Exploiting this phenomenon is the basis for studying and developing



Contents lists available at ScienceDirect

International Journal of Pharmaceutics

journal homepage: www.elsevier.com/locate/ijpharm



Bacteriophage endolysin powders for inhaled delivery against pulmonary infections

Yuncheng Wang^a, Dipesh Khanal^a, Adit B. Alreja^b, Hang Yang^c, Rachel YK Chang^a,
Waiting Tai^a, Mengyu Li^a, Daniel C. Nelson^b, Warwick J Britton^d, Hak-Kim Chan^{a,*}

^a Advanced Drug Delivery Group, School of Pharmacy, University of Sydney, Sydney, NSW, Australia

^b Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, MD, USA

^c CAS Key Laboratory of Special Pathogens and Biosafety, Center for Emerging Infectious Diseases, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, China

^d Centenary Institute and Faculty of Medicine and Health, University of Sydney, Sydney, NSW, Australia

ARTICLE INFO

Keywords:

Bacteriophage endolysin
Spray drying
Pulmonary infection
Streptococcus pneumoniae
Cpl-1

ABSTRACT

Endolysins are bacteriophage-encoded enzymatic proteins that have great potential to treat multidrug-resistant bacterial infections. Bacteriophage endolysins Cpl-1 and ClyJ-3 have shown promising antimicrobial activity against *Streptococcus pneumoniae*, which causes pneumonia in humans. This is the first study to investigate the feasibility of spray-dried endolysins Cpl-1 and ClyJ-3 with excipients to produce inhalable powders. The two endolysins were individually tested with leucine and sugar (lactose or trehalose) for spray drying method followed by characterization of biological and physico-chemical properties. A complete loss of ClyJ-3 bioactivity was observed after atomization of the liquid feed solution (before the drying process), while Cpl-1 maintained its bioactivity in the spray-dried powders. Cpl-1 formulations containing leucine with lactose or trehalose showed promising physico-chemical properties (particle size, crystallinity, hygroscopicity, etc.) and aerosol performances (fine particle fraction values above 65%). The results indicated that endolysin Cpl-1 can be formulated as spray dried powders suitable for inhaled delivery to the lungs for the potential treatment of pulmonary infections.



OPEN **Structure of *Escherichia coli* O157:H7 bacteriophage CBA120 tailspike protein 4 baseplate anchor and tailspike assembly domains (TSP4-N)**

Kinlin L. Chao¹, Xiaoran Shang¹, Julia Greenfield^{1,2}, Sara B. Linden¹, Adit B. Alreja¹, Daniel C. Nelson^{1,3} & Osnat Herzberg^{1,2}

Four tailspike proteins (TSP1-4) of *Escherichia coli* O157:H7 bacteriophage CBA120 enable infection of multiple hosts. They form a branched complex that attaches to the tail baseplate. Each TSP recognizes a different lipopolysaccharide on the membrane of a different bacterial host. The 335 N-terminal residues of TSP4 promote the assembly of the TSP complex and anchor it to the tail baseplate. The crystal structure of TSP4-N₃₃₅ reveals a trimeric protein comprising four domains. The baseplate anchor domain (AD) contains an intertwined triple-stranded β -helix. The ensuing XD1, XD2 and XD3 β -sheet containing domains mediate the binding of TSP1-3 to TSP4. Each of the XD domains adopts the same fold as the respective XD domains of bacteriophage T4 gp10 baseplate protein, known to engage in protein-protein interactions via its XD2 and XD3 domains. The structural similarity suggests that XD2 and XD3 of TSP4 also function in protein-protein interactions. Analytical ultracentrifugation analyses of TSP4-N₃₃₅ and of domain deletion proteins showed how TSP4-N₃₃₅ promotes the formation of the TSP quaternary complex. TSP1 and TSP2 bind directly to TSP4 whereas TSP3 binding requires a pre-formed TSP4-N₃₃₅:TSP2 complex. A 3-dimensional model of the bacteriophage CBA120 TSP complex has been developed based on the structural and ultracentrifuge information.

A unique borrelial protein facilitates microbial immune evasion

Shelby D. Foor,¹ Kalvis Brangulis,^{2,3} Anil K. Shakya,⁴ Vipin S. Rana,¹ Sandhya Bista,¹ Chrysoula Kitsou,¹ Michael Ronzetti,^{1,5} Adit B. Alreja,⁴ Sara B. Linden,⁴ Amanda S. Altieri,⁴ Bolormaa Baljinnyam,⁵ Inara Akopjana,² Daniel C. Nelson,^{1,4} Anton Simeonov,⁵ Osnat Herzberg,^{4,6} Melissa J. Caimano,⁷ Utpal Pal^{1,8}

AUTHOR AFFILIATIONS See affiliation list on p. 21.

ABSTRACT *Borrelia burgdorferi*, the pathogen of Lyme disease, encodes many conserved proteins of unknown structure or function, including ones that serve essential roles in microbial infectivity. One such protein is BB0238, which folds into a two-domain protein, as we have determined by X-ray crystallography and AlphaFold analysis. The N-terminal domain begins with a helix-turn-helix motif (HTH), previously referred to as a tetratricopeptide repeat (TPR) motif, known to mediate protein-protein interactions. The fold of the C-terminal domain has been seen in proteins with a range of unrelated activities and thus does not infer function. In addition to its previously known binding partner BB0323, another essential borrelial virulence determinant, we show that BB0238 also binds a second protein, BB0108, a borrelial ortholog of the chaperone protein SurA and the peptidyl-prolyl *cis/trans* isomerase protein PrsA. An *in vitro* enzymatic assay confirmed the catalytic activity. We also determined the crystal structure of the catalytic domain of BB0108, which revealed the parvulin-type organization of the key catalytic residues. We show that BB0238 influences the proteolytic processing of BB0323, although the TPR/HTH motif is not involved in the process. Instead, we show that the motif stabilizes BB0238 in the host environment and facilitates tick-to-mouse pathogen transmission by aiding spirochete evasion of early host cellular immunity. Taken together, these studies highlight the biological significance of BB0238 and its interactions with multiple *B. burgdorferi* proteins essential for microbial infection.

- **Alreja, A. B.**, Appel, A. E., Zhu, J. C., Riley, S. P., Gonzalez-Juarbe, N., & Nelson, D. C. (2023). SP-CHAP, an endolysin with enhanced activity against biofilm pneumococci and nasopharyngeal colonization. Submitted for publication.
- Shang, X., Kelman, Z., **Alreja, A. B.**, Schiel, J. E., Gonzalez-Juarbe, N., Schmelcher, M., Etobayeya, I., Nelson, D. C. (2023). A novel design to exploit the synergy of PlyC catalytic domains. In preparation.
- Invention disclosure: “A Cysteine, Histidine-dependent Aminopeptidase/Hydrolase Bacteriolytic Against *Streptococcus pneumoniae*”. Invention disclosure no. LS-2022-049.

Bibliography

- Abaev, I., Foster-Frey, J., Korobova, O., Shishkova, N., Kiseleva, N., Kopylov, P., . . . Donovan, D. M. (2013). Staphylococcal phage 2638A endolysin is lytic for *Staphylococcus aureus* and harbors an inter-lytic-domain secondary translational start site. *Appl Microbiol Biotechnol*, 97(8), 3449-3456. doi:10.1007/s00253-012-4252-4
- Abdullahi, O., Karani, A., Tigoi, C. C., Mugo, D., Kungu, S., Wanjiru, E., . . . Scott, J. A. (2012). The prevalence and risk factors for pneumococcal colonization of the nasopharynx among children in Kilifi District, Kenya. *PLoS One*, 7(2), e30787. doi:10.1371/journal.pone.0030787
- Aceil, J., & Avci, F. Y. (2022). Pneumococcal surface proteins as virulence factors, immunogens, and conserved vaccine targets. *Front Cell Infect Microbiol*, 12, 832254. doi:10.3389/fcimb.2022.832254
- Ajuebor, J., McAuliffe, O., O'Mahony, J., Ross, R. P., Hill, C., & Coffey, A. (2016). Bacteriophage endolysins and their applications. *Sci Prog*, 99(2), 183-199. doi:10.3184/003685016X14627913637705
- Aldridge, S. P., J.; Sturchio, J.L. (1999). The discovery and development of penicillin: 1928-1945. In T. R. S. o. Chemistry (Ed.). The Alexander Fleming Laboratory Museum, London, UK.
- Alm, R. A., & Gallant, K. (2020). Innovation in antimicrobial resistance: the CARB-X perspective. *ACS Infect Dis*, 6(6), 1317-1322. doi:10.1021/acsinfecdis.0c00026
- Alreja, A. B., Linden, S. B., Lee, H. R., Chao, K. L., Herzberg, O., & Nelson, D. C. (2023). Understanding the molecular basis for homodimer formation of the pneumococcal endolysin Cpl-1. *ACS Infect Dis*, 9(5), 1092-1104. doi:10.1021/acsinfecdis.2c00627
- Arnoldi, M., Fritz, M., Bauerlein, E., Radmacher, M., Sackmann, E., & Boulbitch, A. (2000). Bacterial turgor pressure can be measured by atomic force microscopy. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*, 62(1 Pt B), 1034-1044. doi:10.1103/physreve.62.1034
- Austrian, R. (1981). Pneumococcus: the first one hundred years. *Rev Infect Dis*, 3(2), 183-189. doi:10.1093/clinids/3.2.183
- Avery, O. T., Macleod, C. M., & McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types : induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type Iii. *J Exp Med*, 79(2), 137-158. doi:10.1084/jem.79.2.137
- Becker, S. C., Dong, S., Baker, J. R., Foster-Frey, J., Pritchard, D. G., & Donovan, D. M. (2009). LysK CHAP endopeptidase domain is required for lysis of live staphylococcal cells. *FEMS Microbiol Lett*, 294(1), 52-60. doi:10.1111/j.1574-6968.2009.01541.x
- Becker, S. C., Foster-Frey, J., & Donovan, D. M. (2008). The phage K lytic enzyme LysK and lysostaphin act synergistically to kill MRSA. *FEMS Microbiol Lett*, 287(2), 185-191. doi:10.1111/j.1574-6968.2008.01308.x
- Becker, S. C., Foster-Frey, J., Stodola, A. J., Anacker, D., & Donovan, D. M. (2009). Differentially conserved staphylococcal SH3b_5 cell wall binding domains confer increased staphylolytic and streptolytic activity to a streptococcal prophage endolysin domain. *Gene*, 443(1-2), 32-41. doi:10.1016/j.gene.2009.04.023

