Type VIIb secretion system effector export and neutralization

Mechanistic insights into type VIIb secretion system effector export and neutralization

Timothy A. Klein, B.Sc.

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Doctor of Philosophy

DESCRIPTIVE NOTE

McMaster University DOCTOR OF PHILOSOPHY (2022), Hamilton, Ontario (Biochemistry and Biomedical Sciences)

TITLE: Mechanistic insights into type VIIb secretion system effector export and neutralization

AUTHOR: Timothy A. Klein, B.Sc.

SUPERVISOR: John C. Whitney, Ph.D.

NUMBER OF PAGES: xxi, 256

FOREWORD

Lay Abstract

Bacteria require space and various nutrients to survive and grow and must therefore compete against other bacteria for access to these resources. To gain advantage over their competitors, many bacteria have developed molecular weapons that target and kill other closely related bacteria. Some of these weapons take the form of protein secretion machines that export antibacterial toxins. Gram-positive bacteria use the type VIIb secretion system (T7SSb) to inhibit the growth of other Gram-positive bacteria. In this work, I explore several aspects of T7SSb including: (1) how toxins are inhibited by immunity proteins, (2) how toxins are secreted through the cell envelope, and (3) how toxins are recognized by the secretion apparatus. The goal of this work is to better understand how T7SSb functions at the molecular level.

Abstract

The type VII secretion system is a protein export pathway linked to diverse phenotypes in both Actinobacteria and Firmicutes. The Actinobacterial subtype of the T7SS, referred to as T7SSa, has been shown to play a critical role in various aspects of Mycobacterial life including virulence, conjugation, and metal homeostasis. The T7SSb of Firmicutes bacteria on the other hand has similarly been shown to influence virulence but by the direct growth inhibition of competitor bacteria. Structure-function analyses of the T7SSa apparatus as well as various effectors and chaperones have begun to build a more mechanistic understanding of how T7SSa functions. In contrast, we know little of how the T7SSb functions despite its noted importance to both pathogens and environmental bacteria such as *Bacillus*, *Staphylococcus*, *Enterococcus*, and *Streptococcus*. During my thesis work, I have addressed several gaps in our understanding of T7SSb function. The three major questions that I have studied are: (1) how do T7SSb immunity proteins inhibit the toxicity of their cognate toxins, (2) how does the T7SSb export effectors through the thick Gram-positive cell wall, and (3) what is the role of chaperone proteins in facilitating T7SSb effector export?

Acknowledgements

To my boss, Dr. John Whitney, I want to thank you for mentoring me through these past five years. Your support and guidance has made my academic journey far more productive and enjoyable than it would have been without you. I view a PhD as being fundamentally about learning new skills and, as such, I am incredibly thankful for everything you have taught me about conducting biochemistry, structural biology, and microbiology research. However, what I believe is more important are the intangible things you taught me including (but not limited to) how to share our research with a broad audience, how to collaborate with other scientists, and how to mentor the next generation of researchers. If you are ever in doubt about your success as a mentor just remember that you got me into a postdoc position and Sheri into medical school. Congrats on going 2 for 2!

To my committee, Dr. Brian Coombes and Dr. Michael Surette, thank you for all your guidance and valuable advice throughout the years. The two of you consistently offered useful additions to our projects and made each of our T7SS papers better. I must also thank you for advising me to look into postdoc positions early as it led to far better opportunities than I would have had otherwise. I would also like to thank my external examiner Dr. Tracy Palmer for her valuable feedback on my thesis. Thank you all for your mentorship.

To my labmates, Dr. Sheri Ahmad, Nathan Bullen, Dirk Grebenc, Amir Azhieh, Kartik Sachar, and Andrea Alexei, thank you for all the love and support you have given me on this wild ride. Each and every one of you has offered me useful advice throughout

this journey and I could not have finished my PhD without you. Doing a PhD can definitely be a slog, but it is made a lot easier by doing it alongside a talented group of friends and colleagues.

To my students, Shil Gandhi, Vraj Shah, Brandon Dickson, Prakhar Shah, and Owen McArthur, thank you for teaching me how to teach. A mentor can only be as good as his students and each of you undoubtedly made me better.

To all of my scientist friends outside of the Whitney lab including (but certainly not limited to), Dr. Bushra Ilyas, Dr. Alex Oberc, Dr. Elizabeth Chau, Dr. Caressa Tsai, Dr. Dustin Little, Dr. Wael Elhenawy, Dr. Dave Sychantha, Dr. Lindsey Marmont, Sharok Shekarriz, Dr. Matt Duda, and Jaimie Bortolotti, thank you all for the support and knowledge that you bestowed upon me. I am incredibly lucky to have such a wise and kind-hearted friend group. To anyone not mentioned above, just know that I am thankful for you too!

To all of the administrators in both the Biochemistry department and the IIDR who have facilitated my work, thank you for all of the timely guidance and assistance. You made my work easier in dozens of different ways and for that I will always be grateful.

To my mom, dad, and sisters Katie and Ellen, and to my entire family, thank you for giving me the faculties to tackle this PhD (and life as a whole). Without you, I would never have had the skill, patience, and confidence to finish my degree. Although my name will go on the diploma, my PhD is as much a reflection of your hard work and dedication as it is of mine.

To Carly, I am so grateful to be on this journey called life with you. You have supported me through every second of my PhD and have been there for all of my successes and my failures. I could not have asked for a more loving and caring partner to have stood by me during my PhD. I am especially thankful that you (usually) remember that I work with *Streptococcus* (and not *Streptomyces*) and study T7SS (rather than T6SS). I cannot wait for the next chapter of our life together: sunny San Francisco!!

To all of the aforementioned people and probably about 100 others who I have neglected, I am incredibly lucky to have been on this path with you and so I owe a piece of my PhD to each of you. Thank you!

Table of Contents

List of Figures

Chapter IV figures

Appendix: supplemental figures

List of Tables and Datasets

Appendix: supplemental tables

List of Abbreviations

BCG……...bacillus of Calmette-Guérin

BME…….……β-mercaptoethanol

CFP-10……culture filtrate protein of 10kDa

Cryo-EM……...cryogenic electron microscopy

CSP……..competence stimulating peptide

DUF…….domain of unknown function

ESAT-6……early secreted antigenic target of 6kDa

Ecc……...ESX core component

Esa……..Esx associated

Esp……...ESX secretion associated proteins

Ess…….……..ESX secretion system

ESX (or Esx)……..ESAT-6 secretion complex

IWZ……...inner wall zone

kDa…….……kilodalton

Lap…….…….LXG-associate α-helical protein

LB…….……..…lysogeny broth

LXG……...leucine-x-glycine

NAD^+……..nicotinamide adenine dinucleotide

His_6……...hexahistidine

Ni-NTA…….nickel nitrilotriacetic acid

(P)PE……...(proline) proline glutamate

Sec……..general secretory system

Tat……..twin arginine translocase

Tel…….toxin exported by Esx with LXG domain

Tip……Tel immunity protein

TH……..Todd Hewitt

TMD…….transmembrane domain

T7S……type VII secretion

T7SS……..type VII secretion system

VSV-G……...vesicular stomatitis virus G protein

WXG……tryptophan-x-glycine

Declaration of Academic Achievement

I have performed all the research in this body of work except where indicated in the preface of each chapter.

CHAPTER I – Introduction

Preface

Although the work presented in this chapter is original and not borrowed from any previous publication it is inspired and directed by the following reviews:

Klein, T.A.*, Ahmad, S.*, and Whitney, J.C. (2020). Contact-dependent interbacterial antagonism mediated by protein secretion machines. *Trends in Microbiology* 28 (5): 387-400.

*These authors contributed equally.

Author contributions: All authors wrote and edited the paper and designed the figures.

Tran, H.R.*, Grebenc, D.W.*, Klein, T.A., and Whitney, J.C. (2021). Bacterial type VII secretion: an important player in host-microbe and microbe-microbe interactions. *Molecular Microbiology* 115: 478-489.

*These authors contributed equally.

Author contributions: All authors wrote and edited the paper. JCW designed the figures.

Permission has been granted by the publishers to reproduce the material from these studies.

The type VII secretion system: one of many antibacterial toxin export pathways

General principles of bacterial secretion

Bacterial life is a complex series of interactions with different environments, hosts, other bacteria, and bacteriophage. To deal with these dynamic interactions, bacteria have evolved several mechanisms that allow for communication with each other and the environment. One of these essential mechanisms is secretion, or the delivery of intracellular molecules, such as DNA, polysaccharides, small molecules and proteins, to the extracellular environment (Llosa et al., 2002; Limoli et al., 2015; Green & Mecsas, 2016). Bacterial DNA export plays a critical role in biofilm formation as extracellular DNA has been shown to seed bacterial biofilms (Yu et al., 2018). DNA secretion also allows for the passage of genes from one bacterium to the next via conjugation (Llosa et al., 2002). Polysaccharides secreted by bacteria are another frequent component of biofilms as well as bacterial capsules both of which protect bacterial cells from external threats (Limoli et al., 2015; Paton et al., 2019). Bacteria can secrete a myriad of small molecules, and while some promote co-operation through quorum sensing, others, such as bacteriocins and antibiotics, directly inhibit bacterial growth (Miller & Bassler, 2001; Cotter et al., 2013). Finally, bacteria secrete proteins through several distinct mechanisms, and these secreted proteins can play various roles including modulation of host cell physiology and growth inhibition of competing bacteria (Green & Mecsas, 2016).

To facilitate the secretion of protein effectors, bacteria rely on various secretion systems, which are macromolecular machines that actively transport proteins through an

otherwise impassable membrane. The most common and widespread of these secretion systems are the general secretory system (Sec) and the Twin-arginine translocase (Tat). Both systems are virtually ubiquitous, being distributed within both Gram-negative and Gram-positive bacteria as well as Archaea and Eukaryotes (Green & Mecsas, 2016). The Sec system of prokaryotes facilitates the secretion of unfolded polypeptides through the cell (or inner) membrane and also plays a role in incorporating integral membrane proteins into the membrane (Tsirigotaki et al., 2017). In contrast to Sec, Tat exports folded proteins through the cell membrane. Proteins secreted by Tat are often ones that require posttranslational modification or are metalloproteins (Palmer & Berks, 2012).

Along with the widespread Sec and Tat systems, Gram-negative bacteria can also have several more specialized secretion systems to export proteins (Green & Mecsas, 2016). Currently eleven of these more specialized secretion systems have been described and given a standardized "type X secretion system (TXSS)" nomenclature where "X" represents a numerical value that relates to the chronological order of discovery. The well characterized type I-VI secretion systems (T1SS to T6SS) as well as the less well studied type VIII-XI secretion systems (T8SS to T11SS) are all found in Gram-negative bacteria (Green & Mecsas, 2016; Lauber et al., 2018; Palmer et al., 2021; Grossman et al., 2021). In general, these systems secrete smaller repertoires of proteins than Sec and Tat and serve more specific roles for bacteria. For example, the T3SS plays an essential role in virulence for various pathogens such as *Salmonella enterica* and *Yersinia pestis*, while a homologous nanomachine, the flagellar secretion system is equally important for swimming motility (Zhang et al., 2018; Colin et al., 2021). Similarly, some bacteria use a

T6SS to target and inhibit the growth of competitors, while other bacteria use the same system to target host cells (Mougous et al., 2006; Pukatzki et al., 2006). Regardless of function, the specialized secretion systems can be divided into one- and two-step secretion mechanisms. One-step secretion systems export effectors through both the inner and outer membrane in a single step and in some cases can further transport these effectors through a third membrane such as that of a host or a competitor. In general, the secretion systems that can deliver effectors through a tertiary membrane are those that have a large extracellular pilus or needle-like appendage such as T3SS, T4SS, or T6SS. In contrast, two-step secretion systems, such as T2SS and T5SS, passage effectors through the outer membrane only and therefore require Sec or Tat to transport effectors through the inner membrane (Green & Mecsas, 2016).

One of the most prevalent functions that secreted proteins have is direct inhibition of competitor bacteria (Klein et al., 2020). This is because many, if not most, bacteria live in highly contested niches and must therefore compete with various competitor bacteria for space and nutrients. While bacterial protein secretion plays an important role in competition (see below), so too does the export of small diffusible molecules such as antibiotics and bacteriocins (Cotter et al., 2013). Antibiotics can be either secondary metabolites or non-ribosomally produced peptides and can have varying target ranges depending on class and mechanism of action (Demain, 1999; Guilhelmelli et al., 2013). Bacteriocins are small peptides produced by ribosomes that generally target a narrow range of bacteria (Riley & Wertz, 2002). Due to the toxic nature of these molecules, bacteria have developed various ways to defend themselves from intoxication.

Bacteriocins generally have immunity proteins which protect producing bacteria but can also provide an advantage to any target bacteria that encode them (Jeon et al., 2009). Similarly, resistance to secondary metabolite antibiotics is now known to be common and can occur through a myriad of methods (Cox & Wright, 2013). Although the production of antibiotics and bacteriocins is widespread, many bacteria also rely on secretion systems that directly inject protein toxins into neighbouring cells to gain control of a niche. Four Gram-negative secretion systems have been shown to conduct contact-dependent antibacterial targeting including T1SS, T4SS, T5SS, and T6SS (Klein et al., 2020, Ruhe et al., 2013). Although contact-dependent mechanisms of inhibition are short range compared to diffusible molecules, they have an advantage in the sense that they are highly targeted and avoid losses in potency due to the effect of dilution.

The best studied of the antibacterial contact-dependent inhibition systems is the T6SS, which is commonly found in Gram-negative bacterial genomes (as high as 25%) and although it has also been linked to virulence, it seems that the primary function of the system is in inhibiting the growth of competitor bacteria (Green & Mecsas, 2016, Mougous et al., 2006). The T6SS forms an inverted phage tail-like structure with a hollow nanotube that is spiked at the distal end and, when fired, injects a large payload of toxic effectors into an adjacent competitor cell (Leiman et al., 2009; Nguyen et al, 2018). Toxic effectors can be loaded onto/into the T6SS spiked nanotube through various ways, including inside the hollow tube, non-covalently linked to the tip, or as a C-terminal extension of the different proteins that make up the nanotube and spike (Klein et al., 2020). Regardless of which compartment of the T6SS they are localized to, the effectors

are delivered into the periplasm of an adjacent competitor cell through contraction of the sheath which, in turn, causes ejection of the spiked T6SS tube (Zoued et al., 2014). While some effectors will target essential molecules found within the periplasm, effectors that have cytoplasmic targets will self-translocate into the cytoplasm before inhibiting growth (Quentin et al., 2018). Various toxin functions have been characterized in the context of T6SS including cell wall targeting amidases and glycosidases, membrane depolarizing toxins, DNases, $NAD(P)^+$ hydrolases, (p)ppApp synthetases, and ADP-ribosyl transferases (Le et al., 2021; Mariano et al., 2019; Jana et al., 2019; Tang et al., 2018, Ahmad et al., 2019, Ting et al., 2018).

Although understudied compared to T6SS, it is now clear that certain T1SSs and T4SSs can also inhibit bacterial growth through contact-dependent mechanisms. T1SS is much simpler than T6SS, requiring only an inner membrane ABC transporter, an outer membrane pore and a membrane fusion protein which links the two (Kanonenberg et al., 2018). Although T1SSs are often required for bacterial virulence, it was recently discovered that a T1SS in *Caulobacter crescentus* secretes a two-peptide bacteriocin called CdzCD that can inhibit the growth of neighbouring cells. CdzCD contains N-terminal glycine zipper motifs that allow for aggregation in the outer membrane of the attacking cell (García-Bayona et al., 2017). These aggregates are then transferred directly to an adjacent competitor and likely kill by forming pores in the target's inner membrane. Like other bacteriocins, CdzCD has a narrow target range and can only inhibit the growth of other α-proteobacteria. Bioinformatic analysis suggests that, although understudied,

these antibacterial T1SSs are widespread and often transferred horizontally (García-Bayona et al., 2017).

T4SSs are highly variable in terms of both their function and cargo. Bacteria use the T4SS to directly transfer DNA cell-to-cell, in a process called conjugation, but T4SS can also be a major virulence factor and directly inject host-modulating proteins and protein-DNA complexes (Christie et al., 2005). More recently, T4SSs that impact bacterial competition have been identified in both *Xanthomonas citri* and *Stenotrophomonas maltophilia*. *X. citri* uses its T4SS to secrete a novel set of effectors called XVIPs which consist of an N-terminal toxin domain and a C-terminal T4SS-targeting domain (Souza et al., 2015; Bayer-Santos et al., 2019). An even larger set of effectors have been predicted from *S. maltophilia* genomes and although antibacterial T4SSs have only been discovered in a small set of γ- and β-proteobacteria, it is a tantalizing possibility that these bacteria-targeting systems are more widespread than currently appreciated (Bayer-Santos et al., 2019; Sgro et al., 2019).

The final characterized example of an antibacterial secretion system is T5SS, which is more commonly referred to as contact dependent growth inhibition (CDI) when it is antibacterial in nature (Ruhe et al., 2013). CDI generally consists of two components CdiA, which is the CDI toxin, and CdiB which is an outer membrane protein (Aoki et al., 2005). Secretion of the CdiA toxin occurs through several steps where first CdiB is transported through the inner membrane by Sec and forms a pore in the outer membrane. CdiA also passes through the inner membrane via Sec but then extends through the outer membrane via the CdiB pore (Ruhe et al., 2018). CdiA proteins often have very long

central regions that can extend over 100 nm beyond the cell surface. At the end of this stock-like appendage is the CdiA C-terminal toxin domain, which is transported into a competitor cell by directly binding to an outer membrane-embedded receptor (Ruhe et al., 2018). Because CdiA toxins require interaction with an outer membrane receptor, these systems generally have a narrow target range, similar to bacteriocins and antibacterial T1SSs (Ruhe et al., 2017). Several toxic activities have been ascribed to CdiA C-termini including DNase, RNase, and pore forming toxins (Hayes et al., 2014). Similar to the previously described antibacterial secretion systems, *cdiAB* genes are generally co-transcribed with a *cdiI* immunity determinant. Unsurprisingly, CdiI proteins have been shown to directly interact with the CdiA C-terminus to inhibit toxic activity (Aoki et al., 2010).

Despite obvious differences between the various antibacterial secretion systems encoded by bacteria, there are some overarching principles that unite the systems. Firstly, and with the distinct exception of CDI, each of the systems consist of a membrane embedded apparatus containing an ATPase that drives secretion. Secondly, the genes that encode for toxin effectors are co-transcribed with immunity factors that specifically inhibit their cognate toxins. Finally, antibacterial toxins target essential bacterial components and pathways such as transcription, translation, and cell wall synthesis. While these universal principals unite the various antibacterial systems of Gram-negative bacteria, they also link these systems to the only specialized secretion system encoded by Gram-positive bacteria, the type VII secretion system (T7SS) (Klein et al., 2020).

Discovery of the T7SSa in Actinobacteria

T7SS was originally discovered through comparative genomic studies of virulent strains of *Mycobacterium tuberculosis* and *Mycobacterium bovis* and the live attenuated vaccine strain for Tuberculosis, *M. bovis* bacillus Calmette-Guérin (BCG). In comparison to virulent Mycobacterial strains, BCG lacks a 9.5kb region of difference (RD1) that effectively renders the strain avirulent (Mahairas et al., 1996). It was later discovered that RD1 encodes a novel secretion system that is widely distributed in Actinobacteria and essential for the virulence of *M. tuberculosis* and other virulent Mycobacteria (Pym et al., 2002; Stanley et al., 2003). Though not named the T7SS until many years later, this system became a dominant focus in understanding mycobacterial physiology and virulence.

The T7SS of Actinobacteria is often referred to as the ESAT-6 secretion system (ESX, ESAT-6 being the first discovered effector) and a single species of Mycobacteria can encode up to five ESX systems. The different ESX systems play somewhat divergent roles for mycobacterial cells but all five have at this point been either tentatively (ESX-2/4) or definitively (ESX-1/3/5) linked to virulence in a *M. tuberculosis* model (Conrad et al., 2017; Abdallah et al., 2011; Tufariello et al., 2016; Pajuelo et al., 2021; Izquierdo Lafuente et al., 2021). Indeed, the primary role of ESX-1 and ESX5 seems to be modulation and evasion of host immune cells (Tiwari et al., 2019; Abdallah et al., 2011). ESX-1 is the original T7SS system discovered as RD1 and is necessary for Mycobacteria to infect their host (Tiwari et al., 2019). More specifically, *M. tuberculosis* ESX-1 has been implicated in phagosomal escape in macrophages and this is believed to be mediated

by the ESAT-6 effector, which causes membrane lysis at low pH (de Jonge et al., 2007; Conrad et al., 2017). ESX-5 has been similarly linked to virulence as avirulent strains of Mycobacteria typically lack ESX-5. ESX-5 has the largest repertoire of effectors of any T7SS (see below) and is essential for Mycobacterial viability as well as virulence (Ates et al., 2016; Di Luca et al., 2012). ESX-3 has similarly been linked to virulence but has also been found to play a critical role in iron homeostasis (Serafini et al., 2009). ESX-4 is the most ancestral of the five ESX systems and was initially found to be important for conjugation (Gray et al., 2016). Similar to ESX-1/3/5, ESX-4 has recently been found to have a role in Mycobacterial virulence as it is necessary for *Mycobacterium abscessus* infection but also for secretion of the NAD^+ glycohydrolase toxin CpnT in *M. tuberculosis* (see below) (Laencina et al., 2018; Pajuelo et al., 2021). The least understood of the five ESX systems is ESX-2. The only phenotype linked to ESX-2 thus far is phagosomal rupture during *M. tuberculosis* infection of macrophages, but this was only observed in conjunction with ESX-4 (Izquierdo Lafuente et al., 2021; Pajuelo et al., 2021).

There are five integral membrane components that make up an ESX secretion system: EccB, EccC, EccD, EccE, and MycP (Fig. 1.1) (Bitter et al., 2009; Houben et al., 2012). Recently there have been several papers reporting the cryogenic electron microscopy (cryo-EM) structure of several ESX apparatuses (Famelis et al., 2019; Poweleit et al., 2019; Bunduc et al., 2021; Beckham et al., 2021). In general, these studies suggest a 1:1:2:1 stoichiometry for the four Ecc components. Interestingly, an ultrastructure that includes the MycP component of ESX-5, shows that this T7SSa forms

a trimer of dimers, where each dimer consists of two copies of the 1:1:2:1 protomer with a single copy of MycP interacting with and stabilizing the outward facing portion of EccB (Bunduc et al., 2021). Structural and biochemical work has also suggested roles for the different components of T7SSa. EccB faces the mycobacterial periplasm and likely anchors T7SSa to the cell wall through a peptidoglycan binding fold (Bunduc et al., 2021). EccC is a FtsK-SpoIIIE family ATPase that energizes T7SSa and also plays a role in effector recognition (Rosenberg et al., 2015). EccD is a large membrane embedded protein and is thought to act as the scaffold that holds the other components of T7SSa together (Beckham et al., 2021). EccE contains non-functional glycosyltransferase domains but is ultimately thought to play a role in protomer stability (Famelis et al., 2019). Finally, MycP interacts with EccB to form a dome-shaped structure in the periplasm (Bunduc et al., 2021). MycP is a protease that has been shown to cleave the EspB effector of *M. tuberculosis*, however, proteolytic activity of MycP is not required for secretion and so it is possible that the main role of MycP is in stabilizing the trimer of dimers through its interaction with the periplasmic domain of EccB (Ohol et al., 2010; van Winden et al., 2016; Bunduc et al., 2021). In conjunction with the five membrane embedded components of T7SSa, there is also a single soluble ATPase called EccA. EccA is thought to act as an "instigator" of type VII secretion, facilitating secretion by passing effectors to the membrane bound ATPase EccC (Fig. 1.1) (Crosskey et al., 2020).

The recent advances in T7SSa macrostructure have led to a more thorough understanding of how these systems exports proteins. The first step of secretion is effector recognition, in which the C-terminal ATPase domain of EccC binds to the EsxB

(described in more detail below) effector (Rosenberg et al., 2015). This interaction stimulates multimerization of EccC which in turn activates the ATPase activity of the translocase. Although it is not entirely understood how multimerization of the system in its entirety occurs, the multimerized T7SSa adopts a hexameric arrangement (Rosenberg et al., 2015; Beckham et al., 2017; Famelis et al., 2019). The membrane pore of the T7SSa is formed by EccC in the cytoplasm and continues through EccB into the periplasm, while EccD and EccE play stabilizing roles (Famelis et al., 2019). ATP binding and hydrolysis then causes a conformational change in the apparatus that facilitates the export of the EccC-bound effectors into the periplasm (Rosenberg et al., 2015). Until very recently, it was not known how effectors transported into the periplasm by T7SSa then bypass the outer mycomembrane, but new evidence suggests that mycomembrane pores are formed by bona fide T7SSa effectors. In this model, some T7SSa effectors oligomerize in the mycomembrane to form pores that facilitates the export of other effectors secreted by the same ESX secretion system (Piton et al., 2020; Tak et al., 2021). This finding partially explains the fact that T7SSa effectors are generally co-secreted with other effectors and the deletion of a single effector can disrupt the function of an entire ESX system. The tuberculosis necrotizing toxin (TNT) effector CpnT is the best example of this phenomenon as its secretion depends on the co-secretion and outer membrane pore formation of EsxEF oligomers (Tak et al., 2021).

ESXs secrete several classes of effectors, all of which fall into the WXG100 superfamily of proteins. Canonical WXGs are ~100 amino acid helix-turn-helix proteins that have a Trp-x-Gly motif in the turn region and are ubiquitous throughout all T7SSs

(Poulsen et al., 2014). Prototypical examples of WXG proteins are 6 kDa early secretory antigenic target (ESAT-6, now called EsxA) and 10 kDa culture filtrate protein (CFP-10, now called EsxB), which heterodimerize with each other and are co-secreted by ESX-1 (Renshaw et al., 2005). EsxA and EsxB are known to be some of the most antigenic proteins secreted by *M. tuberculosis*, although their exact role in type VII secretion is not totally understood (Skjøt et al., 2000). Each ESX system has its own heterodimeric pair of WXG proteins that are secreted by the system and are also essential for the secretion of other effectors. WXG pairs are recognized by the C-terminal ATPase domain of EccC, an interaction that requires a conserved secretion motif (Rosenberg et al., 2015). The secretion motif consists of the WXG motif found in the turn region of EsxA-homologous proteins as well as an unstructured tail containing a Tyr-x-x-x-Asp/Glu (YxxxD/E) motif (collectively called the 'export arm') in the C-terminal region of EsxB-homologous proteins (Champion et al., 2006, Daleke et al., 2012a). In general, any perturbation to the YxxxD/E motif inhibits secretion of the EsxA:EsxB heterodimer (Daleke et al., 2012a).

ESX systems secrete other classes of WXG100 superfamily effectors including PE, PPE, and Esp effectors. Similar to EsxA:EsxB, PE and PPE proteins always form a heterodimeric pair and are co-secreted (Korotkova et al., 2014; Ekiert & Cox, 2014; Williamson et al., 2020). These proteins are primarily α-helical and are named for their Pro-Glu or Pro-Pro-Glu motif although PPE proteins often also have a WXG motif (Korotkova et al., 2014). PE:PPE pairs evidently have important functions for Mycobacterial cells as genes for these two families can represent up to 10% of the coding capacity of *M. tuberculosis* (Cole et al., 1998). Although it is likely that the vast array of

PE:PPE pairs in *Mycobacteria* have an equally vast number of functions for the cell, the PE:PPE proteins studied to date have been shown to facilitate nutrient uptake, subvert immune responses, or play a role in nutrient recycling (Wang et al., 2020; Sayes et al., 2012; Santucci et al., 2018). Unlike canonical WXG pairs like EsxA:EsxB, PE:PPE pairs often require a chaperone, called EspG, for secretion (Korotkova et al., 2014; Ekiert & Cox, 2014; Tuukkanen et al., 2019). Co-crystallization of the PE:PPE:EspG heterocomplex has shown that PPE has a conserved EspG binding face which allows an EspG protein to bind the PE:PPE pair and guide it to a specific ESX system (Korotkova et al., 2014; Ekiert & Cox, 2014). The vast majority of PE:PPE effectors are secreted by ESX-5 and are therefore bound and chaperoned by EspG5. However, it was recently shown that a PE:PPE dimer could be rerouted through ESX-1 by switching the PPE effector's EspG5 binding face to that of EspG1 (Phan et al., 2017). ESXs also secrete a third class of effectors called Esp proteins (not related to the EspG chaperones) with EspB being the best studied. These effectors form a four helix bundle reminiscent of a PE:PPE heterodimer but, unlike PE/PPE dimers, do not seem to interact with ESX-specific chaperones (Solomonson et al., 2015; Korotkova et al., 2015). Although the function of EspB is still somewhat speculative, it is thought that the protein oligomerizes to form a pore in the mycomembrane and facilitates the export of other effectors and the import of signalling lipids and/or DNA (Piton et al., 2020). It is interesting to note that both PE/PPE pairs and EspB maintain the YxxxD/E secretion motif required for EsxAB dimer secretion (Solomonson et al., 2015). As such, this motif is now referred to as the "general secretion signal" for T7SS.