- Becker, S. C., Roach, D. R., Chauhan, V. S., Shen, Y., Foster-Frey, J., Powell, A. M., . . . Donovan, D. M. (2016). Triple-acting lytic enzyme treatment of drug-resistant and intracellular *Staphylococcus aureus*. *Sci Rep*, 6, 25063. doi:10.1038/srep25063
- Bhosale, S., Deen, M. C., Proceviat, C., Hettle, A., Winter, D. K., Brockerman, J., . . . Vocado, D. J. (2023). A fluorogenic disaccharide substrate for alpha-mannosidases enables high-throughput screening and identification of an inhibitor of the GH92 virulence factor from *Streptococcus pneumoniae*. *ACS Chem Biol*, 18(8), 1730-1737. doi:10.1021/acscchembio.3c00053
- Billal, D. S., Feng, J., Leprohon, P., Legare, D., & Ouellette, M. (2011). Whole genome analysis of linezolid resistance in *Streptococcus pneumoniae* reveals resistance and compensatory mutations. *BMC Genomics*, 12, 512. doi:10.1186/1471-2164-12-512
- Binte Muhammad Jai, H. S., Dam, L. C., Tay, L. S., Koh, J. J. W., Loo, H. L., Kline, K. A., & Goh, B. C. (2020). Engineered lysins with customized lytic activities against Enterococci and Staphylococci. *Front Microbiol*, 11, 574739. doi:10.3389/fmicb.2020.574739
- Blanchette-Cain, K., Hinojosa, C. A., Akula Suresh Babu, R., Lizcano, A., Gonzalez-Juarbe, N., Munoz-Almagro, C., . . . Orihuela, C. J. (2013). *Streptococcus pneumoniae* biofilm formation is strain dependent, multifactorial, and associated with reduced invasiveness and immunoreactivity during colonization. *mBio*, 4(5), e00745-00713. doi:10.1128/mBio.00745-13
- Blazquez, B., Fresco-Taboada, A., Iglesias-Bexiga, M., Menendez, M., & Garcia, P. (2016). PL3 amidase, a tailor-made lysin constructed by domain shuffling with potent killing activity against pneumococci and related species. *Front Microbiol*, 7, 1156. doi:10.3389/fmicb.2016.01156
- Borysowski, J., Weber-Dabrowska, B., & Gorski, A. (2006). Bacteriophage endolysins as a novel class of antibacterial agents. *Exp Biol Med (Maywood)*, 231(4), 366-377. doi:10.1177/153537020623100402
- Braun, J. S., Sublett, J. E., Freyer, D., Mitchell, T. J., Cleveland, J. L., Tuomanen, E. I., & Weber, J. R. (2002). Pneumococcal pneumolysin and H(2)O(2) mediate brain cell apoptosis during meningitis. *J Clin Invest*, 109(1), 19-27. doi:10.1172/JCI12035
- Broendum, S. S., Buckle, A. M., & McGowan, S. (2018). Catalytic diversity and cell wall binding repeats in the phage-encoded endolysins. *Mol Microbiol*, 110(6), 879-896. doi:10.1111/mmi.14134
- Broendum, S. S., Williams, D. E., Hayes, B. K., Kraus, F., Fodor, J., Clifton, B. E., . . . McGowan, S. (2021). High avidity drives the interaction between the streptococcal C1 phage endolysin, PlyC, with the cell surface carbohydrates of Group A *Streptococcus*. *Mol Microbiol*, 116(2), 397-415. doi:10.1111/mmi.14719
- Brooks, L. R. K., & Mias, G. I. (2018). *Streptococcus pneumoniae's* virulence and host immunity: aging, diagnostics, and prevention. *Front Immunol*, 9, 1366. doi:10.3389/fimmu.2018.01366
- Brown, E. D., & Wright, G. D. (2016). Antibacterial drug discovery in the resistance era. *Nature*, 529(7586), 336-343. doi:10.1038/nature17042
- Brown, P. H., & Schuck, P. (2006). Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys J*, 90(12), 4651-4661. doi:10.1529/biophysj.106.081372
- Brundish, D. E., & Baddiley, J. (1968). Pneumococcal C-substance, a ribitol teichoic acid containing choline phosphate. *Biochem J*, 110(3), 573-582. doi:10.1042/bj1100573

- Buey, R. M., Monterroso, B., Menendez, M., Diakun, G., Chacon, P., Hermoso, J. A., & Diaz, J. F. (2007). Insights into molecular plasticity of choline binding proteins (pneumococcal surface proteins) by SAXS. *J Mol Biol*, 365(2), 411-424. doi:10.1016/j.jmb.2006.09.091
- Bush, N. G., Diez-Santos, I., Abbott, L. R., & Maxwell, A. (2020). Quinolones: mechanism, lethality and their contributions to antibiotic resistance. *Molecules*, 25(23). doi:10.3390/molecules25235662
- Bustamante, N., Iglesias-Bexiga, M., Bernardo-Garcia, N., Silva-Martin, N., Garcia, G., Campanero-Rhodes, M. A., . . . Menendez, M. (2017). Deciphering how Cpl-7 cell wall-binding repeats recognize the bacterial peptidoglycan. *Sci Rep*, 7(1), 16494. doi:10.1038/s41598-017-16392-4
- Bustamante, N., Rico-Lastres, P., Garcia, E., Garcia, P., & Menendez, M. (2012). Thermal stability of Cpl-7 endolysin from the *streptococcus pneumoniae* bacteriophage Cp-7; cell wall-targeting of its CW_7 motifs. *PLoS One*, 7(10), e46654. doi:10.1371/journal.pone.0046654
- Butler, J. C., Hofmann, J., Cetron, M. S., Elliott, J. A., Facklam, R. R., & Breiman, R. F. (1996). The continued emergence of drug-resistant *Streptococcus pneumoniae* in the United States: an update from the Centers for Disease Control and Prevention's pneumococcal sentinel surveillance system. *J Infect Dis*, 174(5), 986-993. doi:10.1093/infdis/174.5.986
- Caren J. Stark, R. P. B., James T. Hoopes, Daniel C. Nelson. (2010). *Bacteriophage lytic enzymes as antimicrobials*. Washington DC: ASM Press.
- Cassini, A., Hogberg, L. D., Plachouras, D., Quattrocchi, A., Hoxha, A., Simonsen, G. S., . . . Burden of, A. M. R. C. G. (2019). Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the european economic area in 2015: a population-level modelling analysis. *Lancet Infect Dis*, 19(1), 56-66. doi:10.1016/S1473-3099(18)30605-4
- Catalao, M. J., Gil, F., Moniz-Pereira, J., & Pimentel, M. (2011). The endolysin-binding domain encompasses the N-terminal region of the mycobacteriophage Ms6 Gp1 chaperone. *J Bacteriol*, 193(18), 5002-5006. doi:10.1128/JB.00380-11
- CDC. (2019). *Antibiotic resistance threats in the United States*. Retrieved from Atlanta, GA: <https://www.cdc.gov/drugresistance/pdf/threats-report/2019-ar-threats-report-508.pdf>
- CDC. (2022). Global pneumococcal disease and vaccination. Retrieved from <https://www.cdc.gov/pneumococcal/global.html>
- Celia, L. K., Nelson, D., & Kerr, D. E. (2008). Characterization of a bacteriophage lysin (Ply700) from *Streptococcus uberis*. *Vet Microbiol*, 130(1-2), 107-117. doi:10.1016/j.vetmic.2007.12.004
- Chang, B., Nariai, A., Sekizuka, T., Akeda, Y., Kuroda, M., Oishi, K., & Ohnishi, M. (2015). Capsule switching and antimicrobial resistance acquired during repeated *Streptococcus pneumoniae* pneumonia episodes. *J Clin Microbiol*, 53(10), 3318-3324. doi:10.1128/JCM.01222-15
- Cheng, Q., Nelson, D., Zhu, S., & Fischetti, V. A. (2005). Removal of group B streptococci colonizing the vagina and oropharynx of mice with a bacteriophage lytic enzyme. *Antimicrob Agents Chemother*, 49(1), 111-117. doi:10.1128/AAC.49.1.111-117.2005
- Cillóniz C., G.-V. C., Ceccato A., Torres A. (2018). Antimicrobial resistance among *Streptococcus pneumoniae*. In *Antimicrobial Resistance in the 21st Century. Emerging Infectious Diseases of the 21st Century*: Springer, Cham.