Perhaps the most well studied virulence factor of *M. tuberculosis* is the tuberculosis necrotizing toxin (TNT) effector CpnT, which hydrolyzes the essential molecule NAD^+ (Sun et al., 2015). Although it is now established that CpnT is a T7SS effector, there is still a question as to which of the five ESX systems are required for its export. Using *M. tuberculosis* as a model, Pajuelo et al. observed that ESX-4 is required for the export of CpnT and for the surface accessibility of its TNT toxin domain (Pajuelo et al., 2021). This model requires the secretion of EsxEF dimers through ESX-4 that then form oligomeric pores in the outer membrane through which CpnT can be exported (Tak et al., 2021). They further showed that both ESX-1 and ESX-2 are required for membrane permeabilization of the phagosome and suggested that all three of the systems are therefore required for CpnT transport into the macrophage cytoplasm (Pajuelo et al., 2021). In contrast to these findings, Izquierdo-Lafuente et al. found that in a *Mycobacterium marinum* model, CpnT is secreted by ESX-5, although they also noted a requirement for ESX-1 and ESX-4 for intracellular secretion (Izquierdo Lafuente et al., 2021). This work also found that CpnT has the conserved YxxxD/E secretion motif common to WXG100 effectors and that this motif is necessary for export. CpnT therefore is a T7SSa secreted effector but which ESX system(s) is required for its export seems to be species specific.

While much research has gone into understanding how the different ESX systems impact the lifestyle, and especially the virulence, of *M. tuberculosis*, a related system was discovered in Firmicutes bacteria and has since proven to be equally important to various bacterial genera including *Staphylococcus*, *Streptococcus*, *Bacillus*, and *Enterococcus*. To

differentiate the T7SS of Actinobacteria and Firmicutes, the two systems are generally referred to as the T7SSa and the T7SSb, respectively.

T7SSb: a functionally diverse secretion system of Firmicutes

T7SSb was originally discovered in 2002 when it was observed that WXG100 genes were present in Firmicutes genomes including genetically tractable organisms like *Staphylococcus aureus* and *Bacillus subtilis* (Pallen, 2002). In the following years, it was also observed that these bacteria encode EccC homologues and ultimately that the WXG100 proteins are secreted by these novel systems. The best studied T7SSb is that of *S. aureus* and the system has been linked to both *S. aureus* virulence and interbacterial antagonism (Bowman et al., 2021). Both of these phenotypes have also been linked to T7SSb in the human commensal/opportunistic pathogen *Streptococcus intermedius* (Whitney et al., 2017, Hasegawa et al., 2017). More recently, *B. subtilis* and *Enterococcus faecalis* have also been shown to have an active T7SSb and these systems can both export antibacterial toxins (Kobayashi et al., 2021, Chatterjee et al., 2021). Although the interbacterial antagonism aspect of the T7SSb is at this point well-established, it is still unclear if T7SSb-mediated virulence is a direct effect of secreted proteins damaging host cells or simply a secondary effect of niche control through killing of commensal bacteria.

T7SSb is distinguished from T7SSa because, although functionally similar, the two secretion systems are composed of different sets of proteins (Tran et al., 2021). There are two ubiquitous components that are found in all functional T7SSs. The first conserved

component is the apparatus protein EccC/EssC (T7SSa/T7SSb), which is a FtsK-SpoIIIE family ATPase that energizes the systems (Zoltner et al., 2016; Jäger et al., 2018). The second conserved component is at least one (but possibly multiple) WXG effectors which are secreted by all T7SSs and are also essential for the secretion of other effectors (Poulsen et al., 2014). Except for EccC, the T7SSb apparatus is made up of entirely different proteins from T7SSa and beyond WXG proteins, the two systems secrete different classes of effectors.

The T7SSb apparatus is composed of four integral membrane proteins: EssA, EssB, EssC, and EsaA, all of which are essential for effector export (Fig. 1.1) (Aly et al., 2017; Tran et al., 2021). EssA is the smallest of the four proteins, and little is known about the role it plays in secretion. EssB is larger and has both a cytoplasmic pseudokinase domain and an extracellular domain (Zoltner et al., 2012). EssB has been shown to form extensive contacts with the other T7SSb subunits and therefore seems to play a scaffolding and stabilizing role (Tassinari et al., 2020). EssC, like EccC, is an ATPase that is predicted to drive effector recognition and secretion (Zoltner et al., 2016; Mietrach et al., 2020b). The EssC N-terminus has two forkhead-associated domains that interact with the pseudokinase domain of EssB (Tassinari et al., 2020). The EssC C-terminus consists of three ATPase domains and a domain of unknown function which, based on cryo-EM structures of the T7SSa, may actually be a fourth (albeit less conserved) ATPase domain (Famelis et al., 2019). Finally, EsaA has a large soluble domain flanked by N- and C-terminal transmembrane helices (Ahmed et al., 2018). The structure and function of this massive but poorly understood protein is the focus of

Chapter III. The T7SSb has an additional conserved core component, called EsaB, which is localized to the cytoplasm. EsaB is a ubiquitin-like protein that is essential for T7SSb secretion but the exact role it plays is unknown (Fig. 1.1) (Casabona et al., 2017).

Similar to T7SSa, all T7SSbs secrete WXG effectors but in the context of T7SSb, WXG proteins often homodimerize rather than heterodimerize (Poulsen et al., 2014; Sundaramoorthy et al., 2008). The function of T7SSb WXGs is not fully known but they are necessary for the secretion of other effectors and have been linked to modulating host immune factors and facilitating an intracellular lifestyle in *S. aureus* (Anderson et al., 2017; Korea et al., 2014). The T7SSb also secretes LXG effectors which, like PE/PPE/EspB effectors, are also part of the WXG100 superfamily of proteins. LXG proteins are polymorphic toxins named for their conserved N-terminal LXG domain which contains a Leu-x-Gly motif (Zhang et al., 2012). The C-termini of LXG effectors harbour functional toxin domains. To date, LXG toxins have been shown to have lipid II phosphatase, NAD-hydrolase, and membrane depolarizing activities (Whitney et al., 2017; Ulhuq et al., 2020). Like Gram-negative antibacterial effectors, LXG effectors are co-transcribed with immunity proteins that specifically block the toxicity of their corresponding effector (Whitney et al., 2017). Chapter II will explore the interaction between one of these immunity proteins, TipC, and its corresponding toxin, the lipid II phosphatase effector TelC from *S. intermedius*. The *S. aureus* effector EsaD is the only example of a non-WXG/LXG effector secreted by a T7SSb that has been studied to date. EsaD is a nuclease toxin that seemingly mediates both virulence and interbacterial antagonism (Cao et al., 2016, Ohr et al., 2017). EsaD toxicity is inhibited by an immunity

protein called EsaG which directly interacts with the C-terminal toxin domain of EsaD (Cao et al., 2016). While the N-terminal domain of EsaD lacks a Leu-x-Gly motif, it has been shown to interact with EsaE, a chaperone protein that facilitates the export of EsaD (Cao et al., 2016; Anderson et al., 2017). In Chapter IV, I begin structural and functional characterization of two new families of LXG-specific chaperones, although it is important to note that EsaE is unlikely to be structurally similar to these chaperones based on secondary structure analysis.

Research goals: towards a more detailed understanding of T7SSb function

While research into the structure and function of T7SSb proteins has advanced steadily over the past couple of years, there is still much to learn about how the secretion system functions on a molecular level. Through my doctoral work, I have attempted to address several of these knowledge gaps, and they are briefly prefaced below.

In chapter II, I focus on the interaction of the TipC1 immunity protein with its cognate effector TelC. My findings suggest that TipC1 is a membrane bound protein that faces the inner-wall-zone (IWZ) of the Gram-positive cell. The directionality of TipC1 is critical as its cognate effector, TelC, is toxic in the IWZ rather than the cytoplasm. TipC1 directly binds to TelC through its concave face and mutation of conserved residues on this interacting surface abrogates TipC1's ability to interact with TelC and protect against TelC-mediated toxicity. Importantly, my work on TipC1 also provided initial evidence that the T7SSb exports effectors through both the cell membrane and the cell wall in a single step, reminiscent of the one-step secretion systems of Gram-negative bacteria.

To further explore the T7SSb secretion mechanism, in Chapter III, I conduct a structural analysis of the apparatus protein EsaA. EsaA is the second largest T7SSb apparatus protein and our work suggests that it may form a conduit through which secretion occurs. In this work, I prove unequivocally that EsaA is an essential part of T7SSb. I also show that the soluble domain of EsaA faces outwards. Finally, by way of X-ray crystallography, I present the first structure of the EsaA soluble domain. I find that EsaA forms an extended structure that dimerizes *in crystallo*, *in vitro*, and *in vivo*. Importantly, this work shows that the most extended portion of EsaA corresponds to a phage receptor, suggestive of it being surface exposed. Together, these data support the hypothesis that EsaA forms a conduit through which effector secretion occurs. We are hopeful that future cryo-EM studies will further substantiate this hypothesis.

In Chapter IV, I switch my focus from how type VIIb secretion occurs to how effectors are specifically recognized for secretion. In this work, I characterize the Lap1 and Lap2 families of proteins, which I found interact directly with the LXG domain of T7SSb effectors. I find that Lap1 and Lap2 are both necessary for secretion of LXG effectors but are not secreted themselves. Using X-ray crystallography and structural modelling, I show that both Lap1 and Lap2 proteins form helix-turn-helix folds that are highly reminiscent of WXG proteins. I also report a conserved C-terminal motif in Lap1 that is necessary for LXG export and therefore represents a possible recognition motif for LXG effectors. From this work we suggest a model of secretion where step one is the formation of an LXG effector-Lap1-Lap2 pre-secretion complex.

Figures

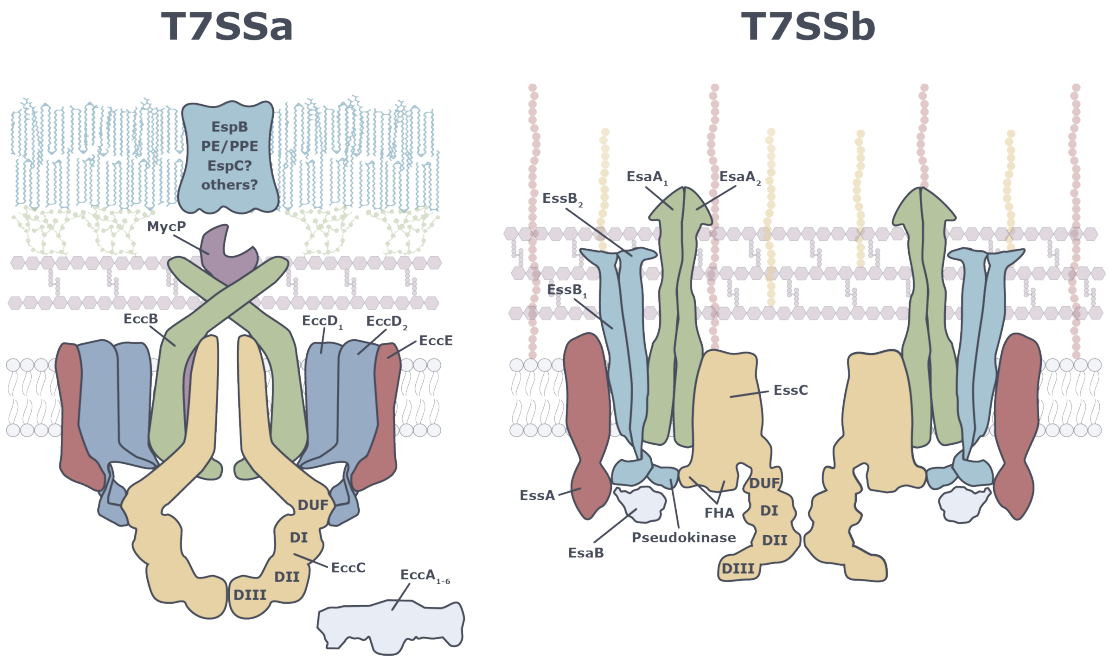


Figure 1.1. T7SSa and T7SSb apparatuses require divergent sets of proteins to facilitate protein secretion across the cell envelope. (Left) The T7SSa apparatus is comprised of EccB, shown in green; EccC, yellow; EccD, blue; EccE, red; MycP, purple; and EccA, white. Schematic is based on the recent structures of the *M. smegmatis* ESX-3 and the *M. tuberculosis* ESX-5 secretion systems, which show that the T7SSa core complex has a 2:2:4:2:1 (EccB:EccC:EccD:EccE:MycP) stoichiometry (Famelis *et al.*, 2019, Bunduc *et al.*, 2020). EccC contains three C-terminal ATPase domains (DI-DIII) as well as a domain of unknown function (DUF) that is proposed to be a fourth ATPase domain. A mycomembrane pore may be needed for T7SSa substrates to reach the extracellular milieu and PE/PPE proteins as well as EspB and EspC have been posited to assist in fulfilling this role. (Right) The T7SSb apparatus consists of five peripheral and integral membrane components: EssA, red; EssB, blue; EsaB, white; EssC, yellow; and

EsaA, green. EssB has a wing-shaped periplasmic domain and a cytoplasmic pseudokinase domain involved in numerous protein-protein interactions. EssC has two N-terminal forkhead-associated domains (FHA) and three C-terminal ATPase domains as well as a DUF. EsaA is potentially analogous to EccB because it contains a large extracellular domain. In the case of the monoderm Firmicutes where no mycomembrane pore is required, the extracellular domain of EsaA is speculated to form a cell wall spanning conduit through which T7SSb substrates are exported.

Chapter II – Molecular basis for immunity protein recognition of a type VII secretion system exported antibacterial toxin

Preface

The work presented in this chapter was previously published in the following study:

Klein, T.A., Pazos, M., Surette, M.G., Vollmer, W., and Whitney, J.C. (2018). Molecular basis for immunity protein recognition of a type VII secretion system exported antibacterial toxin. *Journal of Molecular Biology* 430 (21): 4344-4358.

Author contributions: T.A.K. and J.C.W. designed the research; M.G.S provided strains and reagents; T.A.K., M.P. and J.C.W. performed the research; T.A.K., M.P., W.V. and J.C.W. analyzed the data and T.A.K. and J.C.W. wrote the paper. All authors provided feedback and revised the manuscript.

Permission has been granted by the publishers to reproduce the material from this study.

Abstract

Gram-positive bacteria deploy the type VII secretion system (T7SS) to facilitate interactions between eukaryotic and prokaryotic cells. In recent work, we identified the TelC protein from *Streptococcus intermedius* as a T7SS-exported lipid II phosphatase that mediates interbacterial competition. TelC exerts toxicity in the inner wall zone of Gram-positive bacteria; however, intercellular intoxication of sister cells does not occur because they express the TipC immunity protein. In the present study, we sought to characterize the molecular basis of self-protection by TipC. Using sub-cellular localization and protease protection assays, we show that TipC is a membrane protein with an N-terminal transmembrane segment and a C-terminal TelC-inhibitory domain that protrudes into the inner wall zone. The 1.9Å X-ray crystal structure of a non-protective TipC paralogue reveals that the soluble domain of TipC proteins adopts a crescent-shaped fold comprised of three α-helices and a seven-stranded β-sheet. Subsequent homology-guided mutagenesis demonstrates that a concave surface formed by the predicted β-sheet of TipC is required for both its interaction with TelC and its TelC-inhibitory activity. *S. intermedius* cells lacking the *tipC* gene are susceptible to growth inhibition by TelC delivered between cells; however, we find that the growth of this strain is unaffected by endogenous or overexpressed TelC even though the toxin accumulates in culture supernatants. Together, these data indicate that the TelC-inhibitory activity of TipC is only required for intercellularly-transferred TelC and that the T7SS apparatus transports TelC across the cell envelope in a single step, bypassing the cellular compartment in which it exerts toxicity en route.

Introduction

Bacteria employ a variety of mechanisms to transport macromolecules across membranes. One of the ways this process is accomplished in Gram-positive bacteria is through a multi-subunit membrane protein complex known as the type VII secretion system (T7SS) (Abdallah et al., 2007). T7SSs are best studied in the phylum Actinobacteria where they have been shown to facilitate the transport of molecules involved in a wide array of biological processes. For example, the mycobacterial ESX-1, ESX-3 and ESX-4 T7SSs have been implicated in the lysis of host cell membranes, siderophore-mediated iron uptake and conjugal DNA transfer, respectively (Conrad et al., 2017; Siegrist et al., 2009; Gray et al., 2016). The T7SS has also been characterized in the low G+C Gram-positive phylum Firmicutes, which possesses an evolutionarily distant subfamily of this pathway referred to as T7SSb (Abdallah et al., 2007). The T7SSb apparatus is comprised of fewer protein subunits than Actinobacterial T7SSs and functions to mediate protein export from the cell (Unnikrishnan et al., 2017). Among T7SSb-containing bacteria, the *ess* locus of *Staphylococcus aureus* is the most extensively characterized. This system exports four small non-enzymatic proteins of unknown function named EsxA, EsxB, EsxC and EsxD, which belong to the WXG100 family of T7SS effectors (Burts et al., 2005; Anderson et al., 2013). Additionally, the large nuclease toxin EsaD is exported in a T7SS-dependent manner, and phenotypic characterization of *S. aureus* strains lacking the *esaD* gene indicate that this toxin contributes to both the bacteria and host cell-targeting capabilities of this pathway (Ohr et al., 2017; Cao et al., 2016).

Recently, we demonstrated that *Streptococcus intermedius* uses its T7SS for antagonistic bacterial cell-cell interactions, further substantiating the bacteria-targeting capability of the T7SSb pathway (Whitney et al., 2017). *S. intermedius* is a commensal bacterium found within the densely populated microbial flora of the human oral cavity and is also an opportunistic pathogen (Macey et al., 2001). In addition to the WXG100 protein EsxA, we demonstrated that the T7SS of *S. intermedius* exports three effector proteins named TelA, TelB and TelC (Whitney et al., 2017). The Tel proteins belong to the large and broadly distributed LXG family of polymorphic toxins and the discovery of these effectors provided the first experimental evidence that this family of toxins transits the T7SS (Zhang et al., 2012). While the mode of action of TelA is unknown, biochemical characterization of TelB demonstrated that it exerts toxicity by degrading the electron carrying dinucleotide NAD^+ whereas TelC functions as a phosphatase that cleaves peptidoglycan precursor lipid II.

Concomitant with our discovery of the Tel proteins was the finding that each of these effectors is encoded in close proximity to a gene that encodes a toxin-specific immunity protein (Whitney et al., 2017). For example, TelA and TelB are toxic in the bacterial cytoplasm and their cognate immunity proteins, TipA and TipB, confer immunity to their respective toxins when expressed in this cellular compartment. Furthermore, TelB-expressing strains of *S. intermedius* exhibit a fitness advantage when grown in co-culture with *S. intermedius* strains lacking TipB (Whitney et al., 2017). These observations suggest that cytoplasmic immunity proteins protect bacteria from both self-produced toxins and toxins delivered by sister cells via the T7SS. Since TipA

and TipB were not identified as substrates of the T7SS, cytoplasmic TelA-TipA and TelB-TipB complexes are presumably dissociated prior to toxin export as has been observed for other interbacterial polymorphic toxin delivery systems (Li et al., 2012).

TelC is distinct from characterized Gram-positive antibacterial toxins because it acts in the inner wall zone when delivered into target bacteria by the T7SS (Whitney et al., 2017). Consequently, the TelC-specific immunity protein TipC1 may differ from TipA and TipB in that it likely localizes to this cellular compartment to enable it to confer immunity to TelC. If this prediction is true, TipC1 would be physically separated from its cognate effector in TelC-producing cells by the plasma membrane. T7SS effector translocation across the plasma membrane is catalyzed by the FtsK/SpoIIIE-like motor ATPase EssC (Rosenberg et al., 2015); however, it is not known if the T7SS apparatus additionally facilitates effector transport across the thick Gram-positive cell wall. Thus, it is unclear if TipC1 is required for protection from self-produced TelC or if it is only needed to confer immunity to intercellularly delivered TelC.

In the present work, we sought to uncover the site of action and mode of TelC inhibition by the TipC immunity protein. To this end, we used subcellular localization and protease accessibility assays to show that TipC is a membrane protein with an extracellular TelC-inhibitory domain. We then determined the structure of a non-protective TipC paralogue, which allowed for homology modelling of TipC. Mutagenesis analysis informed by this structural model suggests that TipC inhibits TelC toxicity via a concave surface formed by a seven-stranded β-sheet. Finally, mutational inactivation of *tipC* does not render *S. intermedius* cells susceptible to self-produced TelC, even though

the toxin is exported from the cell via the T7SS. Taken together, these data point to a model in which TipC is required for protection from competitor delivered but not self-produced TelC toxin.

Results

TipC localizes to the plasma membrane via an N-terminal transmembrane domain.

We previously showed that the soluble region of TipC is sufficient to inhibit the toxic lipid II phosphatase activity of TelC *in vitro* (Whitney et al., 2017). In these biochemical assays, a truncated form of TipC that excluded its hydrophobic N-terminus was employed in order to reduce TipC aggregation in aqueous buffer and consequently, a functional role for this region of the protein was not determined (Fig. 2.1A). Lipid II exists in both the inner and outer leaflet of the plasma membrane; however, for reasons that are unclear, TelC only exerts toxicity when targeted to the inner wall zone (Whitney et al., 2017). We hypothesized that the ability of TipC to effectively neutralize a toxin that acts in the inner wall zone on a membrane-embedded substrate arises because the protein itself also localizes to the plasma membrane. To test this, we performed subcellular fractionation experiments on *S. intermedius* B196 (Si^{B196}) cells expressing vesicular stomatitis virus glycoprotein G (VSV-G) epitope tagged TipC (TipC-V). The characterized streptococcal proteins manganese-dependent superoxide dismutase (SodA) and lantibiotic Smb receptor-like function in streptococci (LsrS) were used as cytoplasmic and membrane protein fractionation controls, respectively (Biswas et al., 2014; Crump et al., 2014). Consistent with our hypothesis, we found that TipC localizes

to the membrane fraction (Fig. 2.1B). Furthermore, this localization was dependent on the hydrophobic N-terminus of TipC because a truncated form of TipC lacking this region of the protein ($TipC_{\Delta TMD}$) was found exclusively in the cytosol.

We next examined the orientation of the TelC-inhibitory domain of TipC in the plasma membrane. Our previous finding that TipC abrogates toxicity caused by Sec translocon-targeted TelC suggested that this domain exists in the inner wall zone (Whitney et al., 2017). To test this prediction, we performed protease accessibility assays on spheroplasts generated via lysozyme digestion of Si^{B196} cells expressing TipC or $TipC_{\Delta TMD}$ (Fenton et al., 2018). As shown in Figure 2.1C, only full-length TipC was readily degraded by the added protease. In contrast, cytoplasmic $TipC_{\Delta TMD}$ was susceptible to proteolysis only after spheroplast rupture by detergent. Together, these data are consistent with the prediction that TipC is a membrane protein with a TelC-inhibitory soluble domain that protrudes into the inner wall zone.

telC-tipC operons harbour multiple *tipC* paralogous genes.

Having established a functional role for the N-terminal TMD of TipC, we next sought to identify the region of its C-terminal domain responsible for its TelC-inhibitory activity. Similar to the T7SS-exported Tel proteins, antibacterial toxins delivered by other pathways involved in interbacterial antagonism possess cognate immunity proteins that protect toxin-producing bacteria from the activity of their own toxins and/or toxins delivered intercellularly by sister cells (Russell et al., 2012). These immunity proteins are highly specific towards their cognate toxin as pairs of homologous immunity proteins

with greater than 50% identity between them have been shown to have opposing abilities to neutralize a given toxin (Russell et al., 2013). We sought to exploit this observation to identify amino acid residues critical for the TelC-inhibitory activity of TipC by locating variable positions between TipC homologous sequences. BLASTp analysis of the NCBI non-redundant sequence database identified 286 TipC homologous proteins whose distribution is restricted to species belonging to the order *Lactobacillales*. Examination of the genomic context of *tipC* ORFs revealed that the vast majority of these genes exist in operons with similar synteny to that of Si^{B196} (Fig. 2.2). Additionally, we noted two examples of *tipC* genes found in gene clusters that may represent heterogeneous arrays of immunity genes as defined by Aravind and colleagues (Zhang et al., 2012). Of particular utility to this work, we also found that the majority of *tipC*-containing bacteria possess multiple *tipC* paralogous genes. Si^{B196} possesses one such *tipC* paralogous gene (SIR_1486), which encodes a protein with 58% identity to TipC. To disambiguate these two proteins, we henceforth refer to TipC (SIR_1488) as TipC1 and the protein encoded by SIR_1486 as TipC2.

<u>TipC2 does not protect against TelC-mediated toxicity.</u>

Given the high degree of homology between TipC1 and TipC2, and the observation that slight divergence in immunity protein sequence is sufficient to abrogate toxin-inhibitory activity (Russell et al., 2013), it seemed reasonable that this protein could guide our identification of TipC1 residues that mediate TelC inhibition. Toward this end, we first tested whether TipC2 could protect Si^{B196} from TelC-based toxicity. In contrast

to cells expressing TipC1, we found that TipC2 expression could not prevent toxicity caused by constitutive expression of the TelC toxin domain ($TelC_{tox}$) targeted to the inner wall zone of Si^{B196} (ss-$TelC_{tox}$) (Fig. 2.3A). To rule out that the failure of TipC2 to inhibit TelC activity is a result of inherent instability of the protein, we next performed nickel affinity co-purification experiments using his_6-tagged $TelC_{tox}$ co-expressed with VSV-G epitope tagged TipC1 or TipC2. To simplify the purification process, we used $TipC1_{\Delta TMD}$ and a similarly truncated TipC2 variant that also lacks its N-terminal transmembrane domain ($TipC2_{\Delta TMD}$) because we previously showed that this region of TipC1 is not required for its ability to inhibit the toxic lipid II phosphatase activity of TelC *in vitro* (Whitney et al., 2017). As shown in Figure 2.3B, these experiments demonstrated that although both $TipC1_{\Delta TMD}$ and $TipC2_{\Delta TMD}$ accumulate to substantial levels in cells, only $TipC1_{\Delta TMD}$ is capable of interacting with TelC. We further expanded this line of inquiry to an organism possessing more than two *tipC* paralogous genes. *Streptococcus gallolyticus* ATCC 43143 contains four adjacently encoded TipC proteins (*Sg*TipC1-4). Using bacterial two-hybrid analysis, we found that only *Sg*$TipC1_{\Delta TMD}$ is capable of interacting with *S. gallolyticus* TelC (*Sg*TelC) (Fig. 2.3C). Together, these results indicate that in two different bacteria TipC1 proteins, but not downstream encoded paralogous TipCs, possesses the molecular determinants for cognate TelC inhibition.

X-ray crystal structure of $TipC2_{\Delta TMD}$.