- Cockeran, R., Durandt, C., Feldman, C., Mitchell, T. J., & Anderson, R. (2002). Pneumolysin activates the synthesis and release of interleukin-8 by human neutrophils in vitro. *J Infect Dis*, 186(4), 562-565. doi:10.1086/341563
- Collaborators, G. B. D. A. R. (2022). Global mortality associated with 33 bacterial pathogens in 2019: a systematic analysis for the global burden of disease study 2019. *Lancet*, 400(10369), 2221-2248. doi:10.1016/S0140-6736(22)02185-7
- Corsini, B., Diez-Martinez, R., Aguinagalde, L., Gonzalez-Camacho, F., Garcia-Fernandez, E., Letrado, P., . . . Yuste, J. (2018). Chemotherapy with phage lysins reduces pneumococcal colonization of the respiratory tract. *Antimicrob Agents Chemother*, 62(6). doi:10.1128/AAC.02212-17
- Cundell, D. R., Gerard, N. P., Gerard, C., Idanpaan-Heikkila, I., & Tuomanen, E. I. (1995). *Streptococcus pneumoniae* anchor to activated human cells by the receptor for platelet-activating factor. *Nature*, 377(6548), 435-438. doi:10.1038/377435a0
- Daniel, A., Euler, C., Collin, M., Chahales, P., Gorelick, K. J., & Fischetti, V. A. (2010). Synergism between a novel chimeric lysin and oxacillin protects against infection by methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother*, 54(4), 1603-1612. doi:10.1128/AAC.01625-09
- Danis-Wlodarczyk, K. M., Wozniak, D. J., & Abedon, S. T. (2021). Treating bacterial infections with bacteriophage-based enzybiotics: *in vitro*, *in vivo* and clinical application. *Antibiotics (Basel)*, 10(12). doi:10.3390/antibiotics10121497
- Darby, E. M., Trampari, E., Siasat, P., Gaya, M. S., Alav, I., Webber, M. A., & Blair, J. M. A. (2023). Molecular mechanisms of antibiotic resistance revisited. *Nat Rev Microbiol*, 21(5), 280-295. doi:10.1038/s41579-022-00820-y
- Davies, N. G., Flasche, S., Jit, M., & Atkins, K. E. (2021). Modeling the effect of vaccination on selection for antibiotic resistance in *Streptococcus pneumoniae*. *Sci Transl Med*, 13(606). doi:10.1126/scitranslmed.aaz8690
- Dehart, H. P., Heath, H. E., Heath, L. S., Leblanc, P. A., & Sloan, G. L. (1995). The lysostaphin endopeptidase resistance gene (epr) specifies modification of peptidoglycan cross bridges in *Staphylococcus simulans* and *Staphylococcus aureus*. *Appl Environ Microbiol*, 61(7), 2811. doi:10.1128/aem.61.7.2811-2811.1995
- Denapate, D., Bruckner, R., Hakenbeck, R., & Vollmer, W. (2012). Biosynthesis of teichoic acids in *Streptococcus pneumoniae* and closely related species: lessons from genomes. *Microb Drug Resist*, 18(3), 344-358. doi:10.1089/mdr.2012.0026
- Diez-Martinez, R., de Paz, H. D., Bustamante, N., Garcia, E., Menendez, M., & Garcia, P. (2013). Improving the lethal effect of cpl-7, a pneumococcal phage lysozyme with broad bactericidal activity, by inverting the net charge of its cell wall-binding module. *Antimicrob Agents Chemother*, 57(11), 5355-5365. doi:10.1128/AAC.01372-13
- Diez-Martinez, R., De Paz, H. D., Garcia-Fernandez, E., Bustamante, N., Euler, C. W., Fischetti, V. A., . . . Garcia, P. (2015). A novel chimeric phage lysin with high in vitro and in vivo bactericidal activity against *Streptococcus pneumoniae*. *J Antimicrob Chemother*, 70(6), 1763-1773. doi:10.1093/jac/dkv038
- Disease, G. B. D., Injury, I., & Prevalence, C. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the global burden of disease study 2017. *Lancet*, 392(10159), 1789-1858. doi:10.1016/S0140-6736(18)32279-7

- Djurkovic, S., Loeffler, J. M., & Fischetti, V. A. (2005). Synergistic killing of *Streptococcus pneumoniae* with the bacteriophage lytic enzyme Cpl-1 and penicillin or gentamicin depends on the level of penicillin resistance. *Antimicrob Agents Chemother*, 49(3), 1225-1228. doi:10.1128/AAC.49.3.1225-1228.2005
- Doehn, J. M., Fischer, K., Reppe, K., Gutbier, B., Tschernig, T., Hocke, A. C., . . . Witzentrath, M. (2013). Delivery of the endolysin Cpl-1 by inhalation rescues mice with fatal pneumococcal pneumonia. *J Antimicrob Chemother*, 68(9), 2111-2117. doi:10.1093/jac/dkt131
- Doern, G. V., Richter, S. S., Miller, A., Miller, N., Rice, C., Heilmann, K., & Beekmann, S. (2005). Antimicrobial resistance among *Streptococcus pneumoniae* in the United States: have we begun to turn the corner on resistance to certain antimicrobial classes? *Clin Infect Dis*, 41(2), 139-148. doi:10.1086/430906
- Doi, Y., Wachino, J. I., & Arakawa, Y. (2016). Aminoglycoside resistance: the emergence of acquired 16S ribosomal RNA methyltransferases. *Infect Dis Clin North Am*, 30(2), 523-537. doi:10.1016/j.idc.2016.02.011
- Domenech, M., Garcia, E., & Moscoso, M. (2011). In vitro destruction of *Streptococcus pneumoniae* biofilms with bacterial and phage peptidoglycan hydrolases. *Antimicrob Agents Chemother*, 55(9), 4144-4148. doi:10.1128/AAC.00492-11
- Donovan, D. M., Dong, S., Garrett, W., Rousseau, G. M., Moineau, S., & Pritchard, D. G. (2006). Peptidoglycan hydrolase fusions maintain their parental specificities. *Appl Environ Microbiol*, 72(4), 2988-2996. doi:10.1128/AEM.72.4.2988-2996.2006
- Donovan, D. M., Lardeo, M., & Foster-Frey, J. (2006). Lysis of staphylococcal mastitis pathogens by bacteriophage phi11 endolysin. *FEMS Microbiol Lett*, 265(1), 133-139. doi:10.1111/j.1574-6968.2006.00483.x
- Dunne, M., Leicht, S., Krichel, B., Mertens, H. D., Thompson, A., Krijgsveld, J., . . . Meijers, R. (2016). Crystal structure of the CTP1L endolysin reveals how its activity is regulated by a secondary translation product. *J Biol Chem*, 291(10), 4882-4893. doi:10.1074/jbc.M115.671172
- Durando, P., Faust, S. N., Fletcher, M., Krizova, P., Torres, A., & Welte, T. (2013). Experience with pneumococcal polysaccharide conjugate vaccine (conjugated to CRM197 carrier protein) in children and adults. *Clin Microbiol Infect*, 19 Suppl 1, 1-9. doi:10.1111/1469-0691.12320
- During, K. (1993). Can lysozymes mediate antibacterial resistance in plants? *Plant Mol Biol*, 23(1), 209-214. doi:10.1007/BF00021432
- Dye, B., Thornton-Evans, G., Li, X., & Iafolla, T. (2015). Dental caries and tooth loss in adults in the United States, 2011-2012. *NCHS Data Brief*(197), 197. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/25973996>
- Dzuvor, C. K. O., Shanbhag, B. K., Younas, T., Shen, H. H., Haritos, V. S., & He, L. (2022). Engineering self-assembled endolysin nanoparticles against antibiotic-resistant bacteria. *ACS Appl Bio Mater*. doi:10.1021/acsabm.2c00741
- Entenza, J. M., Loeffler, J. M., Grandgirard, D., Fischetti, V. A., & Moreillon, P. (2005). Therapeutic effects of bacteriophage Cpl-1 lysin against *Streptococcus pneumoniae* endocarditis in rats. *Antimicrob Agents Chemother*, 49(11), 4789-4792. doi:10.1128/AAC.49.11.4789-4792.2005

- Farrand, A. J., LaChapelle, S., Hotze, E. M., Johnson, A. E., & Tweten, R. K. (2010). Only two amino acids are essential for cytolytic toxin recognition of cholesterol at the membrane surface. *Proc Natl Acad Sci U S A*, 107(9), 4341-4346. doi:10.1073/pnas.0911581107
- Feldman, C., & Anderson, R. (2021). The role of co-infections and secondary infections in patients with COVID-19. *Pneumonia (Nathan)*, 13(1), 5. doi:10.1186/s41479-021-00083-w
- Fernandez-Ruiz, I., Coutinho, F. H., & Rodriguez-Valera, F. (2018). Thousands of novel endolysins discovered in uncultured phage genomes. *Front Microbiol*, 9, 1033. doi:10.3389/fmicb.2018.01033
- Fernandez-Tornero, C., Garcia, E., Lopez, R., Gimenez-Gallego, G., & Romero, A. (2002). Two new crystal forms of the choline-binding domain of the major pneumococcal autolysin: insights into the dynamics of the active homodimer. *J Mol Biol*, 321(1), 163-173. doi:10.1016/s0022-2836(02)00596-x
- Finn, R. D., Bateman, A., Clements, J., Coghill, P., Eberhardt, R. Y., Eddy, S. R., . . . Punta, M. (2014). Pfam: the protein families database. *Nucleic Acids Res*, 42(Database issue), D222-230. doi:10.1093/nar/gkt1223
- Finn, R. D., Clements, J., & Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res*, 39(Web Server issue), W29-37. doi:10.1093/nar/gkr367
- Fischer, W. (2000). Phosphocholine of pneumococcal teichoic acids: role in bacterial physiology and pneumococcal infection. *Res Microbiol*, 151(6), 421-427. doi:10.1016/s0923-2508(00)00174-1
- Fischer, W., Behr, T., Hartmann, R., Peter-Katalinic, J., & Egge, H. (1993). Teichoic acid and lipoteichoic acid of *Streptococcus pneumoniae* possess identical chain structures. A reinvestigation of teichoid acid (C polysaccharide). *Eur J Biochem*, 215(3), 851-857. doi:10.1111/j.1432-1033.1993.tb18102.x
- Fischetti, V. A. (2003). Novel method to control pathogenic bacteria on human mucous membranes. *Ann N Y Acad Sci*, 987, 207-214. doi:10.1111/j.1749-6632.2003.tb06050.x
- Fischetti, V. A. (2005). Bacteriophage lytic enzymes: novel anti-infectives. *Trends Microbiol*, 13(10), 491-496. doi:10.1016/j.tim.2005.08.007
- Fischetti, V. A. (2006). Using phage lytic enzymes to control pathogenic bacteria. *BMC Oral Health*, 6 Suppl 1, S16. doi:10.1186/1472-6831-6-S1-S16
- Fischetti, V. A. (2008). Bacteriophage lysins as effective antibacterials. *Curr Opin Microbiol*, 11(5), 393-400. doi:10.1016/j.mib.2008.09.012
- Fischetti, V. A. (2010). Bacteriophage endolysins: a novel anti-infective to control Gram-positive pathogens. *Int J Med Microbiol*, 300(6), 357-362. doi:10.1016/j.ijmm.2010.04.002
- Gaeng, S., Scherer, S., Neve, H., & Loessner, M. J. (2000). Gene cloning and expression and secretion of *Listeria monocytogenes* bacteriophage-lytic enzymes in *Lactococcus lactis*. *Appl Environ Microbiol*, 66(7), 2951-2958. doi:10.1128/AEM.66.7.2951-2958.2000
- Gagetti, P., Lo, S. W., Hawkins, P. A., Gladstone, R. A., Regueira, M., Faccone, D., . . . Corso, A. (2021). Population genetic structure, serotype distribution and antibiotic resistance of *Streptococcus pneumoniae* causing invasive disease in children in Argentina. *Microb Genom*, 7(9). doi:10.1099/mgen.0.000636
- Gallego-Paramo, C., Hernandez-Ortiz, N., Buey, R. M., Rico-Lastres, P., Garcia, G., Diaz, J. F., . . . Menendez, M. (2021). Structural and functional insights Into Skl and Pal endolysins, two cysteine-amidases with anti-pneumococcal activity. Dithiothreitol (DTT) effect on lytic activity. *Front Microbiol*, 12, 740914. doi:10.3389/fmicb.2021.740914

- Ganaie, F., Saad, J. S., McGee, L., van Tonder, A. J., Bentley, S. D., Lo, S. W., . . . Nahm, M. H. (2020). A new pneumococcal capsule type, 10D, is the 100th serotype and has a large cps fragment from an oral streptococcus. *mBio*, 11(3). doi:10.1128/mBio.00937-20
- Ganguly, J., Low, L. Y., Kamal, N., Saile, E., Forsberg, L. S., Gutierrez-Sanchez, G., . . . Kannenberg, E. L. (2013). The secondary cell wall polysaccharide of *Bacillus anthracis* provides the specific binding ligand for the C-terminal cell wall-binding domain of two phage endolysins, PlyL and PlyG. *Glycobiology*, 23(7), 820-832. doi:10.1093/glycob/cwt019
- Garcia, E., Garcia, J. L., Garcia, P., Arraras, A., Sanchez-Puelles, J. M., & Lopez, R. (1988). Molecular evolution of lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. *Proc Natl Acad Sci U S A*, 85(3), 914-918. doi:10.1073/pnas.85.3.914
- Garcia, J. L., Diaz, E., Romero, A., & Garcia, P. (1994). Carboxy-terminal deletion analysis of the major pneumococcal autolysin. *J Bacteriol*, 176(13), 4066-4072. doi:10.1128/jb.176.13.4066-4072.1994
- Garcia, J. L., Garcia, E., Arraras, A., Garcia, P., Ronda, C., & Lopez, R. (1987). Cloning, purification, and biochemical characterization of the pneumococcal bacteriophage Cp-1 lysin. *J Virol*, 61(8), 2573-2580. doi:10.1128/JVI.61.8.2573-2580.1987
- Garcia, P., Lopez, R., Ronda, C., Garcia, E., & Tomasz, A. (1983). Mechanism of phage-induced lysis in pneumococci. *J Gen Microbiol*, 129(2), 479-487. doi:10.1099/00221287-129-2-479
- Garcia, P., Martinez, B., Rodriguez, L., & Rodriguez, A. (2010). Synergy between the phage endolysin LysH5 and nisin to kill *Staphylococcus aureus* in pasteurized milk. *Int J Food Microbiol*, 141(3), 151-155. doi:10.1016/j.ijfoodmicro.2010.04.029
- Garcia, P., Mendez, E., Garcia, E., Ronda, C., & Lopez, R. (1984). Biochemical characterization of a murein hydrolase induced by bacteriophage Dp-1 in *Streptococcus pneumoniae*: comparative study between bacteriophage-associated lysin and the host amidase. *J Bacteriol*, 159(2), 793-796. doi:10.1128/jb.159.2.793-796.1984
- Gisch, N., Kohler, T., Ulmer, A. J., Muthing, J., Pribyl, T., Fischer, K., . . . Zahringer, U. (2013). Structural reevaluation of *Streptococcus pneumoniae* Lipoteichoic acid and new insights into its immunostimulatory potency. *J Biol Chem*, 288(22), 15654-15667. doi:10.1074/jbc.M112.446963
- Gonzales, J., Chakraborty, T., Romero, M., Mraheil, M. A., Kutlar, A., Pace, B., & Lucas, R. (2021). *Streptococcus pneumoniae* and its virulence factors H(2)O(2) and pneumolysin are potent mediators of the acute chest syndrome in sickle cell disease. *Toxins (Basel)*, 13(2). doi:10.3390/toxins13020157
- Graham, C. J. (2017). The global threat of antibiotic resistance: what can be done? *Journal of Global Health Reports*, 1(2).
- Grandgirard, D., Loeffler, J. M., Fischetti, V. A., & Leib, S. L. (2008). Phage lytic enzyme Cpl-1 for antibacterial therapy in experimental pneumococcal meningitis. *J Infect Dis*, 197(11), 1519-1522. doi:10.1086/587942
- Guariglia-Oropeza, V., & Helmann, J. D. (2011). *Bacillus subtilis* sigma(V) confers lysozyme resistance by activation of two cell wall modification pathways, peptidoglycan O-acetylation and D-alanylation of teichoic acids. *J Bacteriol*, 193(22), 6223-6232. doi:10.1128/JB.06023-11

- Hackel, M., Lascols, C., Bouchillon, S., Hilton, B., Morgenstern, D., & Purdy, J. (2013). Serotype prevalence and antibiotic resistance in *Streptococcus pneumoniae* clinical isolates among global populations. *Vaccine*, *31*(42), 4881-4887. doi:10.1016/j.vaccine.2013.07.054
- Hammond, R. W., Swift, S. M., Foster-Frey, J. A., Kovalskaya, N. Y., & Donovan, D. M. (2019). Optimized production of a biologically active *Clostridium perfringens* glycosyl hydrolase phage endolysin PlyCP41 in plants using virus-based systemic expression. *BMC Biotechnol*, *19*(1), 101. doi:10.1186/s12896-019-0594-7
- Harhala, M., Nelson, D. C., Miernikiewicz, P., Heselpoth, R. D., Brzezicka, B., Majewska, J., . . . Dabrowska, K. (2018). Safety studies of pneumococcal endolysins Cpl-1 and Pal. *Viruses*, *10*(11). doi:10.3390/v10110638
- Harhala, M. A., Gembara, K., Rybicka, I., Kazmierczak, Z. M., Miernikiewicz, P., Majewska, J. M., . . . Dabrowska, K. (2023). Immunogenic epitope scanning in bacteriolytic enzymes Pal and Cpl-1 and engineering Pal to escape antibody responses. *Front Immunol*, *14*, 1075774. doi:10.3389/fimmu.2023.1075774
- Hermoso, J. A., Garcia, J. L., & Garcia, P. (2007). Taking aim on bacterial pathogens: from phage therapy to enzybiotics. *Curr Opin Microbiol*, *10*(5), 461-472. doi:10.1016/j.mib.2007.08.002
- Hermoso, J. A., Monterroso, B., Albert, A., Galan, B., Ahrazem, O., Garcia, P., . . . Menendez, M. (2003). Structural basis for selective recognition of pneumococcal cell wall by modular endolysin from phage Cp-1. *Structure*, *11*(10), 1239-1249. doi:10.1016/j.str.2003.09.005
- Hirst, R. A., Rutman, A., Sikand, K., Andrew, P. W., Mitchell, T. J., & O'Callaghan, C. (2000). Effect of pneumolysin on rat brain ciliary function: comparison of brain slices with cultured ependymal cells. *Pediatr Res*, *47*(3), 381-384. doi:10.1203/00006450-200003000-00016
- Hong, H. W., Kim, Y. D., Jang, J., Kim, M. S., Song, M., & Myung, H. (2022). Combination effect of engineered endolysin EC340 with antibiotics. *Front Microbiol*, *13*, 821936. doi:10.3389/fmicb.2022.821936
- Hoopes, J. T., Stark, C. J., Kim, H. A., Sussman, D. J., Donovan, D. M., & Nelson, D. C. (2009). Use of a bacteriophage lysin, PlyC, as an enzyme disinfectant against *Streptococcus equi*. *Appl Environ Microbiol*, *75*(5), 1388-1394. doi:10.1128/AEM.02195-08
- Howard, L. V., & Gooder, H. (1974). Specificity of the autolysin of *Streptococcus (Diplococcus) pneumoniae*. *J Bacteriol*, *117*(2), 796-804. doi:10.1128/jb.117.2.796-804.1974
- Huang, Y., Yang, H., Yu, J., & Wei, H. (2015). Molecular dissection of phage lysin PlySs2: integrity of the catalytic and cell wall binding domains is essential for its broad lytic activity. *Viol Sin*, *30*(1), 45-51. doi:10.1007/s12250-014-3535-6
- Jado, I., Lopez, R., Garcia, E., Fenoll, A., Casal, J., Garcia, P., & Spanish Pneumococcal Infection Study, N. (2003). Phage lytic enzymes as therapy for antibiotic-resistant *Streptococcus pneumoniae* infection in a murine sepsis model. *J Antimicrob Chemother*, *52*(6), 967-973. doi:10.1093/jac/dkg485
- Jain, S., Self, W. H., Wunderink, R. G., & Team, C. E. S. (2015). Community-acquired pneumonia requiring hospitalization. *N Engl J Med*, *373*(24), 2382. doi:10.1056/NEJMc1511751
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583-589. doi:10.1038/s41586-021-03819-2
- Jun, S. Y., Jang, I. J., Yoon, S., Jang, K., Yu, K. S., Cho, J. Y., . . . Kang, S. H. (2017). Pharmacokinetics and tolerance of the phage endolysin-based candidate drug SAL200 after