Our finding that $TipC1_{\Delta TMD}$, but not $TipC2_{\Delta TMD}$, interacts with and confers immunity to TelC substantially reduces the number of potential residues that could be

involved in TelC inhibition. However, structure prediction algorithms were unable to generate a high confidence model of TipC1$_{\Delta TMD}$ that would allow us to predict which candidate residues are surface exposed and thus be more likely to interact with and inhibit TelC. Crystallization efforts failed to yield diffraction quality crystals of TipC1$_{\Delta TMD}$ or TelC–TipC1$_{\Delta TMD}$ complex; however, TipC2$_{\Delta TMD}$ readily crystallized in the space group *C2* and we were able to determine its X-ray crystal structure to 1.8Å resolution using selenium-incorporated protein and the selenium single wavelength anomalous dispersion technique (Hendrickson et al., 1990). The resulting electron density maps allowed for complete model building of TipC2$_{\Delta TMD}$ (residues 23-203) and a vector-encoded proline residue derived from the linker region connecting a his$_6$-tag to the N-terminus of TipC2$_{\Delta TMD}$. The final model was refined to an R_{work}/R_{free} of 17.0% and 19.4%, respectively (Table 2.1).

TipC2$_{\Delta TMD}$ adopts a mixed α/β fold consisting of three α-helices and seven β-strands that fold together to give the protein a distinct crescent-shaped appearance (Fig. 2.4A). This shape is characterized by a concave surface formed by a seven-stranded β-sheet and a convex surface generated by the positions of three α-helices. Using the DALI webserver to compare our structure with all deposited structures in the PDB, we determined that TipC2$_{\Delta TMD}$ does not bear strong resemblance to proteins of known structure (Holm & Laakso, 2016). Many of the top scoring proteins from this analysis were outer membrane proteins from Gram-negative bacteria whose β-strands loosely resemble the β-sheet of TipC2$_{\Delta TMD}$. For example, the amyloid secretion protein FapF

from *Pseudomonas sp.* UK4 and the oligogalacturonate-specific porin KdgM from *Dickeya dadantii* superimpose with TipC2$_{\Delta TMD}$ with Cα RMSDs of 4.4Å and 3.9Å, respectively, over 91 equivalent Cα positions. Also identified in this analysis was the polo box 1 (PB-1) domain of ZYG-1 Plk4 kinase from *C. elegans* (Cα RMSDs of 2.8Å over 82 equivalent Cα positions). Though also functionally unrelated to TipC2, this crescent-shaped domain mediates a protein-protein interaction with the centriole duplication protein SPD-2 via its concave surface suggesting that the equivalent surface on TipC1 may interact with TelC (Shimanovskaya et al., 2014).

Because TipC2 does not protect cells from TelC-mediated toxicity, we next employed the I-Tasser structure threading server to generate a homology model of TipC1$_{\Delta TMD}$ (Fig. 2.4B) (Yang et al., 2015). The resulting TipC1$_{\Delta TMD}$ model (residues 23-204) had a template modelling score of 0.65, indicating that the probability that our TipC1$_{\Delta TMD}$ model has the same overall topology and fold as TipC2$_{\Delta TMD}$ is greater than 95% (Zhang & Skolnick, 2004; Xu & Zhang, 2010). Additionally, circular dichroism spectroscopy demonstrates that TipC1$_{\Delta TMD}$ and TipC2$_{\Delta TMD}$ have very similar secondary structure composition (Fig. S2.1). We next mapped the amino acids that vary between TipC1 and TipC2 onto the surface of the TipC1$_{\Delta TMD}$ model, restricting our selection to amino acid R-groups with differing polarity (Fig. 2.5A). This analysis revealed that the majority of conserved residues are found on the convex surface while the variable residues were predominantly found on the concave surface. Taken together with our finding that TipC2$_{\Delta TMD}$ does not interact with TelC$_{tox}$, these findings support the idea that the concave surface of TipC1 facilitates its interaction with TelC.

The predicted concave surface of TipC1 harbors the molecular determinants for TelC binding.

To dissect the interaction between $TipC1_{\Delta TMD}$ and TelC, we next performed homology model-guided mutagenesis on $TipC1_{\Delta TMD}$. A structure of a T7SS effector-immunity pair has not yet been determined; however, a number of co-crystal structures exist of effector–immunity complexes from Gram-negative polymorphic toxin systems (Tang et al., 2018; Beck et al., 2014). These structures show that the buried surface area between effectors and their cognate immunity proteins is substantial, typically exceeding 1000 $Å^2$, and thus these interactions may be difficult to disrupt by a conservative mutagenesis approach. Therefore, we mutated surface-exposed hydrophobic and small hydrophilic residues to the large, hydrophilic amino acid glutamine whereas charged amino acids were substituted with a residue of opposite charge. Each site-specific $TipC1_{\Delta TMD}$ variant was co-expressed with his_6-tagged $TelC_{tox}$ and binding was assessed via pull-down analysis. Importantly, all $TipC1_{\Delta TMD}$ variants tested expressed to comparable levels in *E. coli*, indicating that these amino acid substitutions did not adversely impact the stability of the protein (Fig. 2.5B). In line with our structural analyses, we found that site-specific substitution of residues on the convex surface of $TipC1_{\Delta TMD}$ had no effect on the ability of the protein to interact with $TelC_{tox}$. In contrast, mutation of arginine 56 (R56E), phenylalanine 71 (F71Q), arginine 87 (R87E), lysine 93 (K93E) and arginine 96 (R96E), all of which lie on the $TipC1_{\Delta TMD}$ concave surface, substantially reduced $TelC_{tox}$ binding (Fig. 2.5A and Fig. 2.5B).

We next selected two of the identified TipC1$_{\Delta TMD}$ point mutants defective in TelC$_{tox}$ binding, F71Q and K93E, and tested if these variants could rescue Si^{B196} cells from TelC-based toxicity. Individually, we found that these TipC variants exhibited a partial reduction in their ability to protect cells from the toxic activity of TelC while a TipC1 variant bearing both of these amino acid substitutions displayed a substantially greater reduction in TelC-neutralizing capability (Fig. 2.5C). Consistent with these findings, we found that only the TipC1$_{\Delta TMD}$ double mutant lacked no inhibitory activity towards the lipid II phosphatase activity of TelC$_{tox}$ (Fig. 2.5D and Fig. 2.5E). This defect in TelC inhibition by the TipC1$_{\Delta TMD}$ double mutant is not due misfolding of the protein as its circular dichroism spectrum was indistinguishable from wild-type TipC1$_{\Delta TMD}$ (Fig. S2.2). Together, these data indicate that the concave surface of TipC1 is required for direct inhibition of the toxic lipid II phosphatase activity of TelC.

<u>TelC bypasses the inner wall zone via the T7SS in TelC-producing cells.</u>

Having established that TipC1 is a membrane protein with a soluble TelC-inhibitory domain that exists in the inner wall zone, we next wanted to exploit the unique site of action of TelC to gain insight into the export mechanism of the T7SS. Our prior finding that TelC is toxic to both *Staphylococcus aureus* and Si^{B196} cells when artificially targeted to the Sec translocon but not when milligram quantities of purified, active TelC toxin are added to susceptible cells suggests that the Gram-positive cell wall prevents the diffusion of TelC between the extracellular milieu and the inner wall zone (Whitney et al., 2017). Taking these observations into consideration, we posited that the T7SS

apparatus likely facilitates the export of effector proteins across the entire Gram-positive cell envelope in a single step. In this model, deletion of *tipC1* would be expected to have no detrimental effect on Si^{B196} growth in liquid media because TelC and TipC1 would be physically separated by the plasma membrane in toxin-producing cells and the T7SS would allow TelC to bypass the inner wall zone during export. Importantly, T7SS-dependent intercellular intoxication would not occur because this requires growth on a solid surface (Whitney et al., 2017). In contrast, if the T7SS only functions to export TelC from the cytoplasm to the inner wall zone, a *tipC1* deficient strain would likely be susceptible to intoxication by self-produced TelC. To distinguish between these two possibilities, we generated a Si^{B196} strain lacking *tipC* genes and assessed whether this strain is susceptible to TelC-mediated toxicity by comparing its growth rate in liquid monoculture to that of its parent strain (Fig. 2.6A). Under these conditions, the immunity-deficient strain showed no significant growth impairment even though substantial amounts of the TelC toxin could be detected in culture supernatants (Fig. 2.6B). To rule out the possibility that endogenous levels of TelC are insufficient to observe intoxication by self-produced toxin, we also employed the plasmid-based expression system used for our Sec translocon-targeting TelC toxicity assays to express TelC in our immunity deficient strain. Despite elevated levels of TelC accumulation in culture supernatants, this strain also exhibited no measurable growth defect in monoculture compared to immunity-expressing strains (Fig. 2.6B and Fig. 2.6C). When contrasted with our previous observation that TelC is toxic when targeted to the inner wall zone via a sec leader peptide (Whitney et al., 2017), these data suggest that the T7SS apparatus forms a

continuous channel that facilitates TelC export from the cytoplasm into the extracellular milieu in a single step (Fig. 2.6D).

DISCUSSION

This study describes the first biochemical characterization of a T7SS immunity protein. We have shown that TipC1 is a membrane protein with a soluble domain that localizes to the inner wall zone and is responsible for its TelC-inhibitory activity. Furthermore, using structural and informatic approaches, we identified a concave surface on TipC1 that mediates its direct interaction with TelC. By showing the dispensability of TipC1 in TelC-producing cells, we also provide evidence that T7SS effectors bypass the inner wall zone as they transit the secretory apparatus.

TipC1 is distinct from the other identified T7SS immunity proteins TipA, TipB and EsaG in that it neutralizes a toxin that acts from outside the cell. In Gram-negative bacteria, the antibacterial type VI secretion system (T6SS) has been shown to deliver toxins into the periplasm that similarly disrupt cell surface structures (Russell et al., 2014). For example, the T6SS-delivered toxin Tse1 is a peptidoglycan hydrolase that, like TelC, possesses a cognate immunity determinant (Russell et al., 2011). This immunity protein, named Tsi1, is a soluble periplasmic protein that inhibits Tse1 despite not being anchored to the cellular structure that it protects, presumably because the confines of the Gram-negative periplasm allow Tsi1 to accumulate to levels that confer resistance to Tse1-mediated toxicity (Chou et al., 2012). Our finding that TipC1 is anchored to the plasma membrane not only increases the proximity of its TelC-inhibitory domain to the

lipid II substrate of TelC but also prevents its diffusion into the extracellular milieu through the estimated 50 kDa molecular weight cut-off pores of the peptidoglycan layer (Demchick & Koch, 1996). Though peptidoglycan hydrolyzing toxins with cognate immunity proteins have yet to be identified in Gram-positive bacteria, should these toxins exist, the diffusion of their associated immunity proteins away from the cell could similarly be prevented via covalent tethering to the cell wall via an LPXTG sorting motif (Navarre & Schneewind, 1999).

Like TelC, the Colicin M family of proteins are antibacterial toxins with lipid II phosphatase activity (El Ghachi et al., 2006). Colicins differ from T7SS-exported toxins in that they act between closely related Gram-negative bacteria and they do not require a specialized secretion system for delivery; however, they are similar in that they possess cognate immunity proteins that confer resistance to toxin activity (Olschläger et al., 1991). The structure of a colicin M immunity protein (Cmi) from *Escherichia coli* has been solved in both monomeric and domain-swapped dimeric states (Usón et al., 2012; Gérard et al., 2011). The Cmi dimer is approximately the same molecular weight as TipC; however, it does not bear any significant structural similarity. Furthermore, its overall shape is tetragonal, in contrast to the crescent-shaped appearance of TipC. The weak interaction between Colicin M and Cmi *in vitro* has made mapping their interaction interface challenging and thus it is unclear if the residues responsible for this interaction cluster to a discrete area of Cmi in a manner that is analogous to what we have shown here for TipC1. A lipid II phosphatase–immunity protein co-crystal structure is needed to

provide further mechanistic insight into how this family of enzymes is inactivated by proteinaceous inhibitors.

We exploited inability of the TipC2 protein to inhibit TelC-mediated toxicity to identify TipC1 amino acids critical for its function. However, the observation that many *telC*-containing bacteria possess additional *tipC* genes whose protein products do not interact with the TelC protein of the same organism raises the question of what the function of these genes is. One intriguing possibility is that these additional genes confer immunity to TelC toxins produced by other bacterial species. If this is indeed the case, then these bacteria would be resistant not only to TelC delivered by sister cells but also from divergent TelC toxins delivered by other species of bacteria occupying the same niche. Lending further support to this hypothesis, we identified several bacteria that possess 'orphan' *tipC* genes, which presumably exist to provide protection from intercellularly delivered TelC toxins.

The dispensability of *tipC* immunity genes in TelC-producing strains coupled with our observation that TelC targeted to the Sec translocon is toxic but TelC targeted to the T7SS is not, suggests that the T7SSb secretion apparatus not only exports its substrates across the plasma membrane but also the peptidoglycan layer. One way this might be accomplished is by a continuous proteinaceous channel formed by the structural components of the T7SS apparatus. To date, the best characterized T7SS structural subunit is the EssC ATPase, which exports proteins across the plasma membrane via a mechanism that requires homo-multimerization (Rosenberg et al., 2015). However, the other structural components of the T7SSb pathway, such as EsaA, EssA and EssB, are

less well characterized and it remains to be determined if the complex formed by these proteins forms a channel that penetrates the peptidoglycan sacculus (Aly et al., 2017). Recently, a 'needle-like' structure was shown to be formed by the EspC protein of the mycobacterial T7SS (Lou et al., 2017; Ates & Brosch, 2017); however, a homologous protein does not exist in Firmicutes, perhaps because of the substantial differences in cell envelope architecture between Actinobacteria and Firmicutes. Our data provide evidence that a functionally analogous structure may be formed by the T7SSb system; however, the protein subunits comprising such an assembly remain to be identified. Ultimately, visualization of an intact T7SSb apparatus is required in order to unequivocally demonstrate the existence of a transenvelope complex.

Figures

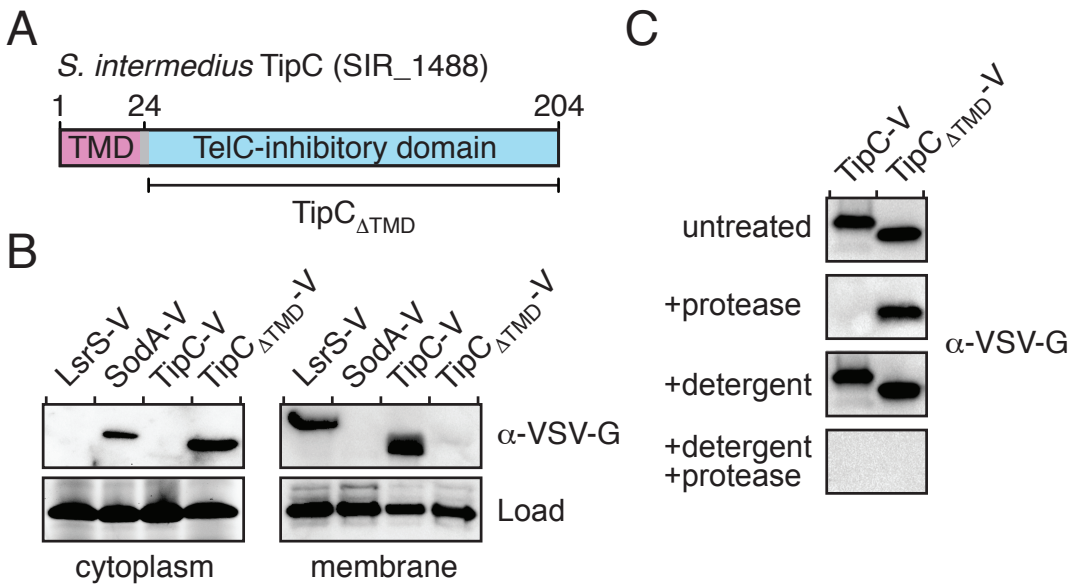


Figure 2.1: TipC is a surface exposed membrane protein. (A) Domain organization of TipC from *S. intermedius* B196. The boundaries for the TelC-inhibitory domain ($TipC_{\Delta TMD}$) and the predicted transmembrane domain (TMD) are indicated. (B) TipC1 is anchored to the plasma membrane via its N-terminal TMD. Western blot analysis of the cytoplasmic and membrane fractions of *S. intermedius* B196 strains expressing the indicated VSV-G epitope (V) tagged proteins. SodA-V and LsrS-V are cytoplasmic and membrane protein controls, respectively. Stain-Free detection (Bio-Rad) was used to ensure equal loading between samples. (C) The TelC-inhibitory domain of TipC is surface exposed. Western blot analysis of *S. intermedius* B196 spheroplasts expressing TipC-V or $TipC_{\Delta TMD}$-V. Spheroplasts were treated with Proteinase K (protease), Triton X-100 (detergent) or both and compared to an untreated control.

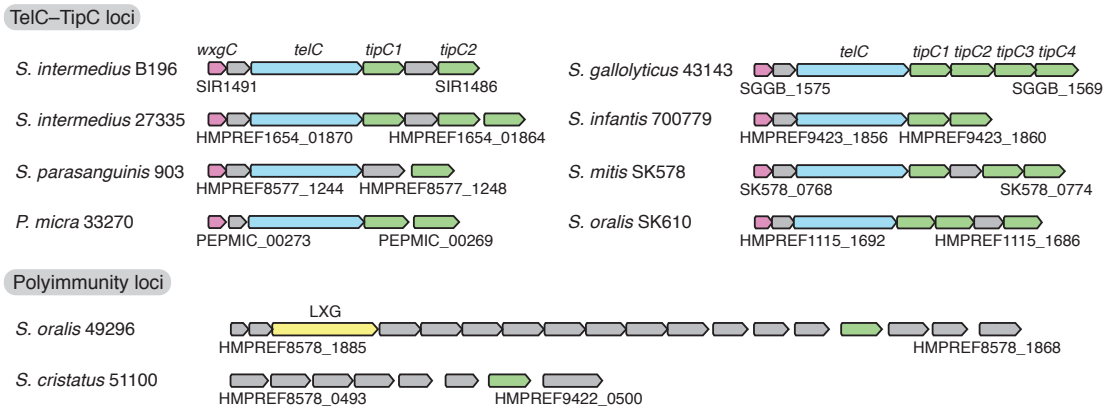


Figure 2.2: *telC* gene clusters possess multiple *tipC* paralogous genes. Genomic context of *tipC* genes from representative Firmicute species. Genes are colored according to homology and by known or predicted function of the encoded protein (TelC-interacting chaperones, purple; TelC toxins, blue; TipC immunity proteins, green; uncharacterized LXG toxin, yellow; other, grey).

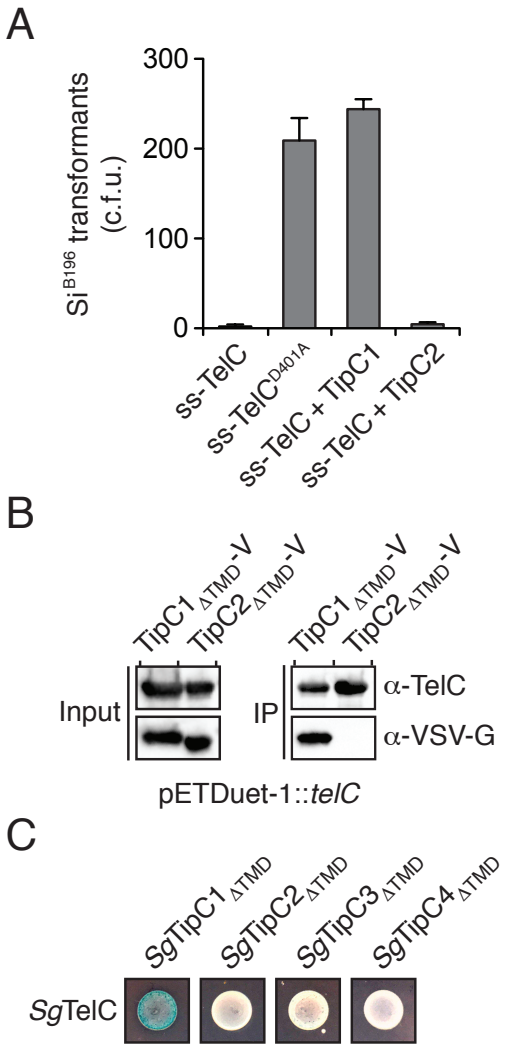


Figure 2.3: TipC2 does interact with TelC or confer immunity to TelC-mediated toxicity. (A) Number of *S. intermedius* B196 colonies after transformation with equimolar amounts of a plasmid constitutively expressing the indicated proteins. TelC fused to a Sec signal peptide (ss-TelC) and an inactive variant thereof (ss-TelC^{D401A}) serve as positive and negative controls, respectively. Details on the construction of these plasmids has been described previously (Whitney et al., 2017). Error bars represent ± SD (n = 3). (B) TipC2$_{\Delta TMD}$ does not interact with TelC. VSV-G epitope tagged TipC1$_{\Delta TMD}$

(TipC1$_{\Delta TMD}$-V) and TipC2$_{\Delta TMD}$ (TipC2$_{\Delta TMD}$-V) were co-expressed with his$_6$-tagged TelC and assessed for copurification by western blot analysis. (C) Only the *telC* adjacent *tipC* gene of *S. gallolyticus* ATCC 43143 encodes a protein (*Sg*TipC1) capable of interacting with the TelC paralogous protein (*Sg*TelC) from this organism. Bacterial two-hybrid analysis of *Sg*TelC and each of the four TipC paralogous proteins from *S. gallolyticus* ATCC 43143. *Sg*TelC was fused to the T25 fragment of adenylate cyclase and co-expressed with each TipC paralogous protein fused to the T18 fragment. Blue color indicates a protein-protein interaction. A schematic of the *S. gallolyticus* ATCC 43143 *telC-tipC* gene cluster can be found in Figure 2.2.

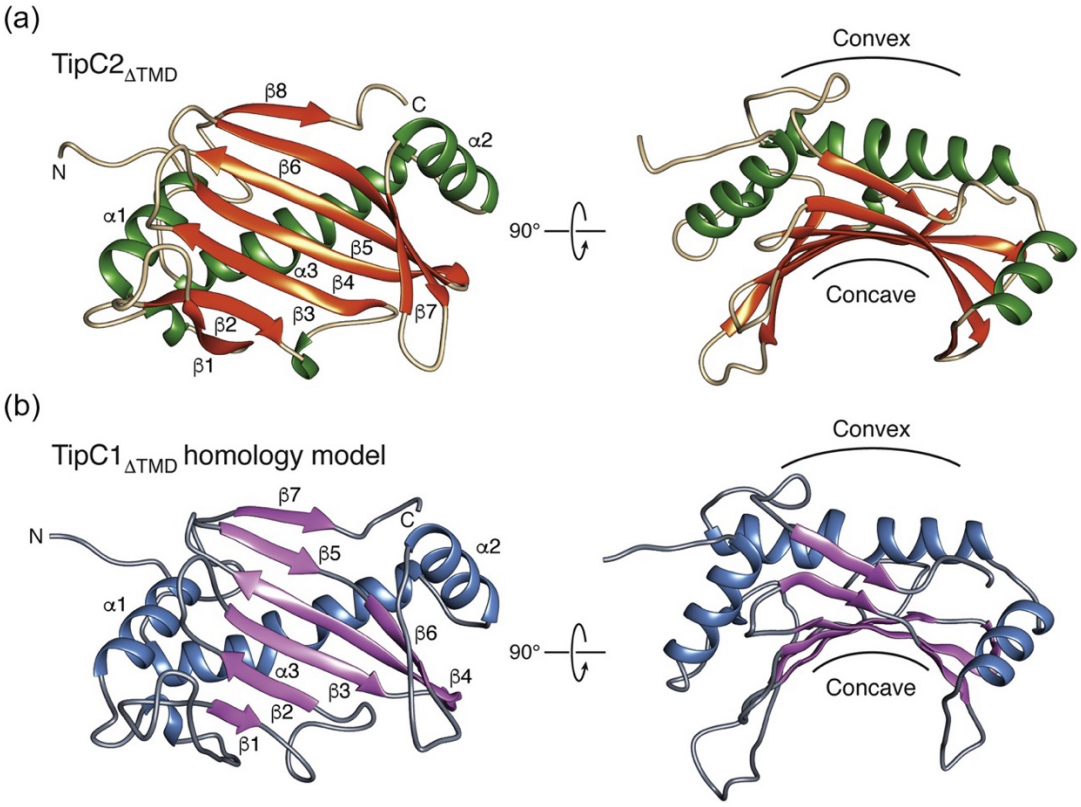


Figure 2.4: X-ray crystal structure of TipC2$_{\Delta TMD}$ and homology model of TipC1$_{\Delta TMD}$. (A) Overall structure of TipC2$_{\Delta TMD}$ shown as a ribbon representation and viewed from two orthogonal angles. (B) I-Tasser generated homology model of TipC1$_{\Delta TMD}$ shown as a ribbon representation and viewed from two orthogonal angles. Secondary structure elements and the concave and convex surfaces of both proteins are indicated.

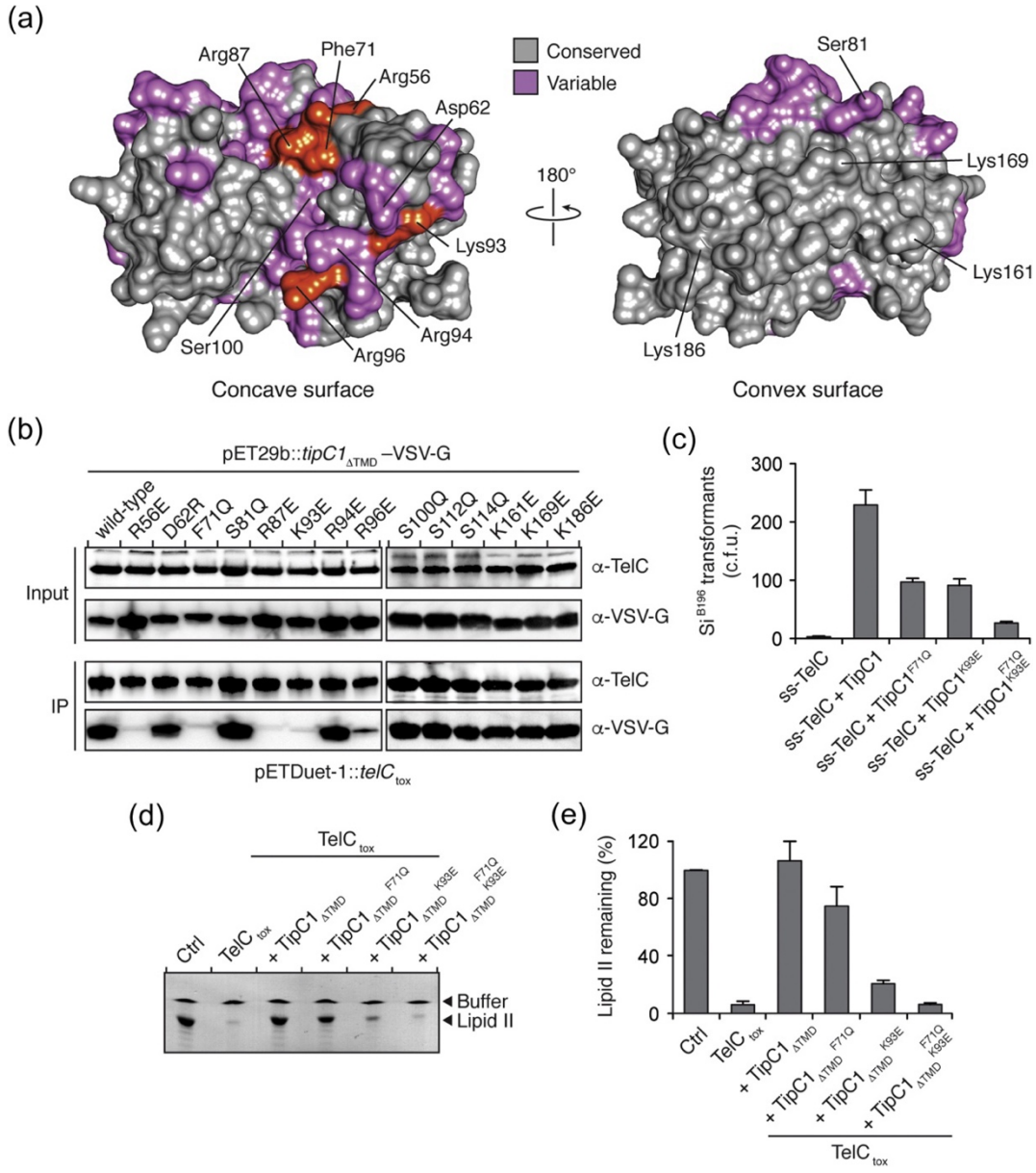


Figure 2.5: A concave surface of TipC1 mediates interaction with TelC. (A) Surface representation of a TipC1$_{\Delta TMD}$ homology model showing the concave and convex surfaces of the protein. Amino acid residues that are conserved (grey) or variable (pink) between TipC1$_{\Delta TMD}$ and TipC2$_{\Delta TMD}$ are depicted. Variable amino acids critical for interaction with TelC (red, defined in B) are labelled. (B) R56E, F71Q, R87E, K93E and

R96E variants of TipC1$_{\Delta TMD}$ do not interact with TelC. VSV-G epitope tagged wild-type TipC1$_{\Delta TMD}$ and the indicated TipC1$_{\Delta TMD}$ site-specific variants were co-expressed with his$_6$-tagged TelC and assessed for copurification by western blot analysis. (C) Number of *S. intermedius* B196 colonies after transformation with equimolar amounts of a plasmid constitutively expressing the indicated proteins. Plasmids expressing ss-TelC and ss-TelC + TipC1 serve as positive and negative controls, respectively. Error bars represent ± SD (*n* = 3). (D) Thin-layer chromatography analysis of reaction products from incubation of synthetic Lys-type lipid II with buffer (Ctrl), TelC$_{tox}$, TelC$_{tox}$ and TipC1$_{\Delta TMD}$ or TelC$_{tox}$ and the indicated TipC1$_{\Delta TMD}$ site-specific variants. (E) Densitometric quantification of (D). Error bars indicate ± SD (n = 3).