- a single intravenous administration among healthy volunteers. *Antimicrob Agents Chemother*, 61(6). doi:10.1128/AAC.02629-16
- Jun, S. Y., Jung, G. M., Yoon, S. J., Choi, Y. J., Koh, W. S., Moon, K. S., & Kang, S. H. (2014). Preclinical safety evaluation of intravenously administered SAL200 containing the recombinant phage endolysin SAL-1 as a pharmaceutical ingredient. *Antimicrob Agents Chemother*, 58(4), 2084-2088. doi:10.1128/AAC.02232-13
- Jun, S. Y., Jung, G. M., Yoon, S. J., Youm, S. Y., Han, H. Y., Lee, J. H., & Kang, S. H. (2016). Pharmacokinetics of the phage endolysin-based candidate drug SAL200 in monkeys and its appropriate intravenous dosing period. *Clin Exp Pharmacol Physiol*, 43(10), 1013-1016. doi:10.1111/1440-1681.12613
- Kadioglu, A., Weiser, J. N., Paton, J. C., & Andrew, P. W. (2008). The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol*, 6(4), 288-301. doi:10.1038/nrmicro1871
- Kapoor, G., Saigal, S., & Elongavan, A. (2017). Action and resistance mechanisms of antibiotics: A guide for clinicians. *J Anaesthesiol Clin Pharmacol*, 33(3), 300-305. doi:10.4103/joacp.JOACP_349_15
- Karlsson, C., Jansson, P. E., & Skov Sorensen, U. B. (1999). The pneumococcal common antigen C-polysaccharide occurs in different forms. Mono-substituted or di-substituted with phosphocholine. *Eur J Biochem*, 265(3), 1091-1097. doi:10.1046/j.1432-1327.1999.00835.x
- Kelleher, D. J., Kreibich, G., & Gilmore, R. (1992). Oligosaccharyltransferase activity is associated with a protein complex composed of ribophorins I and II and a 48 kd protein. *Cell*, 69(1), 55-65. doi:10.1016/0092-8674(92)90118-v
- Kharat, A. S., & Tomasz, A. (2006). Drastic reduction in the virulence of *Streptococcus pneumoniae* expressing type 2 capsular polysaccharide but lacking choline residues in the cell wall. *Mol Microbiol*, 60(1), 93-107. doi:10.1111/j.1365-2958.2006.05082.x
- Khoury, G. A., Baliban, R. C., & Floudas, C. A. (2011). Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci Rep*, 1. doi:10.1038/srep00090
- Kietzman, C. C., Gao, G., Mann, B., Myers, L., & Tuomanen, E. I. (2016). Dynamic capsule restructuring by the main pneumococcal autolysin LytA in response to the epithelium. *Nat Commun*, 7, 10859. doi:10.1038/ncomms10859
- Kikhney, A. G., & Svergun, D. I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett*, 589(19 Pt A), 2570-2577. doi:10.1016/j.febslet.2015.08.027
- Kim, G. L., Seon, S. H., & Rhee, D. K. (2017). Pneumonia and *Streptococcus pneumoniae* vaccine. *Arch Pharm Res*, 40(8), 885-893. doi:10.1007/s12272-017-0933-y
- Kohanski, M. A., Dwyer, D. J., & Collins, J. J. (2010). How antibiotics kill bacteria: from targets to networks. *Nat Rev Microbiol*, 8(6), 423-435. doi:10.1038/nrmicro2333
- Kong, M., Sim, J., Kang, T., Nguyen, H. H., Park, H. K., Chung, B. H., & Ryu, S. (2015). A novel and highly specific phage endolysin cell wall binding domain for detection of *Bacillus cereus*. *Eur Biophys J*, 44(6), 437-446. doi:10.1007/s00249-015-1044-7
- Kretzer, J. W., Lehmann, R., Schmelcher, M., Banz, M., Kim, K. P., Korn, C., & Loessner, M. J. (2007). Use of high-affinity cell wall-binding domains of bacteriophage endolysins for immobilization and separation of bacterial cells. *Appl Environ Microbiol*, 73(6), 1992-2000. doi:10.1128/AEM.02402-06

- Kusuma, C. M., & Kokai-Kun, J. F. (2005). Comparison of four methods for determining lysostaphin susceptibility of various strains of *Staphylococcus aureus*. *Antimicrob Agents Chemother*, 49(8), 3256-3263. doi:10.1128/AAC.49.8.3256-3263.2005
- Labrie, S. J., Samson, J. E., & Moineau, S. (2010). Bacteriophage resistance mechanisms. *Nat Rev Microbiol*, 8(5), 317-327. doi:10.1038/nrmicro2315
- Lane, S., Hilliam, Y., & Bomberger, J. M. (2023). Microbial and immune regulation of the gut-lung axis during viral-bacterial coinfection. *J Bacteriol*, 205(1), e0029522. doi:10.1128/jb.00295-22
- Lewis, K. (2008). Multidrug tolerance of biofilms and persister cells. *Curr Top Microbiol Immunol*, 322, 107-131. doi:10.1007/978-3-540-75418-3_6
- Li, L., Ma, J., Yu, Z., Li, M., Zhang, W., & Sun, H. (2023). Epidemiological characteristics and antibiotic resistance mechanisms of *Streptococcus pneumoniae*: an updated review. *Microbiol Res*, 266, 127221. doi:10.1016/j.micres.2022.127221
- Li, Q., Cheng, W., Morlot, C., Bai, X. H., Jiang, Y. L., Wang, W., . . . Zhou, C. Z. (2015). Full-length structure of the major autolysin LytA. *Acta Crystallogr D Biol Crystallogr*, 71(Pt 6), 1373-1381. doi:10.1107/S1399004715007403
- Li, X., Lin, H., & Wang, J. (2023). Preparation and properties of hydrogels loaded with phage endolysin Lys84. *Sheng Wu Gong Cheng Xue Bao*, 39(9), 3876-3886. doi:10.13345/j.cjb.220925
- Li, Y., Weinberger, D. M., Thompson, C. M., Trzcinski, K., & Lipsitch, M. (2013). Surface charge of *Streptococcus pneumoniae* predicts serotype distribution. *Infect Immun*, 81(12), 4519-4524. doi:10.1128/IAI.00724-13
- Linden, S. B., Alreja, A. B., & Nelson, D. C. (2021). Application of bacteriophage-derived endolysins to combat streptococcal disease: current state and perspectives. *Curr Opin Biotechnol*, 68, 213-220. doi:10.1016/j.copbio.2021.01.012
- Loeffler, J. M., Djurkovic, S., & Fischetti, V. A. (2003). Phage lytic enzyme Cpl-1 as a novel antimicrobial for pneumococcal bacteremia. *Infect Immun*, 71(11), 6199-6204. doi:10.1128/IAI.71.11.6199-6204.2003
- Loeffler, J. M., & Fischetti, V. A. (2003). Synergistic lethal effect of a combination of phage lytic enzymes with different activities on penicillin-sensitive and -resistant *Streptococcus pneumoniae* strains. *Antimicrob Agents Chemother*, 47(1), 375-377. doi:10.1128/AAC.47.1.375-377.2003
- Loeffler, J. M., Nelson, D., & Fischetti, V. A. (2001). Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. *Science*, 294(5549), 2170-2172. doi:10.1126/science.1066869
- Loessner, M. J. (2005). Bacteriophage endolysins--current state of research and applications. *Curr Opin Microbiol*, 8(4), 480-487. doi:10.1016/j.mib.2005.06.002
- Loessner, M. J., Kramer, K., Ebel, F., & Scherer, S. (2002). C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. *Mol Microbiol*, 44(2), 335-349. doi:10.1046/j.1365-2958.2002.02889.x
- Lopez, R., & Garcia, E. (2004). Recent trends on the molecular biology of pneumococcal capsules, lytic enzymes, and bacteriophage. *FEMS Microbiol Rev*, 28(5), 553-580. doi:10.1016/j.femsre.2004.05.002

- Lou, H., Chen, M., Black, S. S., Bushell, S. R., Ceccarelli, M., Mach, T., . . . Naismith, J. H. (2011). Altered antibiotic transport in OmpC mutants isolated from a series of clinical strains of multi-drug resistant *E. coli*. *PLoS One*, 6(10), e25825. doi:10.1371/journal.pone.0025825
- Low, L. Y., Yang, C., Perego, M., Osterman, A., & Liddington, R. (2011). Role of net charge on catalytic domain and influence of cell wall binding domain on bactericidal activity, specificity, and host range of phage lysins. *J Biol Chem*, 286(39), 34391-34403. doi:10.1074/jbc.M111.244160
- Luo, D., Huang, L., Gondil, V. S., Zhou, W., Yang, W., Jia, M., . . . Wei, H. (2020). A choline-recognizing monomeric lysin, ClyJ-3m, shows elevated activity against *Streptococcus pneumoniae*. *Antimicrob Agents Chemother*, 64(12). doi:10.1128/AAC.00311-20
- Maestro, B., & Sanz, J. M. (2016). Choline binding proteins from *Streptococcus pneumoniae*: a dual role as enzybiotics and targets for the design of new antimicrobials. *Antibiotics (Basel)*, 5(2). doi:10.3390/antibiotics5020021
- Marks, L. R., Clementi, E. A., & Hakansson, A. P. (2012). The human milk protein-lipid complex HAMLET sensitizes bacterial pathogens to traditional antimicrobial agents. *PLoS One*, 7(8), e43514. doi:10.1371/journal.pone.0043514
- Marquart, M. E. (2021). Pathogenicity and virulence of *Streptococcus pneumoniae*: cutting to the chase on proteases. *Virulence*, 12(1), 766-787. doi:10.1080/21505594.2021.1889812
- Matthews, B. W., & Remington, S. J. (1974). The three dimensional structure of the lysozyme from bacteriophage T4. *Proc Natl Acad Sci U S A*, 71(10), 4178-4182. doi:10.1073/pnas.71.10.4178
- McCullers, J. A., Karlstrom, A., Iverson, A. R., Loeffler, J. M., & Fischetti, V. A. (2007). Novel strategy to prevent otitis media caused by colonizing *Streptococcus pneumoniae*. *PLoS Pathog*, 3(3), e28. doi:10.1371/journal.ppat.0030028
- Miller, W. R., Munita, J. M., & Arias, C. A. (2014). Mechanisms of antibiotic resistance in enterococci. *Expert Rev Anti Infect Ther*, 12(10), 1221-1236. doi:10.1586/14787210.2014.956092
- Mirdita, M., Schutze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nat Methods*, 19(6), 679-682. doi:10.1038/s41592-022-01488-1
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., . . . Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Res*, 49(D1), D412-D419. doi:10.1093/nar/gkaa913
- Mo, K. F., Li, X., Li, H., Low, L. Y., Quinn, C. P., & Boons, G. J. (2012). Endolysins of *Bacillus anthracis* bacteriophages recognize unique carbohydrate epitopes of vegetative cell wall polysaccharides with high affinity and selectivity. *J Am Chem Soc*, 134(37), 15556-15562. doi:10.1021/ja3069962
- Monterroso, B., Saiz, J. L., Garcia, P., Garcia, J. L., & Menendez, M. (2008). Insights into the structure-function relationships of pneumococcal cell wall lysozymes, LytC and Cpl-1. *J Biol Chem*, 283(42), 28618-28628. doi:10.1074/jbc.M802808200
- Morens, D. M., Taubenberger, J. K., & Fauci, A. S. (2008). Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: implications for pandemic influenza preparedness. *J Infect Dis*, 198(7), 962-970. doi:10.1086/591708
- Morzywolek, A., Plotka, M., Kaczorowska, A. K., Szadkowska, M., Kozłowski, L. P., Wyrzykowski, D., . . . Kaczorowski, T. (2021). Novel lytic enzyme of prophage origin

- from *Clostridium botulinum* E3 strain Alaska E43 with bactericidal activity against clostridial cells. *Int J Mol Sci*, 22(17). doi:10.3390/ijms22179536
- Moscoso, M., Garcia, E., & Lopez, R. (2009). Pneumococcal biofilms. *Int Microbiol*, 12(2), 77-85. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/19784927>
- Naas, T., Oueslati, S., Bonnin, R. A., Dabos, M. L., Zavala, A., Dortet, L., . . . Iorga, B. I. (2017). Beta-lactamase database (BLDB) - structure and function. *J Enzyme Inhib Med Chem*, 32(1), 917-919. doi:10.1080/14756366.2017.1344235
- Nau, R., & Eiffert, H. (2002). Modulation of release of proinflammatory bacterial compounds by antibacterials: potential impact on course of inflammation and outcome in sepsis and meningitis. *Clin Microbiol Rev*, 15(1), 95-110. doi:10.1128/CMR.15.1.95-110.2002
- Nelson, A. L., Roche, A. M., Gould, J. M., Chim, K., Ratner, A. J., & Weiser, J. N. (2007). Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect Immun*, 75(1), 83-90. doi:10.1128/IAI.01475-06
- Nelson, D., Loomis, L., & Fischetti, V. A. (2001). Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. *Proc Natl Acad Sci U S A*, 98(7), 4107-4112. doi:10.1073/pnas.061038398
- Nelson, D., Schuch, R., Chahales, P., Zhu, S., & Fischetti, V. A. (2006). PlyC: a multimeric bacteriophage lysin. *Proc Natl Acad Sci U S A*, 103(28), 10765-10770. doi:10.1073/pnas.0604521103
- Nelson, D. C., Schmelcher, M., Rodriguez-Rubio, L., Klumpp, J., Pritchard, D. G., Dong, S., & Donovan, D. M. (2012). Endolysins as antimicrobials. *Adv Virus Res*, 83, 299-365. doi:10.1016/B978-0-12-394438-2.00007-4
- Nishimoto, A. T., Rosch, J. W., & Tuomanen, E. I. (2020). Pneumolysin: pathogenesis and therapeutic target. *Front Microbiol*, 11, 1543. doi:10.3389/fmicb.2020.01543
- Nollmann, M., Gilbert, R., Mitchell, T., Sferrazza, M., & Byron, O. (2004). The role of cholesterol in the activity of pneumolysin, a bacterial protein toxin. *Biophys J*, 86(5), 3141-3151. doi:10.1016/S0006-3495(04)74362-3
- Nuermberger, E. L., & Bishai, W. R. (2004). Antibiotic resistance in *Streptococcus pneumoniae*: what does the future hold? *Clin Infect Dis*, 38 Suppl 4, S363-371. doi:10.1086/382696
- Nunes, S., Sa-Leao, R., Carrico, J., Alves, C. R., Mato, R., Avo, A. B., . . . de Lencastre, H. (2005). Trends in drug resistance, serotypes, and molecular types of *Streptococcus pneumoniae* colonizing preschool-age children attending day care centers in Lisbon, Portugal: a summary of 4 years of annual surveillance. *J Clin Microbiol*, 43(3), 1285-1293. doi:10.1128/JCM.43.3.1285-1293.2005
- O'Flaherty, S., Coffey, A., Meaney, W., Fitzgerald, G. F., & Ross, R. P. (2005). The recombinant phage lysin LysK has a broad spectrum of lytic activity against clinically relevant staphylococci, including methicillin-resistant *Staphylococcus aureus*. *J Bacteriol*, 187(20), 7161-7164. doi:10.1128/JB.187.20.7161-7164.2005
- Obeso, J. M., Martinez, B., Rodriguez, A., & Garcia, P. (2008). Lytic activity of the recombinant staphylococcal bacteriophage PhiH5 endolysin active against *Staphylococcus aureus* in milk. *Int J Food Microbiol*, 128(2), 212-218. doi:10.1016/j.ijfoodmicro.2008.08.010
- Oey, M., Lohse, M., Kreikemeyer, B., & Bock, R. (2009). Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. *Plant J*, 57(3), 436-445. doi:10.1111/j.1365-313X.2008.03702.x
- Oey, M., Lohse, M., Scharff, L. B., Kreikemeyer, B., & Bock, R. (2009). Plastid production of protein antibiotics against pneumonia via a new strategy for high-level expression of

- antimicrobial proteins. *Proc Natl Acad Sci U S A*, 106(16), 6579-6584. doi:10.1073/pnas.0813146106
- Oliveira, H., Sao-Jose, C., & Azeredo, J. (2018). Phage-derived peptidoglycan degrading enzymes: challenges and future prospects for *in vivo* therapy. *Viruses*, 10(6). doi:10.3390/v10060292
- Pastagia, M., Euler, C., Chahales, P., Fuentes-Duculan, J., Krueger, J. G., & Fischetti, V. A. (2011). A novel chimeric lysin shows superiority to mupirocin for skin decolonization of methicillin-resistant and -sensitive *Staphylococcus aureus* strains. *Antimicrob Agents Chemother*, 55(2), 738-744. doi:10.1128/AAC.00890-10
- Perez-Dorado, I., Gonzalez, A., Morales, M., Sanles, R., Striker, W., Vollmer, W., . . . Hermoso, J. A. (2010). Insights into pneumococcal fratricide from the crystal structures of the modular killing factor LytC. *Nat Struct Mol Biol*, 17(5), 576-581. doi:10.1038/nsmb.1817
- Pinto, D., Goncalo, R., Louro, M., Silva, M. S., Hernandez, G., Cordeiro, T. N., . . . Sao-Jose, C. (2022). On the occurrence and multimerization of two-polypeptide phage endolysins encoded in single genes. *Microbiol Spectr*, 10(4), e0103722. doi:10.1128/spectrum.01037-22
- Porter, C. J., Schuch, R., Pelzek, A. J., Buckle, A. M., McGowan, S., Wilce, M. C., . . . Whisstock, J. C. (2007). The 1.6 Å crystal structure of the catalytic domain of PlyB, a bacteriophage lysin active against *Bacillus anthracis*. *J Mol Biol*, 366(2), 540-550. doi:10.1016/j.jmb.2006.11.056
- Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., & Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Res*, 46(W1), W200-W204. doi:10.1093/nar/gky448
- Pritchard, D. G., Dong, S., Baker, J. R., & Engler, J. A. (2004). The bifunctional peptidoglycan lysin of *Streptococcus agalactiae* bacteriophage B30. *Microbiology (Reading)*, 150(Pt 7), 2079-2087. doi:10.1099/mic.0.27063-0
- Pritchard, D. G., Dong, S., Kirk, M. C., Cartee, R. T., & Baker, J. R. (2007). LambdaSa1 and LambdaSa2 prophage lysins of *Streptococcus agalactiae*. *Appl Environ Microbiol*, 73(22), 7150-7154. doi:10.1128/AEM.01783-07
- Proenca, D., Velours, C., Leandro, C., Garcia, M., Pimentel, M., & Sao-Jose, C. (2015). A two-component, multimeric endolysin encoded by a single gene. *Mol Microbiol*, 95(5), 739-753. doi:10.1111/mmi.12857
- Rai, P., He, F., Kwang, J., Engelward, B. P., & Chow, V. T. (2016). Pneumococcal pneumolysin induces DNA damage and cell cycle arrest. *Sci Rep*, 6, 22972. doi:10.1038/srep22972
- Ramirez, M. S., & Tolmasky, M. E. (2010). Aminoglycoside modifying enzymes. *Drug Resist Updat*, 13(6), 151-171. doi:10.1016/j.drug.2010.08.003
- Rashel, M., Uchiyama, J., Ujihara, T., Uehara, Y., Kuramoto, S., Sugihara, S., . . . Matsuzaki, S. (2007). Efficient elimination of multidrug-resistant *Staphylococcus aureus* by cloned lysin derived from bacteriophage phi MR11. *J Infect Dis*, 196(8), 1237-1247. doi:10.1086/521305
- Raz, A., Serrano, A., Lawson, C., Thaker, M., Alston, T., Bournazos, S., . . . Fischetti, V. A. (2017). Lysibodies are IgG Fc fusions with lysin binding domains targeting *Staphylococcus aureus* wall carbohydrates for effective phagocytosis. *Proc Natl Acad Sci U S A*, 114(18), 4781-4786. doi:10.1073/pnas.1619249114
- Resch, G., Moreillon, P., & Fischetti, V. A. (2011a). PEGylating a bacteriophage endolysin inhibits its bactericidal activity. *AMB Express*, 1, 29. doi:10.1186/2191-0855-1-29