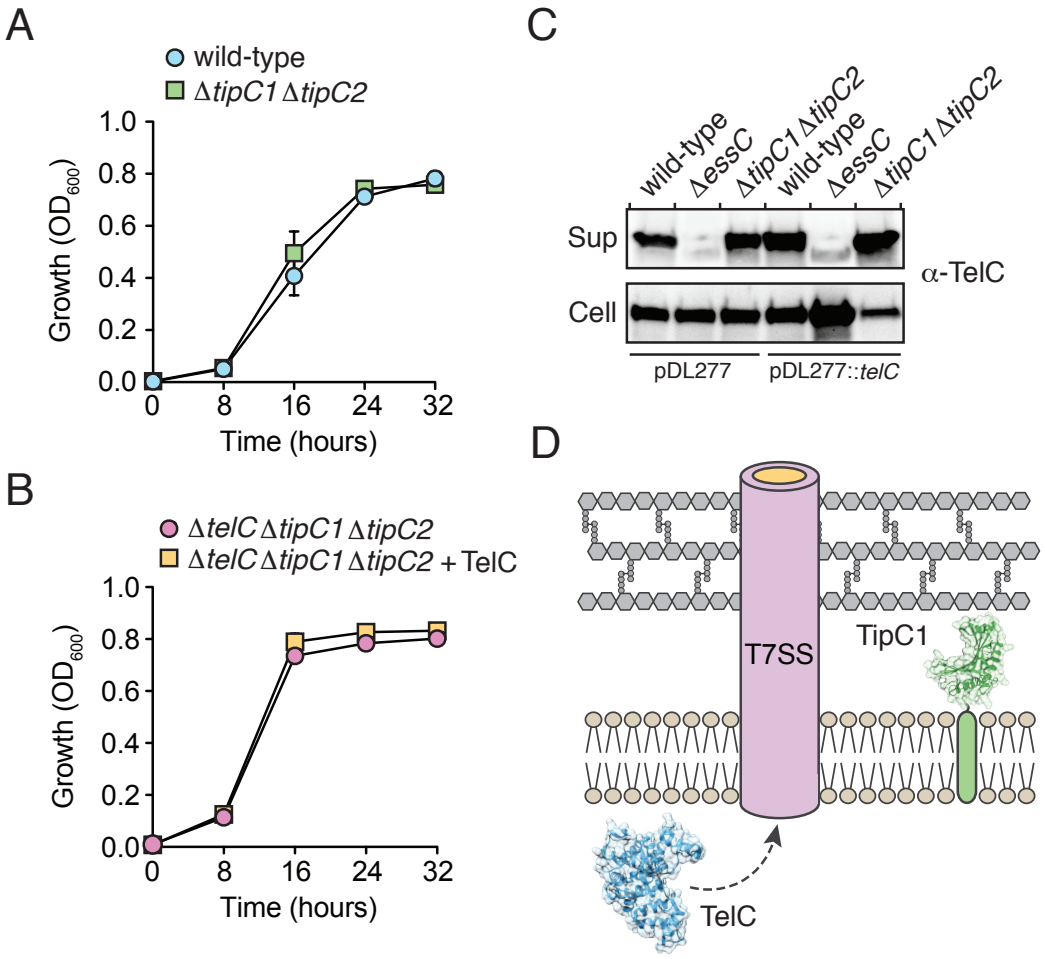


Figure 2.6: TelC does not access the inner wall zone as it transits the T7SS. (A) Mutational inactivation of *tipC* genes does not affect the growth of *S. intermedius* B196. Growth of the indicated *S. intermedius* B196 strains grown in liquid media. Error bars indicate ± SD (n = 3). (C) TelC expressed from its native locus or from a multi-copy plasmid accumulates in culture supernatants. Western blot analysis of TelC levels in supernatant (sup) or cell fractions of the indicated *S. intermedius* B196 strains. (C) Plasmid-borne expression of TelC in strains lacking *tipC* genes does not affect the growth of *S. intermedius* B196 strains grown in liquid media. Error bars indicate ± SD (n = 3).

(D) Model depicting the T7SS-dependent export of TelC across the Gram-positive cell envelope in a single step.

Tables

Table 2.1: X-ray data collection and refinement statistics for TipC2$_{\Delta TMD}$

	TipC2$_{\Delta TMD}$ (selenomethionine)
Data Collection	
Beamline	ALS 5.0.2
Wavelength (Å)	0.979
Space group	*C*2
Cell dimensions	
a, b, c (Å)	159.7, 54.5, 104.4
α, β, γ (°)	90.0, 108.0, 90.0
Resolution (Å)	33.60 - 1.75 (1.78 - 1.75)[a]
Total no. of reflections	85866
R_{merge} (%)[b]	4.8 (140.8)[a]
$I/\sigma I$	21.1 (1.3)[a]
Completeness (%)	99.2 (98.5)[a]
Redundancy	7.3 (6.7)[a]
Refinement	
R_{work} / R_{free} (%)[c]	17.0/19.4
No. atoms	
Protein	4489
Water	489
Average B-factors (Å^2)	
Protein	37.5
Water	32.3
Rms deviations	
Bond lengths (Å)	0.014
Bond angles (°)	1.221
Ramachandran plot (%)[d]	
Total favored	96.2
Total allowed	100.0
Coordinate error (Å)[e]	0.18

[a]Values in parentheses correspond to the highest resolution shell.
[b]$R_{merge} = \Sigma \Sigma |I(k) - \langle I \rangle| / \Sigma I(k)$ where $I(k)$ and $\langle I \rangle$ represent the diffraction intensity values of the individual measurements and the corresponding mean values. The summation is over all unique measurements.
[c]$R_{work} = \Sigma ||F_{obs}| - k|F_{calc}|| / |F_{obs}|$ where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. R_{free} is the sum extended over a subset of reflections excluded from all stages of the refinement.
[d]As calculated using MOLPROBITY (Chen et al., 2010).
[e]Maximum-Likelihood Based Coordinate Error, as determined by PHENIX (Adams et al., 2010).

Methods

Bacterial strains, plasmids and growth conditions

All *S. intermedius* strains used were generated from the sequenced B196 strain (Olson et al., 2013). *E. coli* strains XL-1, BL21 Codon Plus and BTH101 were used for plasmid maintenance, protein expression and Bacterial two-hybrid assays, respectively. A detailed list of bacterial strains and plasmids used in this study can be found in Tables S2.1 and S2.2. *S. intermedius* strains were grown statically in Todd Hewitt broth or on Todd Hewitt agar supplemented with 0.5% yeast extract at 37°C in the presence of 5% CO_2. *E. coli* strains used in this study were grown in LB broth at 37°C in a shaking incubator or on LB agar grown at 37°C in a static incubator. *S. gallolyticus* ATCC 43143 was grown in Brain Heart Infusion broth at 37°C in a shaking incubator. *S. intermedius* mutants were generated by replacing the gene to be deleted with a cassette conferring resistance to spectinomycin or kanamycin as previously described (Whitney et al., 2017). Briefly, the antibiotic resistance cassette was cloned between ~800 bp of sequence homologous to the regions flanking the gene to be deleted. The DNA fragment containing the cassette and flanking sequences was then linearized by restriction digest, gel purified, and ~250 ng of the purified fragment was added to 2 mL of log-phase culture pre-treated for two hours with competence peptide (500 ng/ml) to stimulate natural transformation. Cultures were further grown for four hours before plating on the appropriate antibiotic. All deletions were confirmed by PCR.

DNA manipulation and plasmid construction

S. intermedius and *S. gallolyticus* genomic DNA was prepared using a cell lysis buffer containing 20mg/mL lysozyme (BioShop), 25mM Tris-HCl pH 8.0, 2.5mM EDTA, and the DNA was purified using the Genomic DNA Mini Kit (Invitrogen). Primers were synthesized and purified by Integrated DNA Technologies (IDT). Q5 polymerase, restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (NEB). Site-specific mutants used in this study were generated by overlap extension PCR. All plasmids were sequenced by Genewiz Incorporated.

Subcellular fractionation

One litre of each *S. intermedius* strain was grown to an OD_{600} of 0.8 prior to centrifugation at 5,524 x *g* for 15 min. Pelleted cells were then resuspended in lysis buffer containing 25mM Tris-HCl pH 8.0, 150mM NaCl, 2mg/mL lysozyme and sonicated (4 x 30s pulses at 30% amplitude). Insoluble cellular debris was then cleared by centrifugation at 39,191 *g* for 30 min and the resulting supernatant was spun at 200,000 x *g* for two hours to isolate the membrane fraction. Aliquots of the supernatant fractions were added to Laemmli loading buffer whereas the membrane-containing pellet was washed once using 25mM Tris-HCl pH 8.0, 150mM NaCl buffer prior to dissolving in Laemmli loading buffer. Cytoplasmic and membrane fractions were then subjected to SDS-PAGE and western blot analysis.

Protease protection assay

Protease protection assays were performed as recently described for *Streptococcus pneumoniae* with minor modifications (Fenton et al., 2018). Briefly, 40mL of the indicated *S. intermedius* strains were grown to OD_{600} = 0.3 prior to harvesting by centrifugation at 4,000 x *g* for 15 minutes. Cells were washed once in SMM buffer (20mM maleic acid pH 6.5, 20mM $MgCl_2$, 0.5M sucrose) prior to resuspension in 2mL SMM buffer containing 5mg/mL lysozyme. Lysozyme digestion was carried out for 20 minutes at 37°C followed by washing and resuspension in 1mL SMM buffer. Aliquots of the resulting spheroplasts were either left untreated, treated with Proteinase K (20μg/mL), treated with Triton X-100 (1% *v*/*v*) or treated with Proteinase K and Triton X-100 for 30 min at room temperature. Proteolysis reactions were quenched using 1mM PMSF prior to the addition of Laemmli loading buffer. Samples were analyzed by SDS-PAGE and western blotting.

Western blotting

Western blot analyses were performed as previously described using rabbit α-VSV-G (Sigma, 1:5000) and rabbit α-TelC (1:3000) (Whitney et al., 2017). HRP-conjugated goat α-rabbit secondary antibody (Sigma, 1:5000) and ECL substrate (Clarity Max, Bio-Rad) were used for chemiluminescent detection. Western blots were imaged using a ChemiDoc System (Bio-Rad).

Identification of TipC homologous proteins

To determine TipC1 distribution in bacteria, the amino acid sequence of TipC1 was run through the iterative hidden Markov model search tool JackHMMER against the UniProtKB database. After five iterations, the search converged resulting in the identification of 286 protein sequences. An arbitrary subset of the genes encoding these TipC1 homologous proteins were selected for depiction in Figure 2.2.

Toxicity assays

S. intermedius cells were grown to mid-log phase (OD_{600} of 0.6) before competence was induced by the addition of 500ng of competence-stimulating peptide (CSP) per mL of culture. Cultures were then incubated for two hours prior to the addition of 1μg of the indicated plasmids to the media. After an additional three-hour incubation, 100μL of each culture was plated on selective media.

Co-purification assays

50ml of *E. coli* BL21 cells expressing the indicated plasmids were grown in LB broth to an OD_{600} of 0.6. Protein expression was then induced by adding IPTG to a final concentration of 1mM following by further incubated for three hours. Cells were collected by centrifugation at 5,524 x *g* for ten minutes and subsequently resuspended in lysis buffer (50mM Tris-HCl pH 8.0, 300mM NaCl, 10mM imidazole). Cells were then lysed by sonication and cellular debris was removed by centrifugation at 39,191 *g* for 30 min. Aliquots of the cleared lysate were added to Laemmli loading buffer for downstream

western blot analysis of the input fraction. 100uL of Ni-NTA slurry (Qiagen) was then added to the remaining cell lysate and incubated at room temperature for one hour. The beads were then washed three times with 10 mL of wash buffer (20mM Tris-HCl pH 8.0, 300mM NaCl, 10mM imidazole) by iterative rounds of centrifugation at 700 x *g* for two minutes followed by removal of the supernatant. Proteins bound to the Ni-NTA resin were then eluted by adding 500uL of elution buffer (20mM Tris-HCl, 150mM NaCl, 400mM imidazole) followed by a final spin at 700 x *g* to remove the resin. The eluate was then added to Laemmli sample buffer and was analyzed, along with the input fractions, by Western blot.

Bacterial two-hybrid analyses

E. coli BTH101 cells were co-transformed with plasmids encoding the T25 and T18 fragments of *Bordetella pertussis* adenylate cyclase fused to *Sg*TelC and *Sg*TipC1-4, respectively. Stationary phase cells were then plated on LB agar containing 40 μg/mL X-gal, 0.5 mM IPTG, 50 μg/mL kanamycin and 150 μg/mL carbenicillin and grown for 30 hr at 30°C. Plates were imaged using an iPhone 7 (Apple Inc.). A representative image of each two-hybrid experiment is shown. Three independent replicate experiments were performed for each pairwise combination and yielded comparable results.

Protein expression and purification

Two litres of *E. coli* BL21 CodonPlus cells expressing pETDuet-1::$tipC2_{\Delta TMD}$ were grown at 37°C in 2xYT broth an OD_{600} of 0.6 prior to induction of protein

expression with 1mM IPTG. Following further incubation at 37°C for four hours, cells were harvested by centrifugation and flash frozen. Frozen cells were thawed using lysis buffer (50mM Tris-HCl pH 8.0, 300mM NaCl, 10mM imidazole) and lysed by sonication (6 x 30 second pulses at 30% amplitude). Insoluble cellular debris was then cleared by centrifugation and the TipC2-containing supernatant was applied to a 5mL HisTrap™ FF Ni-NTA cartridge connected to an AKTA FPLC purification system (GE Healthcare). Unbound proteins were removed by extensive washing of the column in lysis buffer and TipC2$_{\Delta TMD}$ was eluted using a linear imidazole gradient to a final concentration of 400mM. Ni-NTA purified fractions of TipC2$_{\Delta TMD}$ were pooled the protein was further purified using a 16/600 HiLoad S200 size exclusion column (GE Healthcare) run in 20mM Tris-HCl pH 8.0, 150mM NaCl. Selenomethionine incorporated TipC2$_{\Delta TMD}$ was expressed an purified in an identical manner except that cells were grown in SelenoMethionine Medium Complete (Molecular Dimensions) and all purification buffers contained 1mM tris(2-carboxyethyl)phosphine (TCEP).

Crystallization and structural analyses

Size exclusion purified TipC2$_{\Delta TMD}$ was concentrated to 25mg/mL by spin filtration prior to crystallization (10kDa MWCO, Millipore). TipC2$_{\Delta TMD}$ at a concentration of 25mg/mL was screened against commercially available sparse matrix crystallization kits (MCSG1-4, Anatrace). After several days of incubation at room temperature, crystals of TipC2$_{\Delta TMD}$ grew in 100mM Tris-HCl pH 8.5, 25% w/v PEG 3350. Optimization of native TipC2$_{\Delta TMD}$ was not pursued because selenomethionine

incorporated TipC2$_{\Delta TMD}$ also readily crystallized in this condition. Single crystals of selenomethionine incorporated TipC2$_{\Delta TMD}$ were obtained by the streak seeding method and following cryoprotection of single crystals in the crystallization buffer supplemented with 20% ethylene glycol, a 1.8Å dataset was collected at beamline 5.0.2 at the Advanced Light Source (360 images, 1.0º Δ φ oscillation, 1.0s exposure and 250mm crystal-to-detector distance). X-ray diffraction data were merged, integrated and scaled using the *xia2* system (Winter et al., 2013).

X-ray phases were obtained by the selenium SAD technique using the AutoSol wizard built into the Phenix GUI (Terwilliger et al., 2009). The resulting electron density map was of sufficient quality to allow for automated model building of the complete structure using Phenix AutoBuild (Terwilliger, 2008). Minor model adjustments were made manually in Coot between iterative rounds of refinement using Phenix.refine (Afonine et al., 2012; Emsley et al., 2010). The final model was refined to an R_{work} of 17.0% and an R_{free} of 19.4%.

Homology modelling

A homology model of the TipC1$_{\Delta TMD}$ was obtained using the structure prediction server I-Tasser using the TipC2$_{\Delta TMD}$ structure as a template. The I-Tasser generated model of TipC1$_{\Delta TMD}$ had sequence coverage of 99% and a normalized Z-score of 10.0 (Yang et al., 2015).

Circular dichroism spectroscopy

Circular dichroism spectra were acquired using an AVIV model 4010 circular dichroism spectrometer (AVIV Associates, Lakewood, NJ). Prior to data acquisition, protein samples were buffer exchanged into 2mM HEPES, 15mM NaCl. Samples were then transferred to a quartz cell with a 1mm path length and data were collected at 25ºC. For each protein sample, spectra were averaged from three scans.

Lipid II phosphatase assay

The digestion of Lys-type lipid II (gift from Eefjan Breukink, University of Utrecht) was assessed by thin-layer chromatography (TLC) as previously described (Pazos et al., 2018). Briefly, $TelC_{tox}$ alone or $TelC_{tox}$ with 1.2 molar equivalents of $TipC1_{\Delta TMD}$, $TipC1_{\Delta TMD}^{F71Q}$, $TipC1_{\Delta TMD}^{K93E}$ or $TipC1_{\Delta TMD}^{F71Q, K93E}$ was incubated in a total volume of 50 µl with 2 nmol lipid II in 150 mM KCl, 0.1% Triton X-100 and 2 mM $CaCl_2$ for 90 min at 37°C. Lipids were extracted with n-butanol/pyridine acetate (2:1) pH 4.2 and resolved on a HPTLC silica gel 60 plate (Millipore) developed with chloroform/methanol/ammonia/water (88:48:1:10). Compounds were stained with iodine and bands were quantified by the ImageJ software.

Growth curves

For *S. intermedius* growth curves, overnight cultures of the indicated strains were sub-inoculated into THYB to a starting OD_{600} of 0.01. Cultures were grown statically at

37°C in the presence of 5% CO2 with OD_{600} measurements being taken at the indicated time points.

Secretion assay

S. intermedius strains were grown to an OD_{600} of 0.7 prior to harvesting by centrifugation at 10,000 x *g* for 10 minutes. Cell and supernatant fractions were prepared as described previously and analyzed by western blot analysis (Whitney et al., 2017).

Data availability

The data supporting Chapter II can be found entirely within this thesis as well as at the link found below. Structure files and information pertaining to the structure of TipC2 are indexed in the protein data bank (PDB: 6DHX). For access to strains and plasmids used in this chapter please contact Dr. John Whitney.

Relevant links:

https://www.sciencedirect.com/science/article/pii/S0022283618305941

https://www.rcsb.org/structure/6DHX

Chapter III – Structure of the extracellular region of the bacterial type VIIb secretion system subunit EsaA

Preface

The work presented in this chapter was previously published in the following study:

Klein, T.A., Grebenc, D.W., Gandhi, S.Y., Shah, V.S., Kim, Y., and Whitney, J.C. (2021). Structure of the extracellular region of the bacterial type VIIb secretion system subunit EsaA. Structure 29 (2): 177-185.

Author contributions: T.A.K. and J.C.W. planned the study. All authors contributed to experimental design. T.A.K. and J.C.W. generated strains and plasmids. T.A.K. performed protein expression, purification and crystallization. S.Y.G. and V.S.S. assisted with protein crystallization. T.A.K., D.W.G., Y.K. and J.C.W. solved and analyzed the crystal structure. T.A.K. and D.W.G. performed biochemical experiments. T.A.K., D.W.G., and J.C.W. analyzed the data. T.A.K., D.W.G. and J.C.W. wrote the paper. All authors provided feedback on the manuscript.

Permission has been granted by the publishers to reproduce the material from this study.

Abstract

Gram-positive bacteria use type VII secretion systems (T7SSs) to export effector proteins that manipulate the physiology of nearby prokaryotic and eukaryotic cells. Several mycobacterial T7SSs have established roles in virulence. By contrast, the genetically distinct T7SSb pathway found in Firmicutes bacteria more often functions to mediate bacterial competition. A lack of structural information on the T7SSb has limited the understanding of effector export by this protein secretion apparatus. Here, we present the 2.4Å crystal structure of the extracellular region of the T7SSb subunit EsaA from *Streptococcus gallolyticus*. Our structure reveals that homodimeric EsaA is an elongated, arrow-shaped protein with a surface-accessible 'tip', which in some species of bacteria serves as a receptor for lytic bacteriophages. Because it is the only T7SSb subunit large enough to traverse the peptidoglycan layer of Firmicutes, we propose that EsaA plays a critical role in transporting effectors across the entirety of the Gram-positive cell envelope.

Introduction

Protein secretion is a critical aspect of bacterial physiology and requires the use of membrane-embedded secretion apparatuses. In addition to the general secretory pathway and the twin-arginine translocase, many species of Gram-positive bacteria use type VII secretion systems (T7SSs) for protein export (Abdallah et al., 2007). T7SSs are used by bacteria belonging to the phyla Actinobacteria and Firmicutes and are divided into T7SSa and T7SSb. This distinction reflects differences in T7SS subunit composition between these two distantly related groups of Gram-positive bacteria (Klein et al., 2020). The T7SSa was originally discovered in *Mycobacterium tuberculosis* where it acts as a virulence factor that facilitates immune evasion and phagosomal escape during infection, whereas the T7SSb was initially characterized in *Staphylococcus aureus* and has been shown to play a dual role in pathogenesis and interbacterial competition (Cao et al., 2016; Gao et al., 2004; Ohr et al., 2017; Ulhuq et al., 2020). The interkingdom-targeting capability of the T7SSb has also been demonstrated in the opportunistic pathogen *Streptococcus intermedius* with the antibacterial activity being attributed to the NAD^+ hydrolase effector TelB and the cell wall precursor degrading effector TelC (Hasegawa et al., 2017; Klein et al., 2018; Whitney et al., 2017). The T7SSb pathways of *Bacillus subtilis* and *Enterococcus faecalis* were also recently shown to antagonize competitor bacteria (Tassinari et al., 2020; Chatterjee et al., 2020).

Much of our current understanding of the T7SS has resulted from studies on effector function, which can often explain the phenotypes associated with a given T7SS pathway. Less well understood is the mechanism of T7SS effector export across the cell

envelope. Recent structural analyses have begun to elucidate the ultrastructure of T7SS apparatuses and provide clues as to how this secretion apparatus facilitates protein export (Famelis et al., 2019; Poweleit et al., 2019; Rosenberg et al., 2015). However, these studies have largely focused on T7SSa apparatuses. Of the four major structural proteins that make up the T7SSa, only the EccC/EssC/YukB ATPase is conserved in T7SSb systems. The other three T7SSa subunits, EccB, EccD, and EccE, possess no sequence homology to the EssA, EssB, and EsaA components of the T7SSb and consequently, the two systems likely form distinct structures that may not share a common mechanism for protein export.

EsaA is perhaps the least understood of the T7SSb structural components. Transposon mutagenesis in *S. aureus* strain Newman initially suggested that *esaA* was dispensable for effector secretion (Burts et al., 2005). However, subsequent characterization of an *S. aureus* RN6390 *esaA* mutant strain generated by allelic replacement showed that this subunit is likely essential for T7SSb-dependent protein export (Kneuper et al., 2014). No structural data exists for EsaA, but analysis of its membrane topology suggests it consists of a large soluble region flanked by N- and C-terminal transmembrane domains (TMDs) (Ahmed et al., 2018; Mietrach et al., 2019). Proteomic analyses of intact *S. aureus* cells has shown that EsaA is surface exposed and that its soluble domain may extend into the extracellular milieu (Dreisbach et al., 2010). Furthermore, studies in *B. subtilis* have shown that the EsaA homologue YueB is a cell surface receptor for the SPP1 bacteriophage (Sao-Jose et al., 2004; Sao-Jose et al., 2006). Similarly, many strains of *E. faecalis* possess the EsaA paralogue Phage Infection Protein

(PIP), which serves as a receptor for Enterococcal phage (Duerkop et al., 2016). The prediction that EsaA extends from the plasma membrane to the cell surface makes it unique among the T7SSb subunits because the other structural proteins have either extracellular domains that are too small to span the estimated 30-50 nm thick peptidoglycan layer of Firmicutes bacteria or are entirely intracellular (Tassinari et al., 2020; Vollmer et al., 2008).

In this study, we present the crystal structure of the extracellular domain of EsaA, revealing a highly elongated, arrow-shaped homodimer comprised of three distinct domains. Using cysteine cross-linking, we show that EsaA dimers occur *in vivo* and propose that upon multimerization with the other subunits of the T7SSb, form a conduit that facilitates effector export across the cell envelope of Gram-positive bacteria.

Results

EsaA is required for the secretion of EsxA and Tel effector proteins from *S. intermedius*

Given the conflicting reports on the essentiality of EsaA for T7SSb function, we first examined the consequences of inactivating *esaA* on effector export using the model T7SSb bacterium *S. intermedius.* Characterized T7SSb systems export two major families of effectors: small, α-helical WXG100 proteins whose precise function is unknown; and large, multi-domain LXG proteins that possess C-terminal toxin domains. *S. intermedius* strain B196 exports a single WXG100 effector, EsxA, and the three LXG effectors TelA, TelB and TelC (Whitney et al., 2017). Consistent with functioning as a core structural subunit of the T7SSb apparatus, we found that replacement of the *esaA* gene with a

kanamycin resistance cassette yielded a *S. intermedius* strain that is unable to export detectable levels of EsxA and TelC into culture supernatants (Fig. 3.1A). Similarly, supernatant NADase activity, which is indicative of TelB secretion, was reduced to levels comparable to that of a T7SSb-inactivated strain, Δ*essC*. (Fig. 3.1B). Importantly, we found that export of EsxA and TelC, as well as TelB-dependent NADase activity could be restored by plasmid-based expression of EsaA indicating that our allelic replacement approach did not affect the expression of genes encoding other structural subunits of the T7SSb, which are part of a five-gene cluster that also contains *esaA*. Together, these data indicate that *esaA* is required for WXG100 and LXG effector export in *S. intermedius*.