- Resch, G., Moreillon, P., & Fischetti, V. A. (2011b). A stable phage lysin (Cpl-1) dimer with increased antipneumococcal activity and decreased plasma clearance. *Int J Antimicrob Agents*, 38(6), 516-521. doi:10.1016/j.ijantimicag.2011.08.009
- Riegler, A. N., Brissac, T., Gonzalez-Juarbe, N., & Orihuela, C. J. (2019). Necroptotic cell death promotes adaptive immunity against colonizing pneumococci. *Front Immunol*, 10, 615. doi:10.3389/fimmu.2019.00615
- Ring, A., Weiser, J. N., & Tuomanen, E. I. (1998). Pneumococcal trafficking across the blood-brain barrier. Molecular analysis of a novel bidirectional pathway. *J Clin Invest*, 102(2), 347-360. doi:10.1172/JCI2406
- Roach, D. R., & Donovan, D. M. (2015). Antimicrobial bacteriophage-derived proteins and therapeutic applications. *Bacteriophage*, 5(3), e1062590. doi:10.1080/21597081.2015.1062590
- Robinson, K. A., Baughman, W., Rothrock, G., Barrett, N. L., Pass, M., Lexau, C., . . . Active Bacterial Core Surveillance /Emerging Infections Program, N. (2001). Epidemiology of invasive *Streptococcus pneumoniae* infections in the United States, 1995-1998: opportunities for prevention in the conjugate vaccine era. *JAMA*, 285(13), 1729-1735. doi:10.1001/jama.285.13.1729
- Roig-Molina, E., Sanchez-Angulo, M., Seele, J., Garcia-Asencio, F., Nau, R., Sanz, J. M., & Maestro, B. (2020). Searching for antipneumococcal targets: choline-binding modules as phagocytosis enhancers. *ACS Infect Dis*, 6(5), 954-974. doi:10.1021/acsinfecdis.9b00344
- Rostol, J. T., & Marraffini, L. (2019). (Ph)ighting phages: how bacteria resist their parasites. *Cell Host Microbe*, 25(2), 184-194. doi:10.1016/j.chom.2019.01.009
- Royet, J., & Dziarski, R. (2007). Peptidoglycan recognition proteins: pleiotropic sensors and effectors of antimicrobial defences. *Nat Rev Microbiol*, 5(4), 264-277. doi:10.1038/nrmicro1620
- Sa-Leao, R., Tomasz, A., Sanches, I. S., Nunes, S., Alves, C. R., Avo, A. B., . . . de Lencastre, H. (2000). Genetic diversity and clonal patterns among antibiotic-susceptible and -resistant *Streptococcus pneumoniae* colonizing children: day care centers as autonomous epidemiological units. *J Clin Microbiol*, 38(11), 4137-4144. doi:10.1128/JCM.38.11.4137-4144.2000
- Sanz, J. M., Garcia, J. L., Laynez, J., Usobiaga, P., & Menendez, M. (1993). Thermal stability and cooperative domains of CPL1 lysozyme and its NH₂- and COOH-terminal modules. Dependence on choline binding. *J Biol Chem*, 268(9), 6125-6130. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/8454587>
- Sao-Jose, C. (2018). Engineering of phage-derived lytic enzymes: improving their potential as antimicrobials. *Antibiotics (Basel)*, 7(2). doi:10.3390/antibiotics7020029
- Sao-Jose, C., Costa, A. R., & Melo, L. D. R. (2022). Editorial: bacteriophages and their lytic enzymes as alternative antibacterial therapies in the age of antibiotic resistance. *Front Microbiol*, 13, 884176. doi:10.3389/fmicb.2022.884176
- Sass, P., & Bierbaum, G. (2007). Lytic activity of recombinant bacteriophage phi11 and phi12 endolysins on whole cells and biofilms of *Staphylococcus aureus*. *Appl Environ Microbiol*, 73(1), 347-352. doi:10.1128/AEM.01616-06
- Schmelcher, M., Donovan, D. M., & Loessner, M. J. (2012). Bacteriophage endolysins as novel antimicrobials. *Future Microbiol*, 7(10), 1147-1171. doi:10.2217/fmb.12.97
- Schmelcher, M., Powell, A. M., Becker, S. C., Camp, M. J., & Donovan, D. M. (2012). Chimeric phage lysins act synergistically with lysostaphin to kill mastitis-causing *Staphylococcus*

- aureus* in murine mammary glands. *Appl Environ Microbiol*, 78(7), 2297-2305. doi:10.1128/AEM.07050-11
- Schmelcher, M., Shabarova, T., Eugster, M. R., Eichenseher, F., Tchang, V. S., Banz, M., & Loessner, M. J. (2010). Rapid multiplex detection and differentiation of *Listeria* cells by use of fluorescent phage endolysin cell wall binding domains. *Appl Environ Microbiol*, 76(17), 5745-5756. doi:10.1128/AEM.00801-10
- Schmelcher, M., Waldherr, F., & Loessner, M. J. (2012). *Listeria* bacteriophage peptidoglycan hydrolases feature high thermoresistance and reveal increased activity after divalent metal cation substitution. *Appl Microbiol Biotechnol*, 93(2), 633-643. doi:10.1007/s00253-011-3372-6
- Schuch, R., Nelson, D., & Fischetti, V. A. (2002). A bacteriolytic agent that detects and kills *Bacillus anthracis*. *Nature*, 418(6900), 884-889. doi:10.1038/nature01026
- Seale, A. C., Bianchi-Jassir, F., Russell, N. J., Kohli-Lynch, M., Tann, C. J., Hall, J., . . . Lawn, J. E. (2017). Estimates of the burden of group B streptococcal disease worldwide for pregnant women, stillbirths, and children. *Clin Infect Dis*, 65(suppl_2), S200-S219. doi:10.1093/cid/cix664
- Seo, H. S., Cartee, R. T., Pritchard, D. G., & Nahm, M. H. (2008). A new model of pneumococcal lipoteichoic acid structure resolves biochemical, biosynthetic, and serologic inconsistencies of the current model. *J Bacteriol*, 190(7), 2379-2387. doi:10.1128/JB.01795-07
- Shak, J. R., Ludewick, H. P., Howery, K. E., Sakai, F., Yi, H., Harvey, R. M., . . . Vidal, J. E. (2013). Novel role for the *Streptococcus pneumoniae* toxin pneumolysin in the assembly of biofilms. *mBio*, 4(5), e00655-00613. doi:10.1128/mBio.00655-13
- Shatalin, K., Nuthanakanti, A., Kaushik, A., Shishov, D., Peselis, A., Shamovsky, I., . . . Nudler, E. (2021). Inhibitors of bacterial H(2)S biogenesis targeting antibiotic resistance and tolerance. *Science*, 372(6547), 1169-1175. doi:10.1126/science.abd8377
- Shen, Y., Barros, M., Vennemann, T., Gallagher, D. T., Yin, Y., Linden, S. B., . . . Nelson, D. C. (2016). A bacteriophage endolysin that eliminates intracellular streptococci. *Elife*, 5. doi:10.7554/eLife.13152
- Shenoy, A. T., Brissac, T., Gilley, R. P., Kumar, N., Wang, Y., Gonzalez-Juarbe, N., . . . Orihuela, C. J. (2017). *Streptococcus pneumoniae* in the heart subvert the host response through biofilm-mediated resident macrophage killing. *PLoS Pathog*, 13(8), e1006582. doi:10.1371/journal.ppat.1006582
- Shenoy, A. T., & Orihuela, C. J. (2016). Anatomical site-specific contributions of pneumococcal virulence determinants. *Pneumonia (Nathan)*, 8. doi:10.1186/s41479-016-0007-9
- Silva, M. D., Andre, C., & Bispo, P. J. M. (2022). Targeted killing of ocular *Streptococcus pneumoniae* by the phage endolysin MSlys. *Ophthalmol Sci*, 2(4), 100193. doi:10.1016/j.xops.2022.100193
- Sims Sanyahumbi, A., Colquhoun, S., Wyber, R., & Carapetis, J. R. (2016). Global disease burden of group A streptococcus. In J. J. Ferretti, D. L. Stevens, & V. A. Fischetti (Eds.), *Streptococcus pyogenes: Basic Biology to Clinical Manifestations*. Oklahoma City (OK).
- Song, J. Y., Nahm, M. H., & Moseley, M. A. (2013). Clinical implications of pneumococcal serotypes: invasive disease potential, clinical presentations, and antibiotic resistance. *J Korean Med Sci*, 28(1), 4-15. doi:10.3346/jkms.2013.28.1.4
- Soto, S. M. (2013). Role of efflux pumps in the antibiotic resistance of bacteria embedded in a biofilm. *Virulence*, 4(3), 223-229. doi:10.4161/viru.23724

- Stapleton, P. D., & Taylor, P. W. (2002). Methicillin resistance in *Staphylococcus aureus*: mechanisms and modulation. *Sci Prog*, 85(Pt 1), 57-72. doi:10.3184/003685002783238870
- Steel, H. C., Cockeran, R., Anderson, R., & Feldman, C. (2013). Overview of community-acquired pneumonia and the role of inflammatory mechanisms in the immunopathogenesis of severe pneumococcal disease. *Mediators Inflamm*, 2013, 490346. doi:10.1155/2013/490346
- Steinegger, M., & Soding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol*, 35(11), 1026-1028. doi:10.1038/nbt.3988
- Summers, W. C. (2011). In the beginning. *Bacteriophage*, 1(1), 50-51. doi:10.4161/bact.1.1.14070
- Sun, H., Chai, X., Xu, G., & Wei, S. (2023). The formation and drug resistance mechanism of biofilm for *Streptococcus pneumoniae* infection in severe respiratory patients. *Cell Mol Biol (Noisy-le-grand)*, 69(1), 75-80. doi:10.14715/cmb/2022.69.1.13
- Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., . . . Group, W. H. O. P. P. L. W. (2018). Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis*, 18(3), 318-327. doi:10.1016/S1473-3099(17)30753-3
- Takac, M., Witte, A., & Blasi, U. (2005). Functional analysis of the lysis genes of *Staphylococcus aureus* phage P68 in *Escherichia coli*. *Microbiology (Reading)*, 151(Pt 7), 2331-2342. doi:10.1099/mic.0.27937-0
- Tedesco, F. J., Barton, R. W., & Alpers, D. H. (1974). Clindamycin-associated colitis. A prospective study. *Ann Intern Med*, 81(4), 429-433. doi:10.7326/0003-4819-81-4-429
- Tomasz, A. (1968). Biological consequences of the replacement of choline by ethanolamine in the cell wall of pneumococcus: choline formation, loss of transformability, and loss of autolysis. *Proc Natl Acad Sci U S A*, 59(1), 86-93. doi:10.1073/pnas.59.1.86
- Usobiaga, P., Medrano, F. J., Gasset, M., Garcia, J. L., Saiz, J. L., Rivas, G., . . . Menendez, M. (1996). Structural organization of the major autolysin from *Streptococcus pneumoniae*. *J Biol Chem*, 271(12), 6832-6838. doi:10.1074/jbc.271.12.6832
- van der Poll, T., & Opal, S. M. (2009). Pathogenesis, treatment, and prevention of pneumococcal pneumonia. *Lancet*, 374(9700), 1543-1556. doi:10.1016/S0140-6736(09)61114-4
- Vander Elst, N., Linden, S. B., Lavigne, R., Meyer, E., Briers, Y., & Nelson, D. C. (2020). Characterization of the bacteriophage-derived endolysins PlySs2 and PlySs9 with *in vitro* lytic activity against bovine mastitis *Streptococcus uberis*. *Antibiotics (Basel)*, 9(9). doi:10.3390/antibiotics9090621
- Varea, J., Monterroso, B., Saiz, J. L., Lopez-Zumel, C., Garcia, J. L., Laynez, J., . . . Menendez, M. (2004). Structural and thermodynamic characterization of Pal, a phage natural chimeric lysin active against pneumococci. *J Biol Chem*, 279(42), 43697-43707. doi:10.1074/jbc.M407067200
- Varea, J., Saiz, J. L., Lopez-Zumel, C., Monterroso, B., Medrano, F. J., Arrondo, J. L., . . . Menendez, M. (2000). Do sequence repeats play an equivalent role in the choline-binding module of pneumococcal LytA amidase? *J Biol Chem*, 275(35), 26842-26855. doi:10.1074/jbc.M004379200
- Vassal-Stermann, E., Lacroix, M., Gout, E., Laffly, E., Pedersen, C. M., Martin, L., . . . Thielens, N. M. (2014). Human L-ficolin recognizes phosphocholine moieties of pneumococcal teichoic acid. *J Immunol*, 193(11), 5699-5708. doi:10.4049/jimmunol.1400127
- Vaughn, V. M., Gandhi, T. N., Petty, L. A., Patel, P. K., Prescott, H. C., Malani, A. N., . . . Flanders, S. A. (2021). Empiric antibacterial therapy and community-onset bacterial coinfection in

- patients hospitalized with coronavirus Disease 2019 (COVID-19): a multi-hospital cohort study. *Clin Infect Dis*, 72(10), e533-e541. doi:10.1093/cid/ciaa1239
- Veiga-Crespo, P., Ageitos, J. M., Poza, M., & Villa, T. G. (2007). Enzybiotics: a look to the future, recalling the past. *J Pharm Sci*, 96(8), 1917-1924. doi:10.1002/jps.20853
- Volanakis, J. E., & Kaplan, M. H. (1971). Specificity of C-reactive protein for choline phosphate residues of pneumococcal C-polysaccharide. *Proc Soc Exp Biol Med*, 136(2), 612-614. doi:10.3181/00379727-136-35323
- Vollmer, W. (2008). Structural variation in the glycan strands of bacterial peptidoglycan. *FEMS Microbiol Rev*, 32(2), 287-306. doi:10.1111/j.1574-6976.2007.00088.x
- Vouillamoz, J., Entenza, J. M., Giddey, M., Fischetti, V. A., Moreillon, P., & Resch, G. (2013). Bactericidal synergism between daptomycin and the phage lysin Cpl-1 in a mouse model of pneumococcal bacteraemia. *Int J Antimicrob Agents*, 42(5), 416-421. doi:10.1016/j.ijantimicag.2013.06.020
- Walcher, G., Stessl, B., Wagner, M., Eichenseher, F., Loessner, M. J., & Hein, I. (2010). Evaluation of paramagnetic beads coated with recombinant *Listeria* phage endolysin-derived cell-wall-binding domain proteins for separation of *Listeria monocytogenes* from raw milk in combination with culture-based and real-time polymerase chain reaction-based quantification. *Foodborne Pathog Dis*, 7(9), 1019-1024. doi:10.1089/fpd.2009.0475
- Wall, R. J., Powell, A. M., Paape, M. J., Kerr, D. E., Bannerman, D. D., Pursel, V. G., . . . Hawk, H. W. (2005). Genetically enhanced cows resist intramammary *Staphylococcus aureus* infection. *Nat Biotechnol*, 23(4), 445-451. doi:10.1038/nbt1078
- Walsh, S., Shah, A., & Mond, J. (2003). Improved pharmacokinetics and reduced antibody reactivity of lysostaphin conjugated to polyethylene glycol. *Antimicrob Agents Chemother*, 47(2), 554-558. doi:10.1128/AAC.47.2.554-558.2003
- Wang, Y., Khanal, D., Alreja, A. B., Yang, H., Yk Chang, R., Tai, W., . . . Chan, H. K. (2023). Bacteriophage endolysin powders for inhaled delivery against pulmonary infections. *Int J Pharm*, 635, 122679. doi:10.1016/j.ijpharm.2023.122679
- Wang, Y., Sun, J. H., & Lu, C. P. (2009). Purified recombinant phage lysin LySMP: an extensive spectrum of lytic activity for swine streptococci. *Curr Microbiol*, 58(6), 609-615. doi:10.1007/s00284-009-9379-x
- Weiser, J. N., Ferreira, D. M., & Paton, J. C. (2018). *Streptococcus pneumoniae*: transmission, colonization and invasion. *Nat Rev Microbiol*, 16(6), 355-367. doi:10.1038/s41579-018-0001-8
- WHO. Immunization, vaccines and biologicals. Retrieved from <https://www.who.int/teams/immunization-vaccines-and-biologicals/diseases/pneumonia>
- WHO. (2019). *Pneumococcal conjugate vaccines in infants and children under 5 years of age: WHO position paper - February 2019*. Retrieved from WHO: <https://apps.who.int/iris/bitstream/handle/10665/310968/WER9408.pdf?ua=1>
- WHO. (2021). *Pneumonia fact sheet*. Retrieved from WHO: <https://www.who.int/news-room/fact-sheets/detail/pneumonia>
- Witzenrath, M., Schmeck, B., Doehn, J. M., Tschernig, T., Zahlten, J., Loeffler, J. M., . . . Rosseau, S. (2009). Systemic use of the endolysin Cpl-1 rescues mice with fatal pneumococcal pneumonia. *Crit Care Med*, 37(2), 642-649. doi:10.1097/CCM.0b013e31819586a6
- Wolcott, R. D., & Ehrlich, G. D. (2008). Biofilms and chronic infections. *JAMA*, 299(22), 2682-2684. doi:10.1001/jama.299.22.2682

- Wong, J. L. C., Romano, M., Kerry, L. E., Kwong, H. S., Low, W. W., Brett, S. J., . . . Frankel, G. (2019). OmpK36-mediated carbapenem resistance attenuates ST258 *Klebsiella pneumoniae* in vivo. *Nat Commun*, 10(1), 3957. doi:10.1038/s41467-019-11756-y
- www.compoundchem.com. Different classes of antibiotics - an overview. Retrieved from <https://i0.wp.com/www.compoundchem.com/wp-content/uploads/2014/09/A-Guide-to-Different-Classes-of-Antibiotics-Aug-15.png?ssl=1>
- Xu, L., Mo, K., Ran, D., Ma, J., Zhang, L., Sun, Y., . . . Zou, X. (2023). An endolysin gene from *Candidatus liberibacter asiaticus* confers dual resistance to huanglongbing and citrus canker. *Hortic Res*, 10(9), uhad159. doi:10.1093/hr/uhad159
- Yahiaoui, R. Y., den Heijer, C., van Bijnen, E. M., Paget, W. J., Pringle, M., Goossens, H., . . . Team, A. S. (2016). Prevalence and antibiotic resistance of commensal *Streptococcus pneumoniae* in nine European countries. *Future Microbiol*, 11, 737-744. doi:10.2217/fmb-2015-0011
- Yamashita, Y., Shibata, Y., Nakano, Y., Tsuda, H., Kido, N., Ohta, M., & Koga, T. (1999). A novel gene required for rhamnose-glucose polysaccharide synthesis in *Streptococcus mutans*. *J Bacteriol*, 181(20), 6556-6559. doi:10.1128/JB.181.20.6556-6559.1999
- Yang, H., Gong, Y., Zhang, H., Etobayeva, I., Miernikiewicz, P., Luo, D., . . . Wei, H. (2019). ClyJ is a novel pneumococcal chimeric lysin with a cysteine- and histidine-dependent amidohydrolase/peptidase catalytic domain. *Antimicrob Agents Chemother*, 63(4). doi:10.1128/AAC.02043-18
- Yang, H., Linden, S. B., Wang, J., Yu, J., Nelson, D. C., & Wei, H. (2015). A chimeolysin with extended-spectrum streptococcal host range found by an induced lysis-based rapid screening method. *Sci Rep*, 5, 17257. doi:10.1038/srep17257
- Yang, H., Luo, D., Etobayeva, I., Li, X., Gong, Y., Wang, S., . . . Wei, H. (2020). Linker editing of pneumococcal lysin ClyJ conveys improved bactericidal activity. *Antimicrob Agents Chemother*, 64(2). doi:10.1128/AAC.01610-19
- Yoong, P., Schuch, R., Nelson, D., & Fischetti, V. A. (2004). Identification of a broadly active phage lytic enzyme with lethal activity against antibiotic-resistant *Enterococcus faecalis* and *Enterococcus faecium*. *J Bacteriol*, 186(14), 4808-4812. doi:10.1128/JB.186.14.4808-4812.2004
- Yoong, P., Schuch, R., Nelson, D., & Fischetti, V. A. (2006). PlyPH, a bacteriolytic enzyme with a broad pH range of activity and lytic action against *Bacillus anthracis*. *J Bacteriol*, 188(7), 2711-2714. doi:10.1128/JB.188.7.2711-2714.2006
- Zhou, B., Zhen, X., Zhou, H., Zhao, F., Fan, C., Perculija, V., . . . Ouyang, S. (2020). Structural and functional insights into a novel two-component endolysin encoded by a single gene in *Enterococcus faecalis* phage. *PLoS Pathog*, 16(3), e1008394. doi:10.1371/journal.ppat.1008394
- Zielnik-Jurkiewicz, B., & Bielicka, A. (2015). Antibiotic resistance of *Streptococcus pneumoniae* in children with acute otitis media treatment failure. *Int J Pediatr Otorhinolaryngol*, 79(12), 2129-2133. doi:10.1016/j.ijporl.2015.09.030
- Zimmer, M., Vukov, N., Scherer, S., & Loessner, M. J. (2002). The murein hydrolase of the bacteriophage phi3626 dual lysis system is active against all tested *Clostridium perfringens* strains. *Appl Environ Microbiol*, 68(11), 5311-5317. doi:10.1128/AEM.68.11.5311-5317.2002
- Zysk, G., Schneider-Wald, B. K., Hwang, J. H., Bejo, L., Kim, K. S., Mitchell, T. J., . . . Heinz, H. P. (2001). Pneumolysin is the main inducer of cytotoxicity to brain microvascular

endothelial cells caused by *Streptococcus pneumoniae*. *Infect Immun*, 69(2), 845-852.
doi:10.1128/IAI.69.2.845-852.2001