Topology mapping of EsaA reveals a large extracellular domain

We next sought to examine the membrane topology of *S. intermedius* EsaA (*Si*EsaA). Though cell surface proteomics conducted on *S. aureus* suggest that the soluble region of EsaA exists extracellularly, this assertion has not been tested directly for any T7SSb-containing bacterium. Furthermore, the number of putative TMDs differs among EsaA homologues with *Si*EsaA having a single predicted TMD on either side of its soluble region whereas EsaA proteins from *S. aureus, E. faecalis, B. subtilis, Bacillus cereus* and *Listeria monocytogenes* possess five putative TMDs at their C-terminus (Fig S3.1).

After confirming that *Si*EsaA localizes to the membrane fraction of lysed *S. intermedius* cells (Fig. 3.2A), we introduced a series of cysteine point mutations spaced approximately 150 amino acids apart within *Si*EsaA to map its membrane topology using

a cysteine-reactive maleimide-conjugated fluorophore (Fig. 3.2B). Plasmid-borne expression of each EsaA cysteine mutant in our *esaA* deletion strain restored T7SSb-dependent export of TelC, demonstrating that these mutations do not significantly affect EsaA function (Fig. 3.2C). *Si*EsaA contains a single native cysteine residue predicted to reside in its N-terminal TMD, and we found that with intact cells this residue was inaccessible to the cysteine-reactive dye when analyzed by SDS-PAGE (Fig. 3.2D). Similarly, *Si*EsaA variants harboring cysteine mutations near the N- (V8C) or C-terminus (F909C) of the protein did not react with the dye. By contrast, we found that cells expressing *Si*EsaA bearing V150C, F302C, S454C or S605C mutations, all of which reside within the predicted soluble region, yielded a prominent fluorescent band at the expected molecular weight of *Si*EsaA (Fig. 3.2D). A fluorescent band absent in the wild-type control was also present in the V762C variant; however, this band migrates at a higher molecular weight than *Si*EsaA making it difficult to interpret. Collectively, our data indicate that *Si*EsaA is a membrane protein with a large extracellular domain and intracellular N- and C- termini.

Structure determination of an extracellular fragment of EsaA

Having mapped the membrane topology of *Si*EsaA, we next initiated structural studies on the large extracellular fragment of the protein to gain more insight into its function. Although we could readily express and purify a truncation of *Si*EsaA encompassing its entire extracellular region (residues 41-871), this protein fragment had a propensity to degrade. To identify a stable fragment of *Si*EsaA that would be more

amenable to crystallization, we performed limited proteolysis with chymotrypsin and isolated a protease-resistant species spanning residues 234-790 (Fig. S3.2). This fragment of *Si*EsaA crystallized readily but despite extensive optimization efforts, diffraction quality crystals could not be obtained. Using the boundary information obtained from our proteolysis experiments, we next tried a homologous EsaA fragment from *Streptococcus gallolyticus* ATCC 43143 (*Sg*EsaA$_{235\text{-}829}$), which has 42.9% pairwise sequence identity to the equivalent region of *Si*EsaA (Fig. 3.3A and Fig. S3.3A). Purified *Sg*EsaA$_{235\text{-}829}$ formed diffraction quality crystals and the 2.4Å structure of SgEsaA$_{235\text{-}829}$ was determined using selenium-incorporated protein and the single-wavelength anomalous dispersion technique (Table 3.1). Interestingly, the resulting electron density map only yielded interpretable density for a model encompassing residues 329-727 (henceforth referred to as *Sg*EsaA$_{329\text{-}727}$) with an unmodeled gap from amino acids 513-554, suggesting that large portions of EsaA are disordered in the crystal lattice (Fig. S3.3B). The final model was refined to a R_{work}/R_{free} of 0.21 and 0.26, respectively.

EsaA forms an elongated, arrow-shaped dimer

*Sg*EsaA$_{329\text{-}727}$ forms a highly elongated structure comprised of two alpha helical domains (AD-I and AD-II) and a beta-sheet domain (BD) (Fig. 3.3B and Fig. 3.3C). The modelled fragment adopts a 'there and back again' topology whereby the first half of *Sg*EsaA$_{329\text{-}727}$ contributes secondary structure elements to each of the three domains over a linear distance of 196Å. Following a 180 degree turn that occurs within the unmodelled region between the b1 and b2 strands of the BD, the C-terminal half of the protein

similarly contributes secondary structure to each domain with the C-terminus being located ~20Å away from the N-terminus at the same pole (Fig. 3.3B). In this arrangement, both the N- and C-terminal TMDs present in full-length EsaA would be connected to the AD-I domain. Given the orientation of the termini, the directionality of the β-strands flanking the central unmodelled region, and the number of unmodelled amino acids in our structure, it is likely that the length of the entire extracellular region of EsaA is well in excess of the ~200Å measured for our model. This finding provides a molecular explanation for how this protein is potentially able to traverse the approximately 30-50nm thick cell wall of Firmicutes bacteria (Vollmer et al., 2008).

Another striking feature of *Sg*EsaA$_{329\text{-}727}$ is that it adopts a head-to-head, belly-to-belly homodimer that gives the protein its arrow-shaped appearance (Fig. 3.3D and Fig. S3.4A). The *Sg*EsaA homodimer was generated by a symmetry operation because the dimer axis is coincident with a crystallographic axis. In this arrangement, all three domains and the intervening connecting regions contribute to the dimerization interface (Fig. 3.3E). Analysis of the dimer interface using the PDBePISA webserver indicates that dimer formation is highly favorable ($\Delta^{i}G$: -61.8kcal/mol) and generates 4436Å^2 of buried surface area (Krissinel & Henrick, 2007). Mapping EsaA sequence conservation onto our structure reveals that the residues comprising the surface of EsaA are highly variable whereas the amino acids involved in homodimerization show a much higher level of conservation (Fig. 3.3F). The amino acids lining the dimer interface are a mixture of hydrophobic, polar and acidic residues with tyrosine, leucine, threonine and glutamate being the most abundant. We speculate that the large surface area of the dimer interface

combined with the abundance of hydrophobic residues participating in homodimerization indicates that EsaA likely exists as an obligate homodimer because solvent exposure of this surface in aqueous environments would bear a large entropic cost.

A comparison of *Sg*EsaA$_{329-727}$ to previously determined structures in the Protein Data Bank using DALILITE revealed that the overall structure of *Sg*EsaA$_{329-727}$ does not resemble proteins of known structure (Holm, 2020). The top hit from this search was the BID domain of the type IV secretion system (T4SS) effector protein Bep9 from *Bartonella clarridgeiae* (Z-score, 8.5; Ca root mean square deviation of 3.5Å over 100 aligned residues), which only shares structural similarity with AD-I of EsaA (Fig. S3.4B)(Stanger et al., 2017). BID domains comprise one part of a bipartite signal sequence found in some T4SS effectors and thus appear unrelated in terms of function. Based on these analyses, we conclude that EsaA adopts a unique protein fold.

EsaA exists as a dimer *in vitro* and *in vivo*

To test the biological significance of the EsaA homodimer observed in our crystal structure, we examined a truncation of *Sg*EsaA that more accurately reflects the modeled boundaries of our structure (*Sg*EsaA$_{332-725}$) as well as the equivalent fragment of *Si*EsaA (*Si*EsaA$_{328-685}$) by size exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALS). SEC-MALS allows for the accurate determination of protein molecular mass in solution and therefore helps identify potentially artefactual oligomeric states induced by protein crystallization. For both proteins, the major peak yielded a molecular mass consistent with dimer formation and no evidence of EsaA monomers was

observed in either case (Fig. 3.4A and Fig. 3.4B). The SEC-MALS analysis of *Sg*EsaA$_{332-725}$ also revealed the presence of high molecular weight aggregates but due to their heterogeneous nature and absence in the *Si*EsaA$_{328-685}$ sample, we concluded that they likely do not represent biologically relevant assemblies of EsaA. In sum, the extracellular fragment of EsaA exists as a dimer in solution.

We next wanted to examine if EsaA dimerizes *in vivo* in a manner that is consistent with our crystal structure. To accomplish this, we inspected our *Sg*EsaA$_{329-727}$ structure for amino acid residues within the dimer interface that would be expected to crosslink if mutated to cysteine. This analysis led to the identification of Thr628, found in in the linker region between the BD and AD-II, Ala654, located within the AD-II, and Leu688, which exists in the linker region between AD-I and AD-II (Fig. 3.4C). We mutated each of these residues, along with the equivalent residues in *Si*EsaA (Asn586, Thr612 and Leu644), to cysteine and examined the ability of these variants to form covalent dimers. In support of the dimeric arrangement observed in our crystal structure, all six variants formed b-mercaptoethanol (BME)-sensitive crosslinks when the purified proteins were examined by SDS-PAGE (Fig. 3.4D and Fig. 3.4E). Furthermore, when we introduced the *Si*EsaA cysteine variants into our *S. intermedius* B196 *esaA* deletion strain, BME-sensitive cysteine cross-links were observed in cells expressing either EsaA^{N586C} or EsaA^{L644C} (Fig. 3.4F). Collectively, our cross-linking data suggest that the structure of *Si*EsaA is likely very similar to that of *Sg*EsaA in terms of overall fold and dimeric arrangement, and that dimeric EsaA represents a biologically relevant form of the protein.

The structure of EsaA predicts the putative binding site for a bacteriophage receptor

EsaA homologous proteins are not only involved in type VII secretion but have also been shown to function as receptors for lytic bacteriophages (Sao-Jose et al., 2004). A recent analysis of Enterococcal phages identified a 160 amino acid hypervariable region within the EsaA homologous protein PIP (Phage Infection Protein) responsible for phage tropism among *E. faecalis* strains (Duerkop et al., 2016). The topology of EsaA combined with the domain organization revealed by our *Sg*EsaA$_{329\text{-}727}$ crystal structure suggest that the β-sheet domain of this protein family is likely the surface exposed region, leading us to speculate that this region of the protein likely serves as the receptor for infecting phage. Indeed, mapping the hypervariable region of PIP proteins onto an EsaA-derived homology model of a representative PIP protein from *E. faecalis* V583 indicates that the phage tropism determining region identified by Duerkop et al. likely exists within the BD of EsaA homologous proteins (Fig. S3.4C).

Discussion

Our structure of the extracellular region of EsaA has revealed the unique architecture of this enigmatic T7SSb subunit. EsaA consists of three distinct domains that each contribute to homodimer formation and give the protein its overall arrow-shaped appearance. The observation that T7SSb subunits form dimers is not without precedent as a recently determined crystal structure of full-length YukC (EssB) from *B. subtilis* found that this T7SSb subunit similarly homodimerizes (Tassinari et al., 2020). EssB/YukC also physically interacts with EsaA/YueB. However, it is important to note that this

interaction occurs independently of the EsaA/YueB extracellular domain. Nevertheless, the physical interaction between these T7SSb subunits suggests that they likely function together to facilitate protein secretion across the cell envelope (Ahmed et al., 2018).

Though T7SS structural components form dimers in crystals, current evidence indicates that the ultrastructure of an assembled T7SS apparatus involves hexamerization of the apparatus components. For example, the ESX-5 T7SSa from *Mycobacterium xenopi* exhibits six-fold symmetry and is proposed to contain 1:1:1:1 stoichiometry of the four T7SSa apparatus components EccB, EccC, EccD and EccE based on a 13Å negative stain electron microscopy (EM) map (Beckham et al., 2017). More recently, higher resolution cryo-EM structures of the ESX-3 T7SSa from *Mycobacterium smegmatis* have suggested a 1:1:2:1 protomer stoichiometry in which two EccD subunits interact with one subunit each of EccB, EccC, and EccE (Famelis et al., 2019; Poweleit et al., 2019). Though they are not homologues, EccB and EsaA could play similar roles in effector export for T7SSa and T7SSb pathways, respectively, because both proteins are the only subunit from their respective system that possess a large extracellular domain. However, our structure shows that both the overall structure and dimerization mode of EsaA is substantially different from that of EccB indicating that these T7SS subunits may have distinct functions. Furthermore, the extracellular region of EsaA is cell surface exposed whereas EccB predominantly exists in the mycobacterial periplasm. This observation suggests that additional factors may be involved in T7SSa-dependent effector export across the mycomembrane such as the EspB protein or members of the proline-glutamate and proline-proline-glutamate families of proteins (Solomonson et al., 2015; Wang et al.,

2020). Ultimately, the structure of an intact T7SSb will be needed for an in-depth comparison between these intriguing protein export machines.

The *S. intermedius* T7SSb antibacterial effector TelC exerts toxicity in the inner wall zone (IWZ) by degrading the cell wall precursor lipid II present in the outer leaflet of the plasma membrane (Whitney et al., 2017). We previously used this unique site of action to provide evidence that the T7SSb exports effectors across the plasma membrane and the cell wall in a manner that bypasses the IWZ during transport (Klein et al., 2018). It is now apparent that EsaA, as the only T7SSb apparatus protein with an extended extracellular domain, may well form the conduit that allows for such transport. One of the defining characteristics of Gram-positive Firmicutes bacteria is the 30-50 nm thick peptidoglycan layer, which would likely prevent the diffusion of large ~70kDa LXG effectors from the IWZ to the extracellular milieu (Vollmer et al., 2008). Our structure of EsaA is 20 nm long and represents only a portion of the full-length protein. It is therefore within reason that EsaA extends across the entire cell wall to facilitate effector export from the cell (Fig. 3.5). These observations, coupled with the abovementioned propensity for T7SS subunits to adopt six-point symmetry, lead us to speculate that EsaA dimers might trimerize to form a hexameric tube-shaped assembly. Such a structure would not only enable effector export from T7SSb-containing bacteria but may also facilitate the delivery of effectors into target cells.

Figures

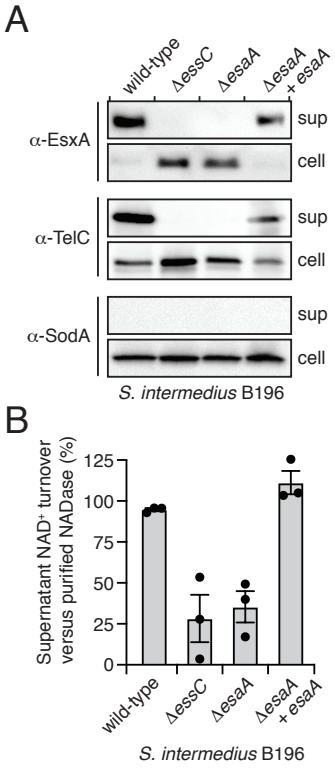


Figure 3.1: EsaA is required for WXG100 and LXG effector export by *S. intermedius*. (A) Western blot analysis of the cell and supernatant fractions of the indicated *S. intermedius* B196 strains. EsxA and TelC belong to the WXG100 and LXG families of T7SSb effectors, respectively. The Δ*essC* strain is used as a secretion deficient control. Superoxide dismutase A (SodA) is used as a cell lysis control. (B) Supernatant NADase activity, indicative of T7SSb-dependent TelB secretion, in cultures of the indicated *S. intermedius* B196 strains. Assay was performed in triplicate and values were calculated as a fraction of NAD^+ turnover compared to the purified NADase Tse6 (Whitney et al., 2015). The data displayed represent three independent replicates. Error bars reflect standard error of the mean (SEM).

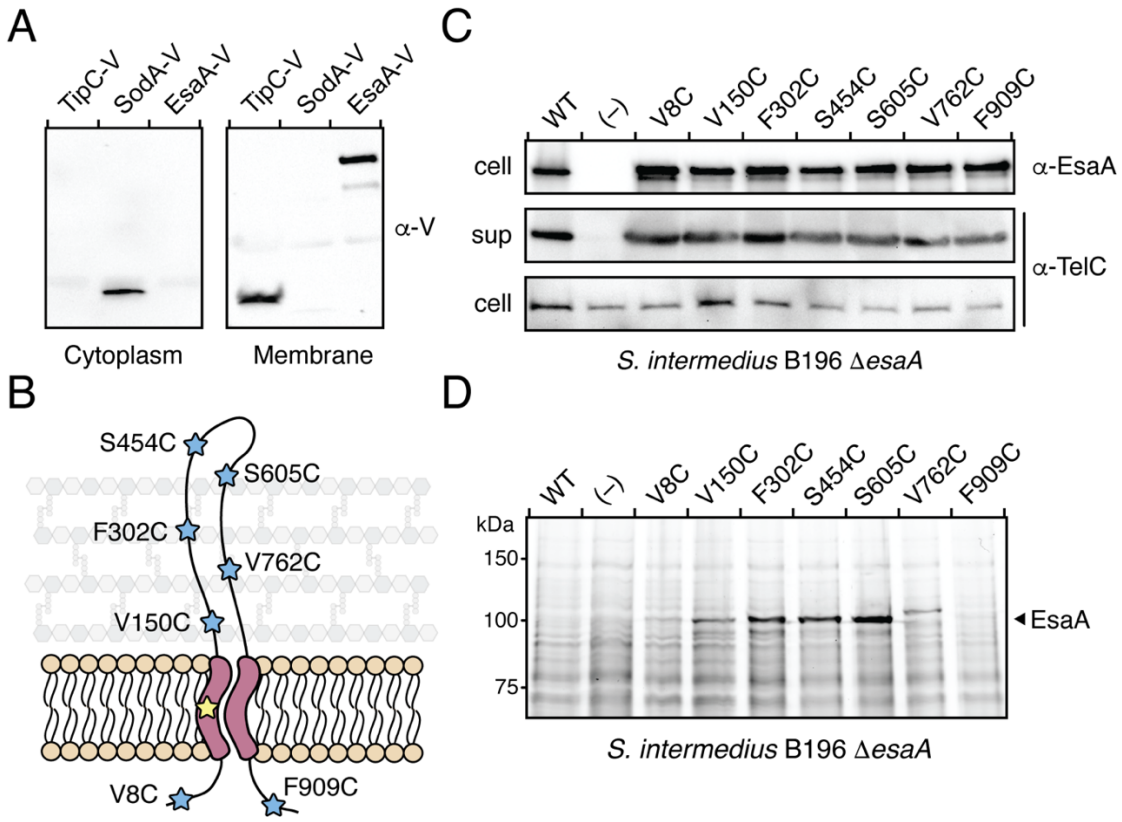


Figure 3.2: EsaA possesses a large extracellular domain. (A) EsaA fractionates with *S. intermedius* membranes. TipC and SodA serve as membrane and cytoplasmic controls, respectively. All proteins contain a C-terminal VSV-G tag and were detected by western blot using an α-VSV-G (α-V) primary antibody. (B) Predicted EsaA membrane topology depicting the location of each cysteine substitution site. The yellow star denotes the native cysteine residue present in EsaA whereas blue stars indicate cysteine mutations generated for topology mapping. (C) EsaA cysteine mutants are expressed and secrete TelC at levels similar to wild-type *S. intermedius*. (D) Cysteine mutations in the predicted extracellular domain of EsaA are accessible to a cysteine-reactive maleimide dye but those located near the N- and C-termini are not. EsaA migrates slightly above the 100kDa marker as indicated.

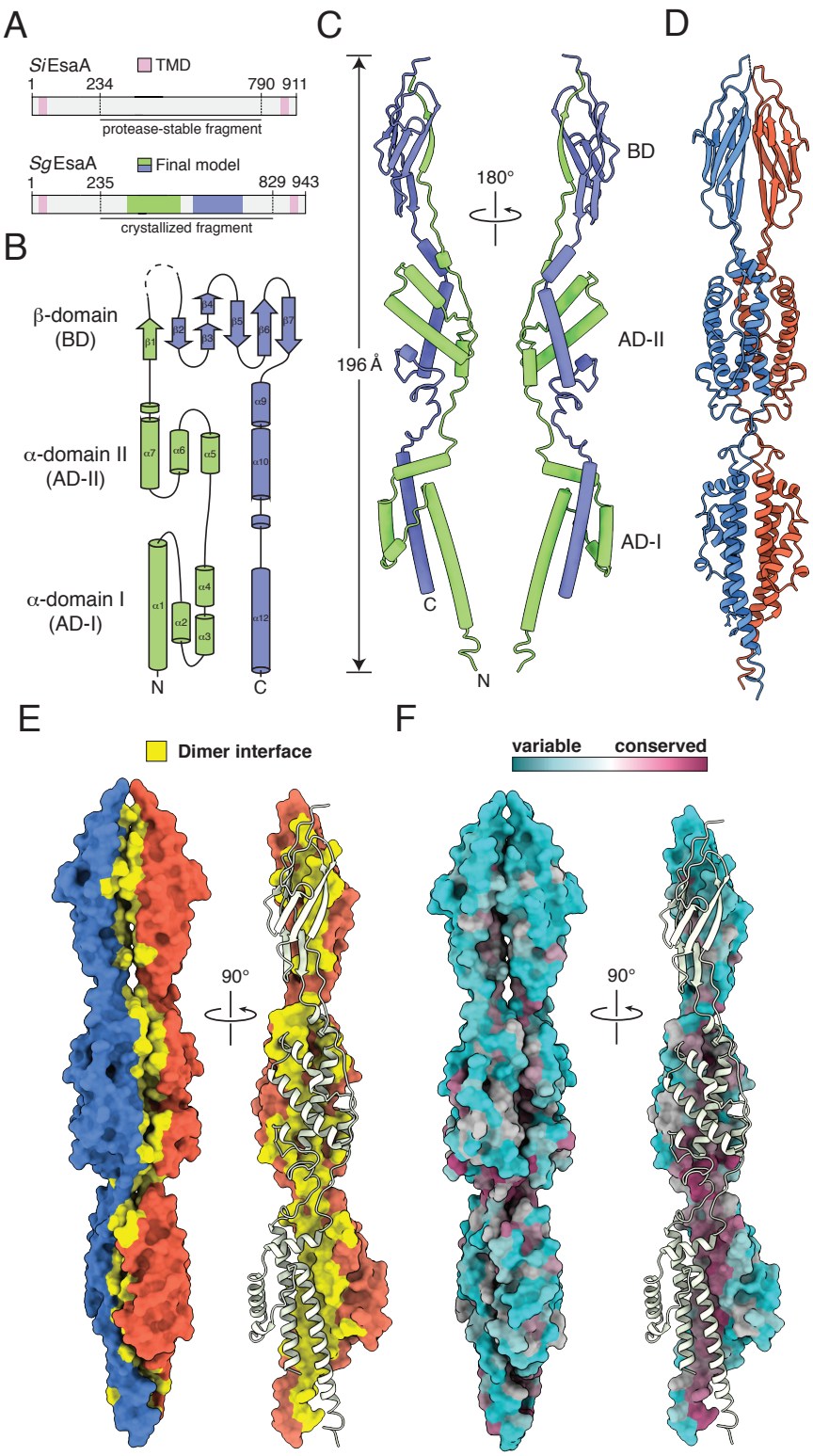

A
SiEsaA
TMD
1
234
790 911
protease-stable fragment
SgEsaA
Final model
1
235
829 943
crystallized fragment
B
β-domain
(BD)
α-domain II
(AD-II)
α-domain I
(AD-I)
N
C
C
196 Å
180°
BD
AD-II
AD-I
C
N
D
E
Dimer interface
90°
F
variable
conserved
90°

Figure 3.3: The extracellular domain of *Sg*EsaA adopts an arrow-shaped structure. (A) Domain architecture of *S. intermedius* B196 EsaA (*Si*EsaA) and *S. gallolyticus* ATCC 43143 EsaA (*Sg*EsaA) depicting the chymotrypsin-stable fragment of *Si*EsaA, the crystallized fragment of *Sg*EsaA, and the regions of *Sg*EsaA for which interpretable electron density was observed in the crystal structure. (B) Topology diagram depicting the secondary structure elements comprising *Sg*EsaA$_{329\text{-}727}$. Blue and green coloring is used to illustrate the 'there and back again' topology of the protein. (C) Model of *Sg*EsaA$_{329\text{-}727}$ shown from two opposing views. α-helices and β-strands are denoted by tubes and arrows, respectively. The N- and C-termini are depicted on the left-hand model. (D) *Sg*EsaA$_{329\text{-}727}$ dimers form an elongated structure. Red and blue ribbon coloring is used to differentiate each monomer within the dimer. (E) Surface representation of an *Sg*EsaA$_{329\text{-}727}$ dimer shown from orthogonal viewpoints. Yellow coloring is used to highlight the buried surface area between *Sg*EsaA$_{329\text{-}727}$ protomers. (F) Surface representation of an *Sg*EsaA$_{329\text{-}727}$ dimer depicting residue-specific sequence conservation among EsaA homologous proteins. Details of the sequences used for conservation analysis can be found in Experimental Procedures. Model was generated using the ConSurf server (Ashkenazy et al., 2016).

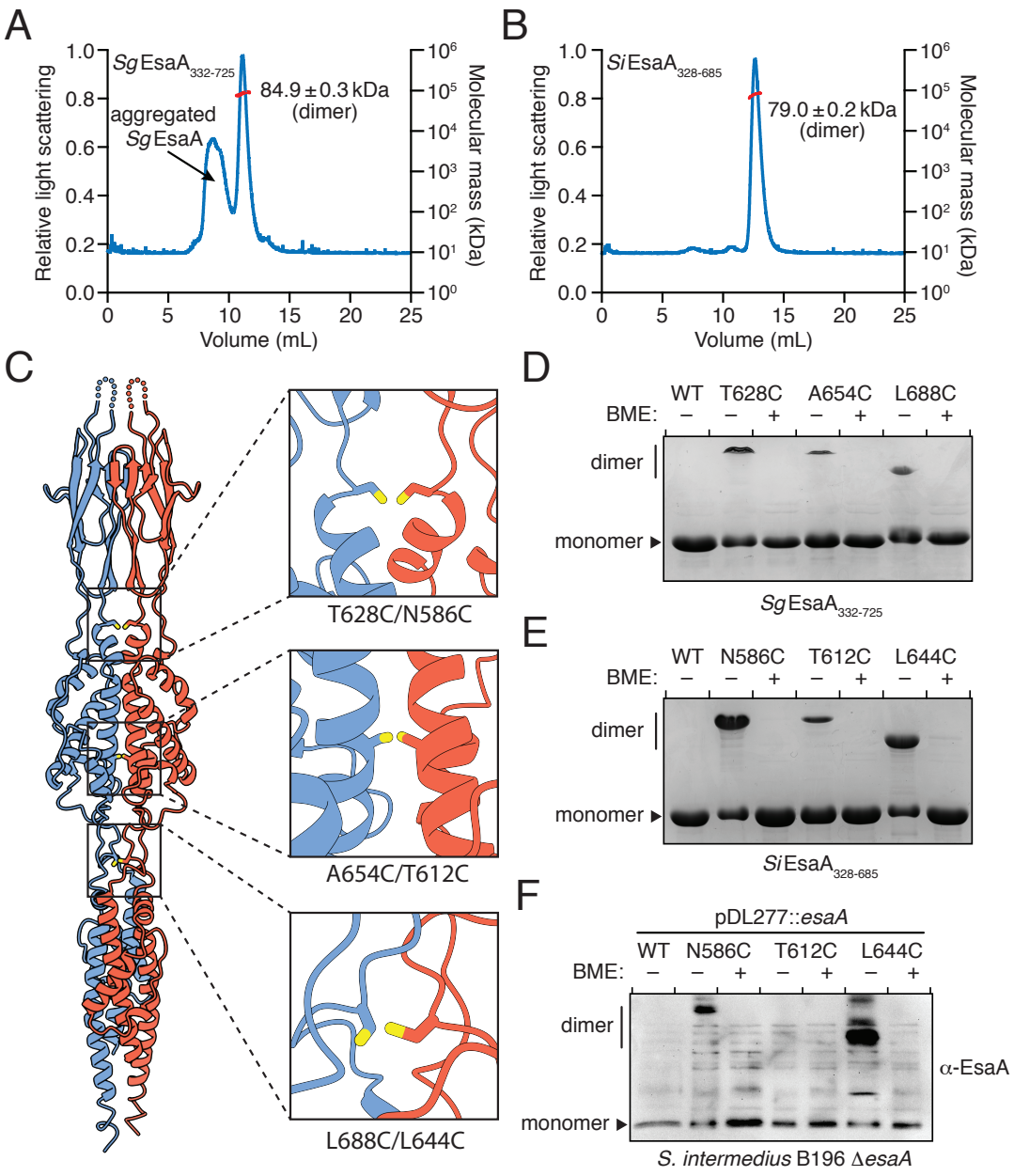


Figure 3.4: EsaA forms dimers *in vitro* and *in vivo*. (A-B) SEC-MALS analysis of *Sg*EsaA$_{332-725}$ (A) and *Si*EsaA$_{328-685}$ (B). Relative light scattering is plotted in blue and molecular weight is plotted in orange. The calculated molecular weights of the dimer peaks for both proteins are indicated. (C) Structure of *Sg*EsaA$_{329-727}$ depicting the cysteine

mutations chosen for cross-linking experiments. *Sg*EsaA$_{329-727}$ protomers are depicted as blue and red ribbons with the hypothetical cysteine mutations shown as sticks. The identities of the residues normally found in these positions are indicated for both *Sg*EsaA (left) and *Si*EsaA (right). (D-E) Coomassie blue-stained gel demonstrating cysteine crosslinking for each of the purified *Sg*EsaA$_{332-725}$ (D) and *Si*EsaA$_{328-685}$ (E) cysteine variants. (F) Western blot analysis of *S. intermedius* B196 Δ*esaA* strains expressing wild-type EsaA or each of the indicated EsaA cysteine variants. BME, b-mercaptoethanol.

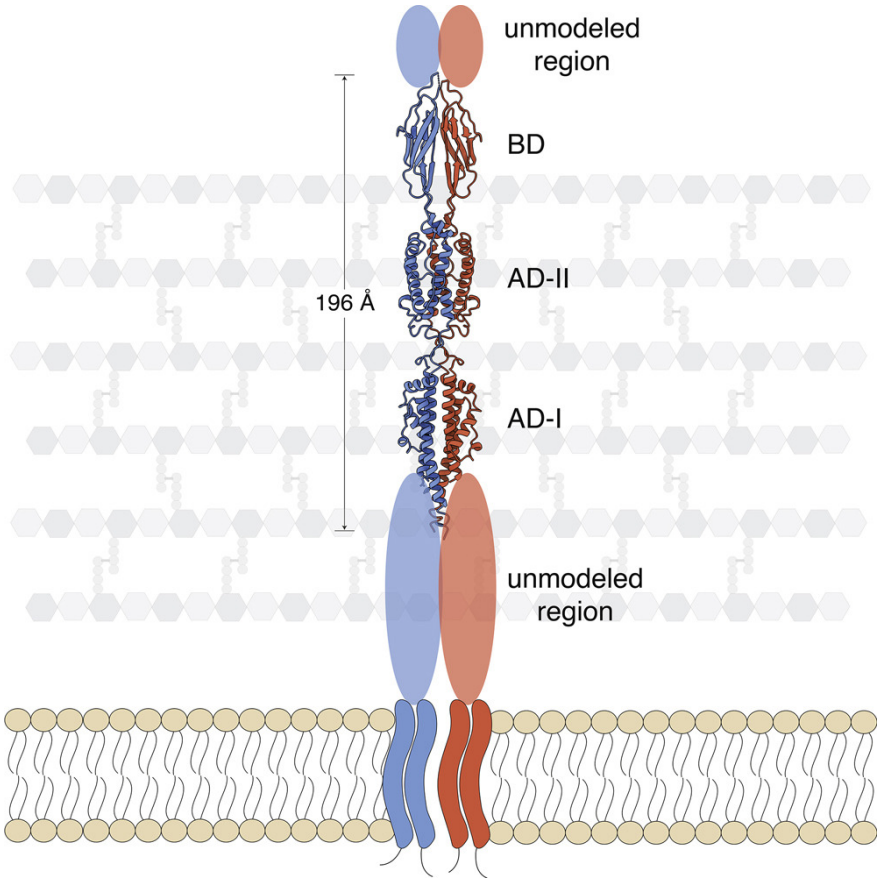


Figure 3.5: Model of EsaA structure and topology in the cell. A *Sg*EsaA dimer (red and blue) extending through the membrane as suggested by the topology data in Figure 3.2. Both the N- and C-terminal regions of EsaA are membrane bound while the soluble domain extends out into the peptidoglycan layer of the cell. There is a distal unmodeled region that corresponds to a bacteriophage receptor. The full extension of EsaA is at minimum 196Å but this calculation does not take into account the two proximal and distal unmodeled regions.

Tables

Table 3.1: X-ray data collection and refinement statistics for *Sg*EsaA$_{329\text{-}727}$

	***Sg*Esa$_{329\text{-}727}$ (selenomethionine)**
Data Collection	
Wavelength (Å)	0.9792
Space group	$C222_1$
Cell dimensions	
a, b, c (Å)	74.4, 248.5, 81.0
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	55.34-2.40 (2.44-2.40)[a]
Total no. of reflections	139995
Total no. of unique reflections	29568
R_{merge} (%)[b]	11.6 (81.0)[a]
$I/\sigma I$	6.1 (1.0)[a]
Completeness (%)	98.7 (89.3)[a]
Redundancy	4.7 (2.8)[a]
$CC_{1/2}$	0.99 (0.51)[a]
Refinement	
Residues modeled	329-512, 555-727
R_{work} / R_{free} (%)[c]	21.46/26.16
No. atoms	
Protein	2947
Water	28
Average B-factors (Å^2)	
Protein	77.69
Water	58.79
Rms deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.427
Ramachandran plot (%)[d]	
Total favored	98.3
Total allowed	1.7
Coordinate error (Å)[e]	0.34

[a]Values in parentheses correspond to the highest resolution shell.
[b]$R_{merge} = \Sigma \Sigma |I(k) - \langle I \rangle| / \Sigma I(k)$ where $I(k)$ and $\langle I \rangle$ represent the diffraction intensity values of the individual measurements and the corresponding mean values. The summation is over all unique measurements.
[c]$R_{work} = \Sigma ||F_{obs}| - k|F_{calc}|| / |F_{obs}|$ where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. R_{free} is the sum extended over a subset of reflections (5%) excluded from all stages of the refinement.
[d]As calculated using MOLPROBITY (Chen et al., 2010).
[e]Maximum-Likelihood Based Coordinate Error, as determined by PHENIX (Adams et al., 2010).

Methods

Bacterial strains, plasmids and growth conditions

All *S. intermedius* strains were generated from the *S. intermedius* B196 wild-type background. *E. coli* XL-1 Blue was used for plasmid maintenance. *E. coli* BL21 (DE3) CodonPlus and B834 (DE3) were used for the expression of methionine and selenomethionine containing proteins, respectively. Genomic DNA isolated from *S. intermedius* B196 and *S. gallolyticus* ATCC 43143 was used for cloning *Si*EsaA and *Sg*EsaA, respectively (Olson et al., 2013; Schlegel et al., 2003). A complete list of bacterial strains can be found in Table S3.1. pET29b and pDL277-derived plasmids were used for protein expression in *E. coli* and *S. intermedius*, respectively. pET29b-derived plasmids were generated by restriction enzyme-based cloning using the NdeI and XhoI restriction endonucleases and T4 DNA ligase. All constructs lacked their native stop codon resulting in the fusion of a vector encoded C-terminal his_6-tag to facilitate protein purification after expression in *E. coli*. Cloning into pDL277 was performed similarly except with the BamHI and SalI restriction endonucleases. Additionally, the P96 promoter sequence from *Streptococcus pneumoniae* was fused upstream of all genes of interest using splicing by overlap extension (SOE) PCR to allow for gene expression in *S. intermedius* (Lo Sapio et al., 2012). All cysteine point mutations were generated by SOE PCR followed by restriction-enzyme based cloning into either pET29b or pDL277 with the abovementioned enzymes. A complete list of plasmids can be found in Table S3.2. All *E. coli* strains were grown overnight in lysogeny broth at 37°C at 225 rpm in a shaking incubator. Kanamycin (50 mg/mL) was added to the growth media for strains

containing pET29b plasmids. All *S. intermedius* strains were grown in Todd Hewitt Broth supplemented with 0.5% yeast extract (THY) in a 37°C stationary 5% CO_2 incubator. To ensure uniform growth rate, all *S. intermedius* strains were grown first on THY agar plates for 1-2 days prior to growth in THY broth. Strains harboring pDL277-derived plasmids were grown in media supplemented with spectinomycin (50mg/mL for *S. intermedius* or 100mg/mL for *E. coli*).

DNA manipulation

S. intermedius and *S. gallolyticus* genomic DNA was prepared by resuspending cell pellets in InstaGene Matrix (Bio-Rad). Primers were synthesized by Integrated DNA Technology (IDT). Molecular cloning was performed using Q5 polymerase, restriction enzymes, and T4 DNA ligase from New England Biolabs (NEB). Sanger sequencing was performed by Genewiz Incorporated.

Transformation of *S. intermedius*

S. intermedius transformation with either plasmid or linear DNA were performed as previously described (Tomoyasu et al., 2010). In short, overnight cultures were back diluted 1:10 into 2 ml THY broth supplemented with 3 mL of 10 mg/ml *S. intermedius* competence stimulating peptide (DSRIRMGFDFSKLFGK, synthesized by Genscript) and incubated at 37°C, 5% CO_2 for 2 hours. Approximately 100-500ng of plasmid, or linear insert DNA was added and cultures were briefly vortexed before incubating for

another 3 hours. 100 ml of culture was then plated on the appropriate selective media (either 50 mg/ml spectinomycin, 250 mg/ml kanamycin, or both).

Gene deletion in *S. intermedius* by allelic replacement

SOE PCR was used generate a pETDuet-1 plasmid containing the *kanR* cassette from the pBAV1K plasmid under the control of the spectinomycin promoter from pDL277 (Bryksin and Matsumura, 2010). The spectinomycin promoter-kanamycin resistance cassette was cloned between the 1000 base pairs of DNA that flank the 5' and 3' ends of *esaA* including the first 15 bases of the *esaA* ORF at the end of 5' flank and the last 15 bases of *esaA* at the start of the 3' flank. The final plasmid for allelic replacement was pETDuet-1::5'*esaA*flank_SpecPromoter_*kanR*_3'*esaA*flank. This plasmid was then digested with BamHI and NotI and the resulting insert (5'*esaA*flank_SpecPromoter_*kanR*_3'*esaA*flank) was gel extracted (Monarch DNA Gel Extraction Kit, NEB). 100 ng of purified insert was transformed into *S. intermedius* B196 and plated onto THY agar plates supplemented with 250 μg/ml of kanamycin. PCR was used to confirm deletion of *esaA*.

Secretion assays

Overnight cultures of *S. intermedius* strains were centrifuged at 7600 *g*, resuspended in 1.6 ml fresh THY broth. These washed cultures were then used to inoculate 5 ml THY broth in 15 ml polypropylene centrifuge tubes to an initial OD of 0.1. Cells were harvested (4000 rpm, 4°C, 15 min) when they reached OD_{600} 0.7-0.9 and

supernatant fractions were prepared as follows. 3.5 ml of supernatant was removed and filtered through a 0.2 µm membrane to remove remaining cells. Proteins were precipitated at 4°C for 30 minutes by adding 700 µl of cold 100% trichloroacetic acid (TCA, final concentration 16.7%). Precipitant was collected by centrifugation (swinging-bucket, 4600 rpm, 4°C, 30 min), and washed 3 times with 500 ul of cold acetone. Precipitant was then air dried in a fume hood for at least 30 minutes before being dissolved in 20 µl resuspension buffer (50 mM Tris:HCl pH 8.0, 150 mM NaCl, 1X protease cocktail inhibitor). Cell fractions were prepared as follows. Cell pellets were washed with 1 ml PBS, transferred to a 2 ml centrifuge tube, re-pelleted (10,000 *g*, 4°C, 10 min), decanted and snap frozen at -80°C. Washed pellets were then resuspended in 50 ml of lysis buffer (50 mM Tris:HCl pH 8.0, 150 mM NaCl, 10 mg/ml lysozyme, 1X protease cocktail inhibitor), and incubated at 37°C for half an hour. Cell numbers were matched across samples by diluting cells in PBS based on final culture OD_{600}. Matched samples were then prepared for western blotting by mixing 2:1 with 4X SDS-PAGE loading dye (125 mM Tris:HCl pH 6.8, 20% *v*/*v* glycerol, 0.01% *w*/*v* bromophenol blue, 4% *v*/*v* BME), heated at 95°C for 10 minutes, and centrifuged (21,000 *g*, room temperature, 15 minutes).

Antibody generation and western blot analyses

Custom polyclonal antibodies for *S. intermedius* EsaA, EsxA and SodA were generated for this study (Customer's Antigen Polyclonal Antibody Package, Genscript). C-terminally his_6-tagged *Si*$EsaA_{41\text{-}871}$, EsxA and SodA were purified as described in

"Protein purification and expression" except that PBS was used in place of Tris:HCl for all purification buffers. 10 mg of each protein was sent to Genscript for antibody production. Generation of the α-TelC antibody has been described previously (Whitney et al., 2017).

With the exception of EsxA, western blot analyses of protein samples were performed using a Tris-glycine gel and buffer system and a standard western blotting protocol. The SDS-PAGE system for EsxA blots required the use of a tris-tricine buffer system, which allows for the electrophoretic separation of low molecular weight proteins. After SDS-PAGE separation, proteins were wet-transferred to 0.45 μm PVDF membranes (80 V for 1 hour, 4°C). Cell and supernatant fractions were analyzed by Western blot using the protein-specific rabbit primary antibodies α-TelC (1:5000 dilution, 1.5 hours), α-EsaA (1:5000, 1 hour), α-EsxA (1:5000, 2 hours), α-SodA (1:5000, 30 minutes), α-VSV-G (1:3000, 1.5 hours) and a goat α-rabbit secondary antibody (Sigma, 1:5000, 45 minutes). Clarity Max Western ECL substrate (Bio-Rad) was used for chemiluminescent detection of the secondary antibody and all blots were imaged with a ChemiDoc XRS+ System (Bio-Rad).

NADase activity assay

The consumption of NAD^+ by *S. intermedius* culture supernatants was assayed as described previously (Whitney et al., 2017). Briefly, culture supernatants taken from mid-log cultures were concentrated 50-fold by spin filtration at 3000 *g* (10kDa MWCO) and then filtered through a 0.2 μm membrane. The samples were then incubated 1:1 with PBS

containing 2 mM NAD^+. Reactions were incubated overnight (approximately 16 hours) at room temperature. 6M NaOH was added to terminate the reaction which was then incubated for 15 minutes in the dark. Fluorescence (ex: 360nm, em: 530nm) was measured using a Synergy 4 Microplate Reader (BioTek Instruments).

Subcellular fractionation by ultracentrifugation

One liter of *S. intermedius* cultures were grown to OD_{600} = 0.8 and pelleted by centrifugation at 6000 *g*. Pellets were resuspended in 20 ml of lysis buffer (20 mM Tris:HCl pH 7.5, 150 mM NaCl, 2 mg/ml lysozyme), incubated at 37°C for one hour, and sonicated at 30% amplitude for three pulses of 30 seconds each. The insoluble cellular debris was cleared by centrifugation at 39,191 *g*. The resulting supernatant was then centrifuged for two hours at 200,000 *g* to isolate the membrane fraction. The supernatant (cytosolic fraction) was mixed 1:1 with Laemmli loading buffer. The membrane pellet was washed once with 20 mM Tris:HCl, pH 7.5, 150 mM NaCl, before being resuspended in Laemmli loading buffer. Cytoplasmic and membrane fractions were then analyzed by SDS-PAGE and Western blot.

Membrane topology mapping

The cysteine labelling experiment was adapted from Ruhe et al. (Ruhe et al., 2018). For our experiment, 20 ml cultures of *S. intermedius* strains were grown to OD_{600} = 0.5 and harvested by centrifugation at 4,000 *g* for 20 minutes. Cell pellets were then washed three times with PBS to remove any extracellular material. The pellets were

resuspended in 35 µl PBS, pH 7.2 and IRDye680LT-maleimide dye (LI-COR Biosciences) was added to cells to a final concentration of 40 µM. The reactions were incubated at room temperature for 20 minutes in a darkroom before being quenched by adding BME to final concentration of 6 mM. Cells were then harvested by centrifugation and washed three times with PBS supplemented with 6 mM BME. Washed pellets were resuspended in minimal SDS-loading dye and boiled for 10 minutes. Samples were run on SDS-PAGE and imaged with a Chemidoc system (Bio-Rad) using a red LED epi-illumination source and a 700nm/50mm band pass filter.

Protein expression and purification

E. coli BL21(DE3) CodonPlus strains containing pET29b-derived plasmids were grown to OD_{600} = 0.4 and protein expression was induced with 1 mM IPTG. The induced strains were incubated overnight (approximately 18-20 hours) in a 225 rpm shaking incubator at 18°C after which the cells were collected by centrifugation at 10,000 *g*. Cell pellets were resuspended in lysis/wash buffer (20 mM Tris-HCl pH 7.5, 300 mM NaCl, 10 mM imidazole) and sonicated four times at 30% amplitude for 30 seconds each to lyse cells. Cleared cell lysates were purified by affinity chromatography using a Ni-NTA agarose column. After passing cell lysates over the column, the Ni-NTA resin was washed four times using wash buffer and eluted with wash buffer supplemented with 400 mM imidazole. Protein samples were further purified by size exclusion chromatography using a HiLoad 16/600 Superdex 200 column connected to an AKTA protein purification system (Cytiva). For selenomethionine incorporated protein, the *E. coli* methionine

auxotroph strain B834 was used and grown in SelenoMethionine Media (Molecular Dimensions) supplemented with 40 mg/l L-selenomethionine. Growth of *E. coli* B834 and protein expression and purification were otherwise carried out similarly to *E. coli* BL21(DE3) CodonPlus as described above.

Crystallization and structure determination

Selenomethionine incorporated *Sg*EsaA$_{234\text{-}829}$ was concentrated to 10mg/ml by spin filtration using an Amicon Ultra Centrifugal filter unit with a 30kDa pore size (Millipore). Concentrated protein was screened for crystallization with the MCSG Crystallization Suite (Anatrace). Long, slender crystals formed in 0.2M $MgCl_2$, 0.1M Tris:HCl, pH 7.0, 10% (w/v) PEG 8000, after three weeks. Protein crystallization was optimized around this condition with crystals forming in 0.1M Tris:HCl pH 7.0-7.8 and 10-15% (*w*/*v*) PEG 8000. Crystals were cryo-protected in similar buffers supplemented with 20% (*v*/*v*) ethylene glycol. X-ray data were collected at the Structural Biology Center (SBC) sector 19-ID at the Advanced Photon Source. A total of 290 diffraction images of 0.5° for 0.5 sec/image were collected on a Dectris Pilatus3 X 6M detector with a crystal to detector distance of 540 mm. Data were indexed, integrated, and scaled using the *xia2* system (Winter et al., 2013).

The structure of selenomethionine incorporated *Sg*EsaA$_{234\text{-}829}$ was solved using the Se-SAD method using the AutoSol package in Phenix (Adams et al., 2010). The AutoBuild wizard was subsequently used for model building and the observed electron density allowed model building for residues 329-727 of *Sg*EsaA$_{234\text{-}829}$ with an unmodeled

gap between residues 514-554 (Terwilliger et al., 2008). Manual adjustments to the model were performed in COOT and model refinement was carried out with Phenix.refine resulting in final R_{work} and R_{free} values of 0.21 and 0.26, respectively (Afonine et al., 2012; Emsley and Cowtan, 2004). X-ray data collection and refinement statistics are listed in Table 3.1. The structural figures presented in this work were generated using the UCSF Chimera or UCSF ChimeraX software (Goddard et al., 2018; Pettersen et al., 2004).

Homology modeling of *Si*EsaA and PIP

Homology models of *Si*EsaA and *E. faecalis* V583 Phage Infection Protein (PIP) were generated based on our solved structure of *Sg*EsaA$_{329\text{-}727}$ using the PHYRE^2 one-to-one threading algorithm (Kelley et al., 2015). *Si*EsaA was modeled with 100% confidence over 325 residues. *E. faecalis* V583 PIP was modeled with 96% confidence over 341 residues.

Sequence alignments and conservation mapping

Protein sequence conservation was mapped onto the structure of *Sg*EsaA using the online ConSurf server (Ashkenazy et al., 2016). The multiple sequence alignment used in the calculation was generated as follows. The full-length protein sequence of *Sg*EsaA was used as a BLASTp query sequence against the NCBI Reference Protein Sequence database, restricted to the phylum Firmicutes, using otherwise default settings (Altschul et al., 1990). Full length sequences for the top 500 hits were downloaded and sequences

shorter than 750 amino acids were filtered out. A multiple sequence alignment using the remaining 434 sequences was generated using Clustal Omega and uploaded with the structure coordinates (Sievers et al., 2011). Dimer interface calculations for *Sg*EsaA, including buried surface area and ΔG^i, were performed by uploading structure coordinates to the PDBePISA server (Krissinel & Henrick, 2007). Pairwise alignment of *Sg*EsaA and *Si*EsaA was generated by M-Coffee, and then visualized alongside the *Sg*EsaA$_{329\text{-}727}$ modelled secondary structure using ESPript3 (Robert and Gouet, 2014; Wallace et al., 2006). The predicted protein disorder plot for *Sg*EsaA$_{329\text{-}727}$ was generated by the IUpred2A server using the default settings (Erdos and Dosztanyi, 2020; Meszaros et al., 2018).

SEC-MALS analysis

Size exclusion chromatography with multi-angle laser static light scattering was performed on *Si*EsaA$_{328\text{-}685}$ and *Sg*EsaA$_{332\text{-}725}$. The proteins were expressed and purified as described above, concentrated to 2 mg/ml by spin filtration and then run on a Superdex 200 column (GE Healthcare). MALS was conducted using a MiniDAWN and Optilab system (Wyatt Technologies). Data was collected and analyzed using the Astra software package (Wyatt Technologies).

Cysteine crosslinking experiments

For in vitro crosslinking experiments, each cysteine mutant was expressed in *E. coli* BL21 (DE3) CodonPlus and the resulting protein was purified by Ni-NTA affinity

chromatography. The eluted protein samples were exposed to environmental oxygen for 16 hours to allow for crosslinking to occur. Samples were then mixed 1:1 with Laemmli buffer either containing or lacking β-mercaptoethanol and analyzed by SDS-PAGE and Coomassie staining. The SDS-PAGE gels were imaged using a ChemiDoc MP system (BioRad).

In vivo cysteine crosslinking was conducted similarly except that a pDL277 plasmid-based system was used to express each cysteine mutant in a *S. intermedius* B196 Δ*esaA* background. *S. intermedius* strains were grown to OD_{600} = 0.8 and centrifuged at 4000 *g*. Pellets were resuspended in 200 µl of lysis buffer (20 mM Tris:HCl, pH 7.5, 150 mM NaCl, 1% *w*/*v* DDM, 10 mg/ml lysozyme) and incubated at 37°C for one hour. Samples were then sonicated three times at 30% amplitude for 15 seconds per pulse. Lysed samples were cleared by centrifugation at 21,130 *g* for 20 minutes. The supernatant was then removed and allowed to sit at room temperature for one hour to allow for crosslinking. Samples were analyzed by Western blot using an α-EsaA primary antibody

Data availability

The data supporting Chapter III can be found entirely within this thesis as well as at the link found below. Structure files and information pertaining to the structure of *Sg*EsaA are indexed in the protein data bank (PDB: 7JQE). For access to strains and plasmids used in this chapter please contact Dr. John Whitney.

Relevant links:

https://www.sciencedirect.com/science/article/pii/S0969212620304147

https://www.rcsb.org/structure/7JQE

Chapter IV – Dual targeting factors are required for LXG toxin export by the bacterial type VIIb secretion system

Preface

The work presented in this chapter was previously published in the following study:

Klein, T.A., Grebenc, D.W., Shah, P.Y., McArthur, O.D., Dickson, B.H., Surette, M.G., Kim, Y., and Whitney, J.C. (2022). Dual targeting factors are required for LXG toxin export by the bacterial type VIIb secretion system. Submitted April 14, 2022 to mBio. Resubmitted July 23, 2022.

Author contributions: T.A.K. and J.C.W conceived the study. All authors contributed to experimental design. T.A.K., O.M., and J.C.W. generated strains and plasmids. T.A.K. and P.Y.S. expressed, purified, and crystallized protein. T.A.K., D.W.G., Y.K., and J.C.W. solved and analyzed the crystal structure. T.A.K. and O.M. performed biochemical experiments. T.A.K., D.W.G., and J.C.W. analyzed the data. D.W.G. and B.H.D. performed structural modeling. M.G.S. provided the *S. intermedius* GC1825 strain and performed whole genome sequencing on this strain. T.A.K., D.W.G. and J.C.W. wrote the paper. All authors provided feedback on the manuscript.

Permission has been granted by the publishers to reproduce the material from this study.

Abstract

Bacterial type VIIb secretion systems (T7SSb) are multi-subunit integral membrane protein complexes found in Firmicutes that play a role in both bacterial competition and virulence by secreting toxic effector proteins. The majority of characterized T7SSb effectors adopt a polymorphic domain architecture consisting of a conserved N-terminal Leu-X-Gly (LXG) domain and a variable C-terminal toxin domain. Recent work has started to reveal the diversity of toxic activities exhibited by LXG effectors; however, little is known about how these proteins are recruited to the T7SSb apparatus. In this work, we sought to characterize genes encoding domains of unknown function (DUFs) 3130 and 3958, which frequently co-occur with LXG effector-encoding genes. Using coimmunoprecipitation-mass spectrometry analyses, *in vitro* copurification experiments and T7SSb secretion assays, we find that representative members of these protein families form heteromeric complexes with their cognate LXG domain and in doing so, function as targeting factors that promote effector export. Additionally, an X-ray crystal structure of a representative DUF3958 protein, combined with predictive modelling of DUF3130 using AlphaFold2, reveals structural similarity between these protein families and the ubiquitous WXG100 family of T7SS effectors. Interestingly, we identify a conserved FxxxD motif within DUF3130 that is reminiscent of the YxxxD/E "export arm" found in Mycobacterial T7SSa substrates and mutation of this motif abrogates LXG effector secretion. Overall, our data experimentally link previously uncharacterized bacterial DUFs to type VIIb secretion and reveal a molecular signature required for LXG effector export.

Introduction

Protein secretion is an essential aspect of bacterial physiology that plays a critical role in diverse cellular activities including interbacterial competition and infection of host cells (Klein et al., 2020; Green & Mecsas, 2016). Bacteria possess several protein export pathways, often referred to as secretion systems, that facilitate protein transport across the cell envelope. In general, these pathways consist of membrane proteins that form the secretion apparatus and effector proteins that transit the secretion system. One important property of protein secretion apparatuses is their ability to recognize and export a specific set of effector proteins among the myriad cytosolic proteins within a cell. In many well-characterized examples, effectors harbour a signal sequence that is recognized by the secretion apparatus and the recruitment of this signal sequence to the apparatus often requires the involvement of molecular chaperones (Parsot et al., 2003; Christie et al., 2014; Sala et al., 2014; Ahmad et al., 2020; Burkinshaw et al., 2018).

Bacteria encode two ubiquitous secretion systems known as the general secretory pathway (Sec) and the Twin-arginine translocase (Tat). In addition to Sec and Tat, many Gram-negative bacteria encode a series of specialized secretion systems, several of which span the entirety of the diderm cell envelope (Green & Mecsas, 2016). By contrast, a substantial number of Gram-positive bacteria possess a single specialized secretion pathway referred to as the type VII secretion system (T7SS) (Abdallah et al., 2007). In recent years, this pathway has been further differentiated into two subtypes, T7SSa and T7SSb, to reflect the substantial differences in protein subunits that comprise each secretion apparatus (Tran et al., 2021). The T7SSa is found in Actinobacteria where it

functions as an essential virulence factor for many pathogenic species of Mycobacteria with specific T7SSa pathways linked to diverse functions including phagosomal escape, metal ion homeostasis, and conjugation (van der Wel et al., 2007; Houben et al., 2007; Serafini et al., 2013; Gray et al., 2016). The T7SSb is found in Firmicutes and is involved in the pathogenesis of *Staphylococcus aureus*, *Streptococcus agalactiae* and *Streptococcus intermedius* (Burts et al., 2005; Hasegawa et al., 2017; Spencer et al., 2021). In addition, several recent studies have uncovered a role for this pathway in mediating antagonistic interbacterial interactions in *S. aureus*, *S. intermedius*, *Enterococcus faecalis* and *Bacillus subtilis* (Cao et al., 2016; Whitney et al., 2017; Chatterjee et al., 2021; Kobayashi, 2021). Both T7SS subtypes have an FtsK-SpoIIIE family ATPase known as EccC/EssC that is thought to energize effector secretion and export one or more small α-helical effectors belonging to the WXG100 protein family (Abdallah et al., 2007). Beyond these similarities, T7SSa and T7SSb require different sets of apparatus proteins and export different families of effector proteins (Tran et al., 2021).

LXG proteins are emerging as the predominant group of effectors exported by T7SSb pathways (Whitney et al., 2017; Chatterjee et al., 2021; Kobayashi, 2021; Ulhuq et al., 2020). These proteins possess a polymorphic domain architecture comprised of a conserved ~200 amino acid N-terminal LXG (Leu-X-Gly) domain and a variable C-terminal toxin domain (Zhang et al., 2012). The toxic activities of several LXG effectors have been biochemically characterized and includes toxin domains that hydrolyze NAD^+, disrupt peptidoglycan biosynthesis, depolarize membranes and degrade essential nucleic acids (Cao et al., 2016; Whitney et al., 2017; Ulhuq et al., 2020; Holberger et al., 2012).

By contrast, little is known about the function of LXG domains. Based on comparisons to other polymorphic toxin systems, Zhang and Aravind propose a role for this domain in effector recruitment to the T7SS apparatus (Zhang et al., 2012). This hypothesis is bolstered by recent bacterial two-hybrid analyses showing that an LXG domain encoded by *B. subtilis* physically interacts with the T7SSb subunit YukC/EssB (Tassinari et al., 2020). However, because this experiment relied on a heterologous expression system, it remains unclear if this interaction is sufficient to promote effector secretion or if other factors are additionally required. In support of the need for additional secretion factors, the three LXG effectors exported by *S. intermedius* B196 interact with effector specific Wxg proteins via their LXG domains. Furthermore, it was shown for the TelC effector that its cognate protein, WxgC, is required for its export (Whitney et al., 2017).

In contrast to LXG proteins, the secretion determinants for T7SSa effectors are better defined. In general, T7SSa effectors exist as obligate heterodimers and heterodimerization is a prerequisite for secretion. The archetypal example is EsxA (ESAT-6) and EsxB (CFP-10), which are secreted as a heterodimer by the ESX-1 system of *Mycobacterium tuberculosis* (Renshaw et al., 2005; Brodin et al., 2005). The co-secretion of these effectors requires a conserved YxxxD/E motif present at the unstructured C-terminus of EsxB (Champion et al., 2006; Daleke et al., 2012a). Biochemical characterization of the interaction between the EccC motor ATPase and EsxB suggests that this secretion signal facilitates effector export by inducing EccC multimerization (Rosenberg et al., 2015). Like EsxB, other families of T7SSa effectors such as EspB and members of the proline-glutamate (PE)/proline-proline-glutamate

(PPE) family possess YxxxD/E motifs and in all tested cases, this motif is required for effector export (Solomonson et al., 2015; Daleke et al., 2012a; Damen et al., 2020).

In the present study, we sought to systematically characterize the secretion determinants of LXG effectors exported by the T7SSb pathway. Using two model effectors from two different strains of *S. intermedius*, we find that members of the DUF3130 (also known as TIGR04197, and "Type VII secretion effector, SACOL2603 family"), and the DUF3958 protein families function as dual targeting factors that physically interact with and promote the secretion of their cognate effector. Using structural analyses, we find that DUF3130 and DUF3958 bear resemblance to WXG100 effectors; however, in contrast to these effectors, they are not exported by the T7SSb. While DUF3958 proteins lack conserved sequence motifs that could provide insight into their precise function, DUF3130 proteins possess a highly conserved FxxxD motif that resembles the secretion signal found in T7SSa substrates. Moreover, site-specific mutation of this motif abrogates LXG effector export. Overall, our work uncovers new intracellular factors involved in LXG effector secretion, provides molecular insights into how these factors function, and demonstrates that effector secretion by T7SSb pathways may share more similarities to their Mycobacterial T7SSa counterparts than previously appreciated.

Results

A DUF3958 protein is required for export of the LXG effector TelC from *S. intermedius* B196.

In a recent bioinformatics study on the LXG effector repertoire of *Listeria monocytogenes*, Bowran and Palmer noted the near ubiquitous existence of two small open reading frames upstream of LXG genes (Bowran & Palmer, 2021). In our initial characterization of the model LXG effector TelC from *S. intermedius* B196 we found that the protein product of one of these genes, WxgC, physically interacts with TelC and is required for its T7SSb-dependent export (Whitney et al., 2017). However, the function of the other LXG effector associated gene, SIR_1490, was not examined. WxgC and SIR_1490 have homology to the DUF3130 and DUF3958 protein families, respectively. Given the frequent co-occurrence of the genes encoding these proteins within LXG effector gene clusters, we hypothesized that SIR_1490 also plays a role in the export of TelC (Fig. 4.1A). To test this, we first generated an *S. intermedius* B196 strain lacking SIR_1490 and examined the ability of this strain to export TelC into culture supernatants. In line with our hypothesis, only intracellular TelC was detected in ΔSIR_1490 and TelC secretion could be restored by plasmid-borne expression of SIR_1490 (Fig. 4.1B). Export of the WXG100 effector EsxA, a hallmark of a functional T7SS apparatus, was unaffected by mutational inactivation of SIR_1490 indicating that the loss of TelC secretion is not due to a defect in T7SSb apparatus function (Fig. 4.1C) (Abdallah et al., 2007). Like EsxA, WxgC and SIR_1490 are predicted to be α-helical proteins of approximately 100 amino acids in length (Fig. S4.1). Therefore, we next considered the

possibility that these proteins are also exported by the T7SSb. In contrast to EsxA and TelC, we were unable to detect either of these proteins in the culture medium (Fig. 4.1D). Based on these data, we conclude that WxgC and SIR_1490 function as cytoplasmic factors that facilitate the T7SSb-dependent export of TelC. In light of these and subsequent findings, we propose to rename WxgC/SIR_1491 and name SIR_1490 to LXG-associated α-helical protein for TelC 1 (LapC1, DUF3130) and 2 (LapC2, DUF3958), respectively.

TelC, LapC1 and LapC2 physically interact to form a heterotrimeric pre-secretion complex.

Given our genetic data linking both *lapC1* and *lapC2* to the T7SSb-dependent export of TelC, we next wanted to examine whether the encoded proteins physically interact with TelC in the context of their native organism. To probe this, we expressed Vesicular Stomatitis Virus G (VSV-G) epitope tagged TelC (TelC-V) in *S. intermedius* B196, performed an immunoprecipitation using anti-VSV-G antibody, and identified proteins that were enriched relative to a control strain by mass spectrometry (Fig. 4.2A and Table 4.1). LapC1 was highly enriched in the TelC-V expressing strain, corroborating previous bacterial two-hybrid data in *E. coli* that indicated these proteins interact directly (Whitney et al., 2017). Interestingly, LapC2 was also highly enriched, suggesting that LapC2 also interacts with TelC. PepC, an aminopeptidase with no known role in type VII secretion, was also present in our TelC-V sample and absent in our control, although it was present in lower overall abundance as measured by total spectral

counts (Chapot-Chartier et al., 1994). We speculate that a small amount of PepC may interact with highly expressed proteins under some conditions but consider it unlikely that PepC is a bonafide interaction partner of TelC. To substantiate this assumption and to validate TelC's interaction with LapC1 and LapC2, we next performed a similar immunoprecipitation using VSV-G tagged LapC1 (LapC1-V) as the bait protein (Fig. 4.2B and Table 4.1). In this experiment, both TelC and LapC2 were enriched relative to the control sample, but PepC was not. Based on these data, we conclude that TelC, LapC1 and LapC2 interact to form an effector pre-secretion complex in *S. intermedius* B196.

Because protein-protein interactions identified by co-immunoprecipitation can be indirect in nature, we next attempted to co-express and purify TelC with LapC1 and LapC2 using an *E. coli* overexpression system. Previous bacterial two-hybrid data showed that the LXG domain of TelC ($TelC_{LXG}$) is both necessary and sufficient for LapC1 interaction (Whitney et al., 2017). Therefore, we similarly used $TelC_{LXG}$ to assess TelC-LapC-LapC2 heteromer formation. Using His_6-tagged $TelC_{LXG}$ to facilitate nickel affinity chromatography, we found that $TelC_{LXG}$ copurified with both LapC1 and LapC2 after nickel affinity and size exclusion chromatography (Fig. 4.2C). Taken together, our results indicate that the physical association of LapC1 and LapC2 with TelC's LXG domain promotes TelC export by the T7SS of *S. intermedius* B196.

TelD is a novel LXG-containing T7SSb effector that also requires a cognate Lap1-Lap2 pair for export.

To test the generalizability of our findings on TelC, LapC1 and LapC2, we next sought to determine if a second LXG effector also requires heterocomplex formation with a cognate Lap1-Lap2 pair to facilitate its secretion by the T7SS. To this end, we examined the recently sequenced GC1825 strain of *S. intermedius* and identified a candidate LXG-domain containing T7SS effector, which we named *telD* to remain consistent with the established *S. intermedius* T7SS LXG effector nomenclature (Whitney et al., 2017). The *telD* gene neighbourhood has similar synteny to that of *telC* in that the effector gene is found immediately downstream of *lapC1* and *lapC2* homologous genes, which we henceforth refer to as *lapD1* and *lapD2*, respectively to reflect their linkage to *telD* (Fig. 4.3A). Downstream of *telD* are two DUF443-encoding genes, which belong to a family of proteins that contain TsaI, a characterized immunity protein for the membrane depolarizing LXG effector TspA of *S. aureus* (Ulhuq et al., 2020). The final ORF in the predicted operon is a DUF4176-encoding gene, members of which are often found among T7SS genes but whose function is unknown (Tran et al., 2021; Bowman et al., 2021).

We first wanted to determine if TelD is indeed a T7SS effector as would be predicted due to it possessing an N-terminal LXG domain. To test this, we deleted the gene encoding the essential T7SSb component, *essB*, and examined TelD secretion by western blot using a TelD-specific antibody. Our results show that in contrast to wild-type *S. intermedius* GC1825, the T7SS-inactivated strain is unable to secrete TelD (Fig. 4.3B). Of note, we observed lower levels of intracellular TelD in the Δ*essB* strain relative

to wild-type and the reason for this is currently unclear. Nonetheless, complementing *essB* in trans resulted in a partial restoration of cellular TelD levels and a complete restoration of TelD export suggesting that TelD is secreted in a T7SS-dependent manner. Antibacterial activity is a property of all LXG toxins characterized to date, so we next wanted to examine if TelD is also toxic to bacterial cells. Consistent with this precedent, we found that expression of the TelD toxin in *E. coli* led to an approximate 100-fold decrease in cell viability (Fig. 4.3C). Furthermore, when grown in liquid culture, we observed that TelD caused *E. coli* growth arrest shortly after induction of toxin expression but did not cause cell lysis (Fig. 4.3D). Finally, co-expression of the adjacent DUF443-encoding gene, henceforth referred to as *tipD* (Tel immunity protein D), substantially restored *E. coli* growth (Fig 4.3D). Given that it shares the same family of predicted immunity proteins as TspA, TelD may similarly inhibit growth via membrane depolarization. However, while their LXG domains possess 29.4% sequence identity and are predicted to have nearly identical secondary structure, the toxin domains are only 13% identical and yield substantially different structural predictions (Fig. S4.2). Therefore, this putative activity will require experimental validation. In sum, these data indicate that TelD is a T7SS effector with antibacterial properties.

Having established that TelD is a substrate of *S. intermedius* GC1825's T7SS, we next examined the dependency of its secretion on *lapD1* and *lapD2*. To this end, we generated *S. intermedius* GC1825 strains lacking either *lapD1* or *lapD2*. Interestingly, and in contrast to TelC, we found that overall TelD levels were greatly diminished in the absence of *lapD1* and below the limit of detection in a *lapD2* deletion strain (Fig. 4.3E

and Fig. 4.3F). Consistent with our findings on TelC, *TelD* export was also abrogated in the strain lacking *lapD1*. Importantly, cellular levels of TelD as well as its export via the T7SS could be restored by complementing each deletion strain with a plasmid-borne copy of the deleted gene (Fig. 4.3E and Fig. 4.3F). The decrease in cellular TelD levels differs from our findings with TelC and suggests that LXG effectors have differing levels of intrinsic stability. In the case of TelD, our data indicate that in addition to being required for effector export, LapD1 and LapD2 are exhibiting chaperone-like properties by stabilizing their cognate effector prior to its export from the cell. Similar to TelC, we found that the LXG domain of TelD ($TelD_{LXG}$) forms a stable heteromeric complex with LapD1 and LapD2 when overexpressed in *E. coli* and copurified using nickel affinity and size exclusion chromatography (Fig. 4.3G). In summary, our TelD data corroborate our findings on TelC by showing that LXG effector secretion, and in some cases LXG effector stability, requires the activities of genetically linked Lap1 and Lap2 proteins.

A crystal structure of LapD2 reveals its structural similarity to the WXG100 family of T7SS effectors.

To better understand the molecular basis for DUF3130 and DUF3958 function, we initiated protein crystallization experiments on six representative members of each protein family including those linked to TelC and TelD export. Unfortunately, most of the Lap1 and Lap2 proteins that we tested expressed poorly or were recalcitrant to crystallization. Despite this discouraging trend, LapD2 was the sole exception and after optimization it formed crystals that diffracted to 2.2Å. The crystallographic phase

problem was overcome using selenomethionine-incorporated protein and the single wavelength anomalous dispersion (SAD) technique. The final model of LapD2 was refined to an R_{work}/R_{free} of 0.23/0.26 using the native diffraction data (Table 4.2).

The overall structure of LapD2 shows that it adopts a helix-turn-helix fold that is reminiscent of the WXG100 family of small secreted T7SS effectors (Fig. 4.4A). However, in contrast to characterized WXG100 proteins, which typically form head-to-toe homodimers mediated by hydrophobic interactions, the turn region of LapD2 contains an intermolecular disulfide bond formed by cysteine 59 that facilitates head-to-head dimerization (Fig. 4.4A and Fig. 4.4B) (Poulsen et al., 2014). The head region of LapD2 also possesses a hydrophobic patch that may also contribute to dimerization (Fig. 4.4C). Not surprisingly, the energy of head-to-head dimer formation as predicted by the PDBePISA webtool is highly favourable (Δ^iG = -21.0) due to the combined effects of burying a hydrophobic patch from the aqueous milieu and possessing a disulfide linkage (Krissinel & Henrick, 2007). PDBePISA also revealed a toe-to-toe homodimer interface and this interaction was also suggested to be favourable, although with a lower energy of formation (Δ^iG = -11.6) (Fig. S4.3). A search for proteins that are structurally homologous to LapD2 using the DALI webserver identified over 12,000 proteins with significant similarity (Z-score > 2) (Holm, 2020). The enormity of this list is due to helix-turn-helix motifs being a common structural element found in numerous proteins of diverse function with the most frequently occurring in our list being DNA-binding proteins. As alluded to above, WXG100 proteins were also well represented with 65 WXG100 family protein structures scoring as significantly similar to LapD2. The top

WXG100 hit was a structure of the EsxB protein exported by the ESX-1 T7SSa of *M. tuberculosis* (PDB: 3FAV, Z-score = 8.4, R.M.S.D. = 2.6Å over 90 aligned residues) (Fig. S4.3).

We next wanted to determine what structural aspects of LapD2 play a role in facilitating TelD secretion. To initiate this, we first generated a sequence alignment of 95 unique homologous proteins identified using three iterations of the JackHMMER algorithm and mapped the resulting sequence conservation onto the structure of LapD2 (Fig. 4.4D and Table 4.3). This analysis revealed that Lap2 proteins generally have low sequence conservation. For example, four randomly selected sequences from our list each share approximately 19% pairwise sequence identity to either LapC2 or LapD2 (Fig. S4.4). An alignment using all identified homologs reveals a pattern of hydrophobic residues, particularly leucine, at conserved positions that are interspersed between regions that favour charged and polar residues (Fig. S4.4). Notably, conservation is very low within the interhelical turn region, which contrasts with the highly conserved WXG motif found within structurally similar WXG100 proteins (Poulsen et al., 2014). We therefore speculate that shape and/or the surface properties of this protein family may be more critical to function than specific motifs within the primary sequence.

One of the more striking features of LapD2 is the disulfide bond formed by Cys59 that contributes to dimerization. This residue is not conserved among Lap2 proteins indicating that an intermolecular disulfide bond is likely not a universal property of this protein family. Nonetheless, we reasoned that its unique involvement in LapD2 dimerization warranted its functional interrogation in the context of TelD stability and

secretion. To accomplish this, we first mutagenized Cys59 to serine (C59S) and confirmed that this variant could no longer form β-mercaptoethanol sensitive dimers *in vitro* (Fig. 4.4E). We next assessed the ability of a strain expressing LapD2^{C59S} to export TelD into culture supernatants. Consistent with not playing an important role in LapD2 function, we found that an *S. intermedius* GC1825 Δ*lapD2* strain expressing plasmid-borne LapD2^{C59S} secretes wild-type levels of TelD (Fig. 4.4F). Furthermore, the ability of LapD2 to form a heteromeric complex with LapD1 and TelD$_{LXG}$ was unaffected by this mutation (Fig. 4.4G). Finally, we noted that although LapD2 readily forms a Cys59 mediated cross-link when purified in isolation, this dimeric species is not observed when it is purified in complex with LapD1 and TelD$_{LXG}$ (Fig. 4.4E and Fig. 4.4G). Together, these data are indicative of the function of Lap2 proteins being less reliant on specific amino acids and more reliant on global aspects of protein structure.

AlphaFold2 predicted models of Lap1 proteins reveals the location of a conserved FxxxD motif required for LXG effector secretion.

Despite extensive efforts, we were unable to solve a crystal structure of a Lap1 protein. In general, we found that Lap1 proteins do not express well and were therefore poor candidates for crystallization experiments. Therefore, to better understand Lap1 function we used the recently released AlphaFold2 network to generate models of LapC1 and LapD1 (Fig. 4.5A and Fig. S4.5) (Jumper et al., 2021). In parallel, we also ran AlphaFold2 on LapD2 and aligned the resulting model with our experimental crystal structure. As might be expected for a small single domain protein, the experimental and

predicted models generally aligned well with a Ca RMSD of 2.0Å (Fig. S4.5). However, we did note that the position of the turn region that connects the two α-helices occurs approximately seven residues earlier in the AlphaFold2 model compared to our crystal structure. Nonetheless, this result gave us reasonable confidence in the ability of AlphaFold2 to accurately predict the overall structure of Lap1 as members of this protein family share a similar size and predicted α-helical content as Lap2 proteins (Fig. S4.1). Additionally, AlphaFold2 suggests that Lap1 proteins may have a propensity to dimerize (Fig S4.5)

Overall, the AlphaFold2 generated Lap1 models adopt a helix-turn-helix arrangement similar to Lap2, with the exception of the first α-helix, which is markedly shorter than the second α-helix (Fig. 4.5A and Fig. S4.5). An alignment of 203 unique homologous Lap1 proteins was generated for LapC1 using one iteration of JackHMMER (Table 4.4) (Finn et al., 2015). In contrast to Lap2, conservation mapping of Lap1 onto the predicted structure of LapC1 revealed two highly conserved regions within this protein family (Fig. 4.5B). The first lies in the interhelical turn region and consists of a DxxTxxxGN motif (Fig. 4.5B and Fig. 4.5C). We speculate that this motif is likely important for protein folding as the conserved residues face inwards towards one another and the side chains of Thr36 and Asn42 are predicted to hydrogen bond to one another based on their 2.9Å proximity. The second conserved region is solvent exposed, exists near the end of the second α-helix, and is punctuated by an FxxxD motif (Fig. 4.5B and Fig. 4.5D). This motif drew our attention because it is remarkably similar to the YxxxD/E ‘export arm’ that serves as a secretion signal for Mycobacterial T7SSa effectors.

Structural alignment of the predicted LapC1 structure with the crystal structures of the characterized Mycobacterial T7SSa effectors EspB and PE25 shows a striking overlap in the three-dimensional position of these residues, despite LapC1 possessing less than 15% sequence identity with either protein (Fig. 4.5E and Fig. S4.6).

Given our data suggesting that LapC1 itself is not secreted, we hypothesized that the FxxxD motif may act as an effector recognition signal that guides LXG proteins to the T7SSb when they are part of a Lap1-Lap2-LXG effector complex. To test this, the residues comprising this motif in LapD1 were targeted for site-specific mutagenesis to probe their role in TelD export. In line with functioning as a T7SSb export motif, we found that the secretion of TelD is not restored by plasmid-borne expression of LapD1^{F77A} or LapD1^{D81A} variants in an *S. intermedius* GC1825 Δ*lapD1* background (Fig. 4.5F). To ensure that the observed lack of TelD secretion in these strains was not due to these site-specific variants compromising TelD-LapD1-LapD2 complex formation, we also introduced these mutations into our *E. coli* co-expression system and purified the LapD1 variant containing protein complexes. The results from this experiment demonstrate that LapD1^{F77A} or LapD1^{D81A} copurify with TelD$_{LXG}$ and LapD2 in a manner that is comparable to wild-type LapD1 (Fig. 4.5G). Collectively, these data show that the FxxxD motif of Lap1 proteins is not required for the formation of an effector pre-secretion complex but that it plays an essential role in LXG effector secretion by the T7SSb apparatus (Fig. 4.6). This conclusion is supported by AlphaFold2 modelling of the TelC-LapC1-LapC2 complex, which places the conserved LxG motif of TelC's LXG domain in close spatial proximity to the FxxxD motif of LapC1 (Fig S4.5E)

Discussion

We have found that representative members of the Lap1/DUF3130 and Lap2/DUF3958 families of proteins function as targeting factors that promote the T7SSb-dependent secretion of cognate LXG effector proteins. Our structural and functional investigation also led us to discover that the former of these protein families possesses a critical sequence motif required for effector export. Altogether, these findings reveal several interesting parallels between the LXG-Lap1-Lap2 complexes defined herein and several well-characterized T7SSa effector families. For example, PPE proteins of *M. tuberculosis* are characterized by N-terminal domains of approximately the same size (~180-200 amino acids) and α-helical content as is predicted for LXG domains (Ulhuq et al., 2020; Gey van Pittius, 2006). Moreover, these proteins are often encoded and expressed alongside members of the PE family of proteins, which like Lap1/Lap2, are ~100 amino acids in length, adopt a helix-turn-helix fold, and physically interact with their adjacently encoded effector (Ates, 2020). Several solved co-crystal structures of PE-PPE heterodimers demonstrates that these protein complexes form elongated α-helical bundles (Strong et al., 2006; Korotkova et al., 2014; Williamson et al., 2020; Ekiert et al., 2014). Like Lap1 and Lap2, LXG domains are predicted to adopt an elongated α-helical structures and thus we speculate that LXG-Lap1-Lap2 heteromers may similarly adopt a side-by-side α-helical packing arrangement (Ulhuq et al., 2020).

Another notable similarity between the PE and Lap1 protein families is the position of a conserved C-terminal motif, which in PE proteins and other T7SSa effectors is defined as Yxxx[D/E] whereas we identified a FxxxD motif in Lap1 proteins

(Solomonson et al., 2015). In T7SSa effectors and EsxA proteins, this motif constitutes the so-called "export arm" and along with a WXG motif on a partner protein, functions as a bipartite secretion signal involved in the recruitment of effectors to the T7SS translocase EccC/EssC (Rivera-Calzada et al., 2021; Mietrach et al., 2020b). However, characterized PE proteins with this export motif are also typically co-secreted along with their partner PPE effector whereas we were unable to detect Lap1 or Lap2 in our secretion assays. While this finding could be due to the sensitivity of our measurements, it is also suggestive of a model in which these targeting factors dissociate from their cognate LXG effector during the secretion process. It is also interesting to note that some PE-PPE effector pairs also require a member of the globular EspG chaperone family to guide them to the T7SSa apparatus (Daleke et al., 2012b). While a globular chaperone, EsaE, has been shown to play a critical role in the T7SSb-dependent secretion of the non-LXG effector EsaD from *S. aureus*, a gene encoding a homologous protein does not exist in the *telC* and *telD* gene clusters (Cao et al., 2016).

The necessity of cognate Lap1–Lap2 targeting factors for the secretion of TelC and TelD is also interesting in the context of other LXG effectors. The TelC-producing B196 strain of *S. intermedius* secretes two additional LXG effectors named TelA and TelB, neither of which are encoded in gene clusters containing *lap1* or *lap2* homologous genes. TelA and TelB are instead encoded downstream of members of the DUF5082 and DUF5344 families of proteins, both of which are predicted helix-turn-helix proteins (Bowman & Palmer, 2021). Bacterial two-hybrid studies on the TelA- and TelB-associated DUF5082 proteins have shown that like Lap1 and Lap2, they physically

interact with the LXG domain of their adjacently encoded LXG effector (Whitney et al., 2017). Therefore, we speculate that these proteins likely play a similar role to the Lap1-Lap2 pairs described herein. The *S. aureus* effector TspA presents yet another intriguing case. In contrast to the LXG effectors of *S. intermedius*, the *tspA* operon consists of the effector gene followed by multiple copies of the immunity factor *tsaI* but no members of the small α-helical DUF families described above (Ulhuq et al., 2020). This may indicate that the secretion of TspA does not require targeting factors for secretion or that TspA secretion requires the presence of small α-helical proteins encoded by ORFs found outside of the *tspA* gene cluster. Interestingly, gene clusters in *S. aureus* strains that encode the T7SSb apparatus often possess multiple genes encoding predicted small α-helical proteins, including EsxA, EsxB, EsxC and EsxD. All four of these proteins are secreted and either homo- or heterodimerize (Anderson et al., 2013; Sundaramoorthy et al., 2008; Burts et al., 2008). Based on our findings, it is conceivable that one or more of these Esx proteins may function as targeting factors for TspA and/or other T7SSb effectors secreted by this bacterium.

While our identification and characterization of the factors required for LXG effector export has yielded new insight into the process of protein secretion by the T7SSb, future work structurally characterizing the identified three-protein complexes is required to better understand how LXG effector recognition by the T7SSb apparatus occurs at the molecular level. Studies on effector recognition by T7SSa pathways suggests that the EccC/EssC translocase may facilitate this recognition (Rosenberg et al., 2015). However, more recent work on the T7SSb of *B. subtilis* found that LXG effectors

directly interact with YukC (EssB), a protein that the authors of this study propose serves as the central interaction hub that holds the T7SSb apparatus together (Tassinari et al., 2020). Regardless of which apparatus protein(s) recognise LXG effectors, our data suggests that the 'signal sequence' that allows for this recognition is likely defined by the quaternary structure of LXG-Lap1-Lap2 complexes and the FxxxD export motif found within Lap1. Upon interaction with the apparatus, we hypothesize that LXG effector export is facilitated by a conformational change in the T7SSb structure that is energetically linked to ATP binding and hydrolysis by EssC. Effectors are then transported through the cell envelope in a single step via a protein channel comprised of the various T7SSb structural subunits, the molecular details of which remain obscure. Several recent cryo-EM structures of Mycobacterial T7SSa apparatuses have provided profound mechanistic insights into the function of T7SSa pathways and it is probable that structures of the T7SSb will similarly inform our understanding of protein export by this complex molecular machine (Famelis et al., 2019; Poweleit et al., 2019; Bunduc et al., 2021; Beckham et al., 2021).

Figures

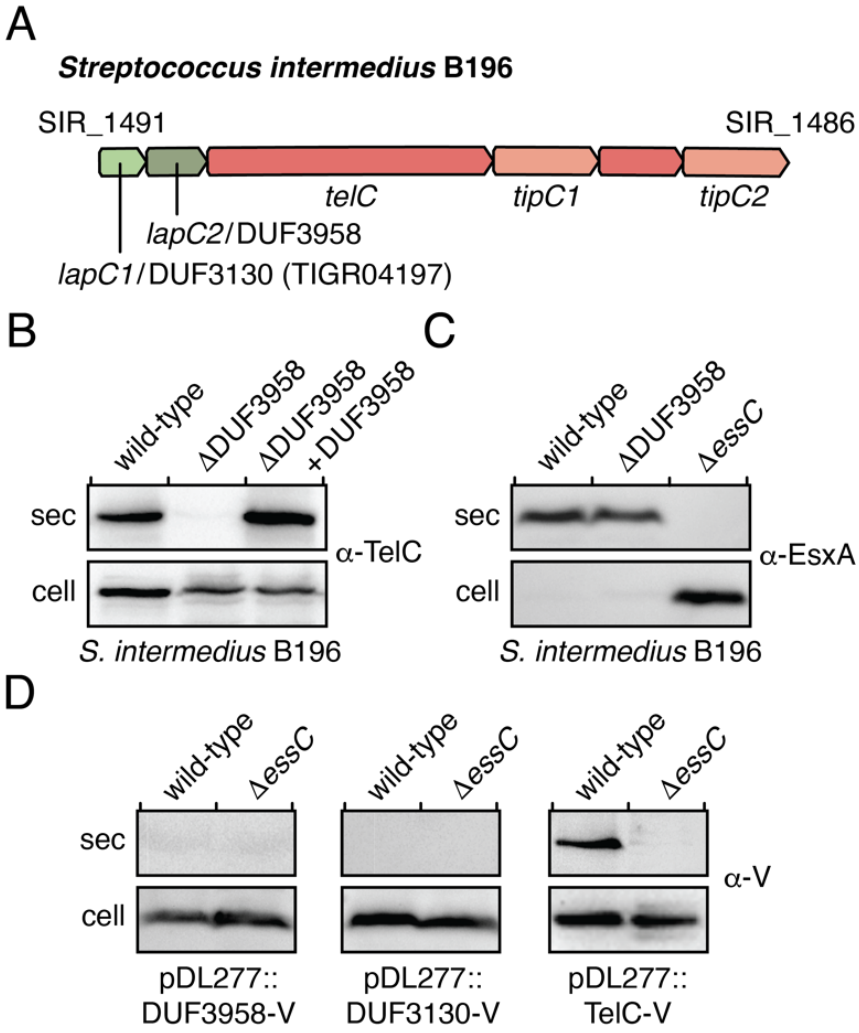


Figure 4.1: SIR_1490 encodes a DUF3958 protein required for the export of LXG effector TelC. (A) Schematic of the *telC* gene cluster from *S. intermedius* B196. Locus tags and gene names/DUF families are provided above and below the gene diagram, respectively. Genes are coloured to signify their function/context: light green – DUF3130 homolog (*lapC1*), dark green – DUF3958 homolog (*lapC2*), orange – LXG effector (*telC*) or orphan toxin domain (SIR_1487), salmon – immunity genes. (B-D) Western blot analysis of the secreted (sec) and cell fractions of the indicated *S. intermedius* B196

strains. Protein specific antibodies were used to detect endogenous TelC and EsxA (B and C) and anti-VSV-G epitope antibody was used to detect ectopically expressed VSV-G-tagged SIR_1490 (DUF3958-V) and VSV-G-tagged SIR_1491 (DUF3130-V) (D). *S. intermedius* B196 Δ*essC* is a T7SS-deficient control. The pDL277::DUF3958-V complementation vector used in (B) is the same as that used to assess secretion in (D).

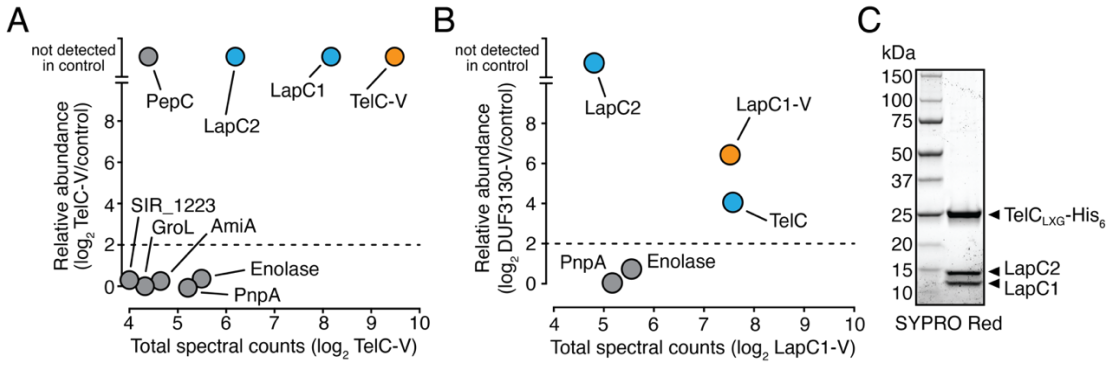


Figure 4.2: LapC1 and LapC2 interact with the LXG domain of TelC to form an effector pre-secretion complex. (A and B) Mass spectrometry analysis of immunoprecipitated VSV-G tagged TelC (TelC-V) (A) and LapC1 (LapC1-V) (B). Total spectral counts of abundantly detected proteins and fold enrichment relative to a control strain are plotted on the X- and Y-axes, respectively. In both panels, the immunoprecipitated protein is coloured orange while interaction partners are coloured blue. (C) SYPRO Red stained gel showing purified $TelC_{LXG}$–LapC1–LapC2 complex. Proteins were co-expressed in *E. coli* and purified using nickel affinity and size exclusion chromatography.

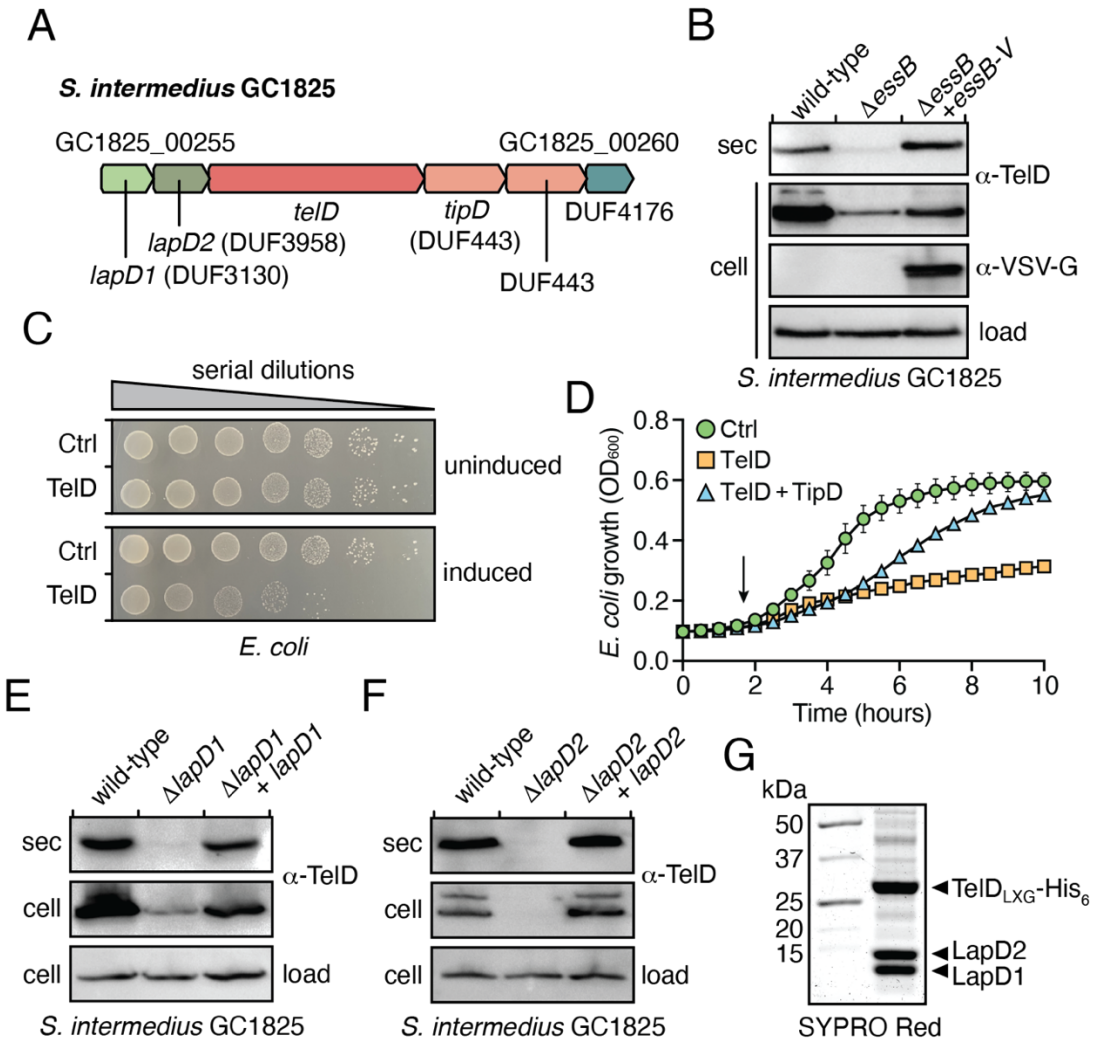


Figure 4.3: Lap1 and Lap2 proteins are required for secretion of the novel LXG effector TelD. (A) Schematic of the *telD* gene cluster from *S. intermedius* GC1825. Locus tags and gene names/DUF families are provided above and below the gene diagram, respectively. Genes are coloured to signify their function/context: light green – DUF3130 homolog (*lapD1*), dark green – DUF3958 homolog (*lapD2*), orange – LXG effector (*telD*), salmon – immunity genes, blue – DUF4176. (B-D) *S. intermedius* GC1825 TelD is an antibacterial toxin that is exported in a T7SS-dependent manner. Western blot analysis of the secreted (sec) and cell fractions of the indicated *S.*

intermedius GC1825 strains (B). CFU plating (C) and growth curves (D) of *E. coli* cells expressing TelD, TelD with the TipD immunity protein or a vector control (Ctrl). In panel D, arrow indicates when inducer was added, and error bars represent SEM. (E and F) Western blot analysis of the secreted and cell fractions of the indicated *S. intermedius* GC1825 strains. (G) SYPRO Red stained gel of purified $TelD_{LXG}$–LapD1–LapD2 complex. In all panels containing western blots, a TelD specific antibody was used to detect endogenous TelD and a cross-reactive band was used as a loading control.

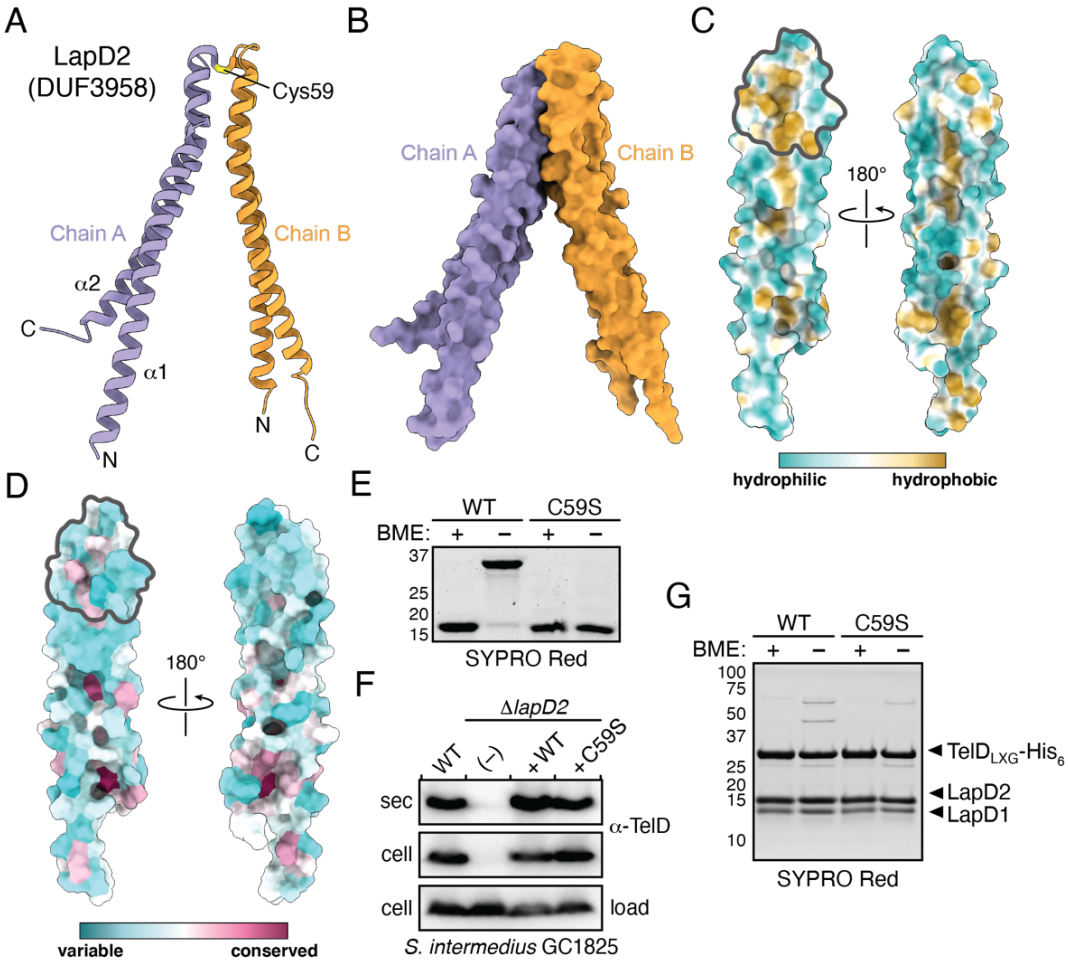


Figure 4.4: LapD2 is a small α-helical protein reminiscent of WXG100 superfamily proteins. (A and B) Overall structure of LapD2. LapD2 is shown as ribbon (A) and space-filling (B) representations with secondary structure elements, intermolecular disulfide bond, chain identities, and termini labelled where appropriate. (C) Hydrophobicity analysis of LapD2's surface as calculated by ChimeraX (Goddard et al., 2018). The LapD2 homodimerization interface is denoted by a grey outline. (D) Surface representation of Lap2 sequence conservation mapped onto the LapD2 structure. Sequences used for conservation analysis are available in Table 4.3. (E-G) Mutation of Cys59 to serine abrogates covalent dimer formation but does not impede TelD secretion or its ability to

interact with LapD1 and LapD2. SYPRO Red staining of purified LapD2 and LapD2^{C59S} in the presence and absence of β-mercaptoethanol (BME) (E). Western blot analysis of the secreted and cell fractions of the indicated *S. intermedius* GC1825 strains (F). SYPRO Red staining of purified TelD$_{LXG}$-LapD1-LapD2 and TelD$_{LXG}$-LapD1-LapD2^{C59S} complexes (G).

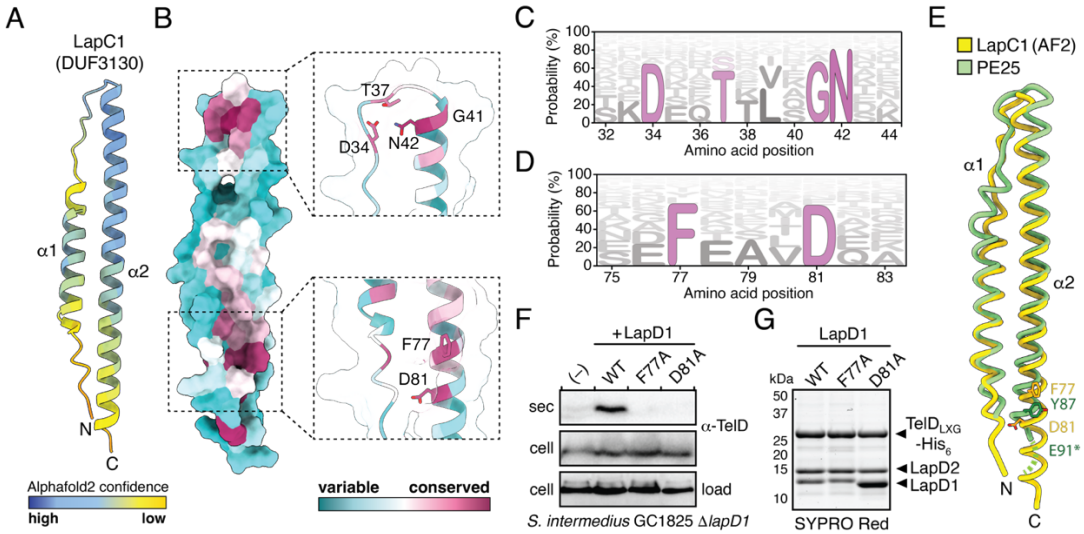


Figure 4.5: Lap1 modelling predicts a small α-helical protein harbouring a T7SSa export motif. (A) AlphaFold2 predicted structure of LapC1. Model is shown as a ribbon representation and coloured according to AlphaFold2 confidence level. (B) Surface representation of Lap1 sequence conservation mapped onto the LapC1 predicted structure. Sequences used for conservation analysis are available in Table 4.4. (C and D) HMM logo representation of the DxxTxxxGN and FxxxD sequence motifs identified in Lap1 family members. Probability is determined as a percent likelihood based on the Lap1 protein sequences in Table 4.4. (E and F) Mutation of the FxxxD motif in LapD1 blocks TelD secretion but does not impact TelD$_{LXG}$-LapD1-LapD2 complex formation. Western blot analysis of the secreted and cell fractions of the indicated *S. intermedius* GC1825 strains (E). SYPRO Red staining of purified TelD$_{LXG}$-LapD1-LapD2 wild-type and indicated variant complexes (F). (G) The FxxxD motif of Lap1 proteins is predicted to exist in a similar three-dimensional position as the YxxxE secretion signal of the T7SSa effector PE25. Noodle representation of LapC1 and PE25 (PDB ID: 4W4L) superposition. Structural models were aligned using the default matchmaker algorithm in

ChimeraX. Asterisk indicates the approximate position of E91 as it was not modelled in the PE25 structure.

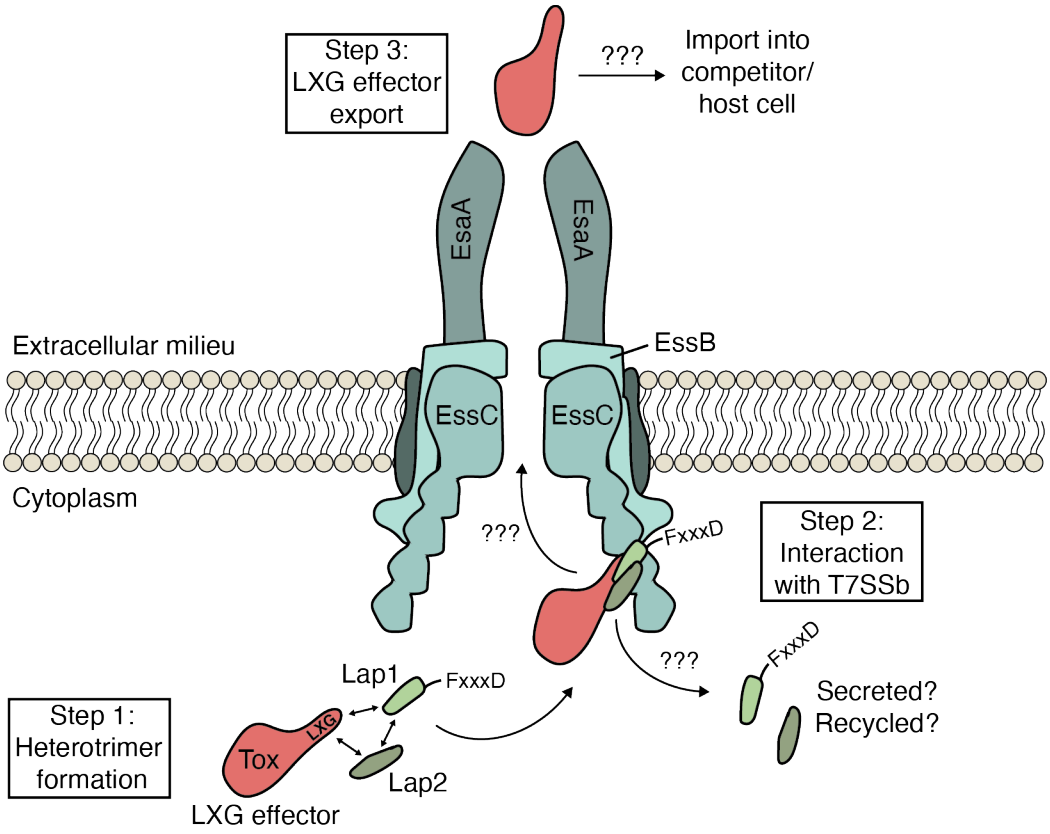


Figure 4.6: Model depicting LXG effector recruitment to the T7SSb apparatus by Lap1 and Lap2 targeting factors. Based on the findings described in this work, we propose that LXG effectors form a pre-secretion complex with cognate Lap1/DUF3130 and Lap2/DUF3958 proteins (step 1). The quaternary structure of this complex, in conjunction with the FxxxD motif found in Lap1 proteins, likely acts as a signal sequence that recruits LXG effectors to the T7SSb apparatus (step 2). The details of how T7SSb apparatuses facilitate protein export across the plasma membrane remain unknown but based on the findings of Rosenberg *et al.* on the ESX-1 T7SSa, this may involve effector-induced multimerization of EssC (step 3) (Rosenberg et al., 2015). Once LXG effectors are released from the bacterial cell, those with cytotoxic activity enter the cytoplasm of their target cell by an unknown molecular mechanism.

Tables

Table 4.1. Spectral counts for TelC-V and LapC1-V immunoprecipitated samples and their respective control samples

#	Identified Proteins	Accession Number	Alternate ID	Δ*telC* ctrl	Δ*telC* + *telC*-V	Δ*wxgC* ctrl	Δ*wxgC* +*wxgC* -V
1	TelC-VSV-G tagged	TelC_VSV-G		0	716	11	191
2	WxgC-VSV-G tagged	WxgC_VSV-G		0	288	2	184
3	Enolase	T1ZFF4	eno	35	45	28	47
4	Polyribonucleotide nucleotidyltransferase	T1ZFD6	pnpA	39	37	35	36
5	Uncharacterized protein	T1ZGI6	SIR_1490	0	73	0	28
6	Oligopeptide-binding protein AmiA	T1ZFZ2	amiA	21	25	12	15
7	60 kDa chaperonin	T1ZGB6	groL	20	20	8	15
8	ABC-type transport system, periplasmic binding protein	T1ZEB9	SIR_1223	13	16	9	11
9	Uncharacterized protein	T1ZDT0	SIR_1033	13	10	6	10
10	Putative extracellular solute-binding protein	T1ZFL7	SIR_1387	11	11	7	11
11	Chaperone protein DnaJ	T1ZG02	dnaJ	10	9	12	13
12	ABC transporter, substrate-binding protein	T1ZG31	SIR_1454	9	13	5	10
13	Isoprenyl transferase	T1ZGQ4	uppS	5	13	7	8
14	Uncharacterized protein	T1ZGR7	SIR_1322	7	13	2	5
15	Uncharacterized protein	T1ZEQ3	SIR_0983	13	7	4	5
16	Foldase protein PrsA	T1ZG93	prsA	8	7	4	7

17	30S ribosomal protein S2	T1ZC88	rpsB	2	15	5	3
18	Aminopeptidase	T1ZEN3	pepC	0	21	0	6
19	Pullulanase, type I	T1ZEI5	pulA	11	5	7	5
20	Elongation factor Tu	T1ZEN1	tuf	4	9	4	5
21	Protein RecA	T1ZFX5	recA	6	15	0	4
22	Ribosomal RNA small subunit methyltransferase H	T1ZGX5	mraW	6	13	5	3
23	Putative rhamnosyltransferase RgpA	T1ZEF3	rgpA	0	9	7	7
24	30S ribosomal protein S5	T1ZGK0	rpsE	4	8	5	7
25	Oxidoreductase	T1ZD40	SIR_0796	5	4	8	2
26	Translation initiation factor IF-3	T1ZDK2	infC	4	8	0	4
27	Uracil phosphoribosyltransferase	T1ZGI0	upp	3	9	2	4
28	Putative lipoprotein	T1ZEA9	SIR_0850	4	6	6	2
29	DNA-directed RNA polymerase subunit beta'	T1ZFL9	rpoC	4	0	5	5
30	Response regulator	T1ZH01	comE	4	11	2	0
31	Beta-N-acetylhexosaminidase	T1ZED9	lacZ	4	5	2	2
32	Beta-N-acetylhexosaminidase	T1ZED9-DECOY		6	0	4	2
33	L-lactate dehydrogenase	T1ZEP5	ldh	3	7	4	4
34	50S ribosomal protein L6	T1ZGX1	rplF	2	9	0	0

35	30S ribosomal protein S12	T1ZCJ3	rpsL	4	3	4	4
36	50S ribosomal protein L4	T1ZFT4	rplD	3	2	2	7
37	DNA-binding protein HU	T1ZCZ7	SIR_0424	6	2	3	6
38	Uncharacterized protein	T1ZGF5	SIR_1455	2	11	0	0
39	Biotin carboxylase	T1ZEV4	accC	3	7	2	3
40	Uncharacterized protein	T1ZBA4	SIR_0113	4	5	4	0
41	Uncharacterized protein	T1ZEG5	SIR_1274	4	6	0	3
42	Signal recognition particle protein	T1ZE30	ffh	5	7	0	0
43	Uncharacterized protein	T1ZGC6	SIR_1156	2	5	3	5
44	Mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase	T1ZEU1	SIR_1072	3	6	2	3
45	DNA-directed RNA polymerase subunit beta	T1ZGS7	rpoB	2	5	4	2
46	Surface antigen	T1ZCQ5	SIR_0054	2	6	5	2
47	Putative adhesion protein	T1ZDS8	fszD	5	5	0	2
48	Hyaluronate lyase	T1ZG27	SIR_1547	3	6	3	2
49	Putative collagen adhesin	T1ZHC4	SIR_1805	3	3	5	2
50	C5a peptidase	T1ZFN2	SIR_1402	3	5	2	2
51	Chaperone protein DnaK	T1ZF47	dnaK	2	5	4	0
52	Putative cell-surface antigen I/II	T1ZHQ3	SIR_1675	0	2	4	0

53	Putative glycosyl transferase	T1ZFQ8	SIR_0933	4	5	0	2
54	Uncharacterized protein	T1ZFV5	SIR_1477	4	3	0	3
55	Formate acetyltransferase	T1ZD63	pfl	2	5	0	3
56	Lysozyme	T1ZF98	SIR_1025	7	2	0	2
57	Translation initiation factor IF-2	T1ZGZ1	infB	0	3	5	2
58	Putative alkaline amylopullulanase	T1ZGL9	pulA2	0	5	4	4
59	Putative stress protein	T1ZDD5	SIR_0040	0	3	3	0
60	3-oxoacyl-[acyl-carrier-protein] synthase 2	T1ZH25	fabF	2	7	0	2
61	DNA polymerase III PolC-type	T1ZGA0	polC	3	5	2	0
62	Uncharacterized protein	T1ZGR7-DECOY		2	0	4	0
63	Glutamine synthetase I alpha	T1ZGF6	glnA	0	4	2	5
64	ATP-dependent zinc metalloprotease FtsH	T1ZDD0	ftsH	3	2	0	4
65	LysM domain-containing protein	T1ZHJ6	SIR_1880	4	4	0	0
66	Uncharacterized protein	T1ZEQ3-DECOY		0	0	3	0
67	Putative phosphoribosylformylglycinamidine synthase	T1ZB52	purL	0	7	2	0
68	Glyceraldehyde-3-phosphate dehydrogenase	T1ZCF7	gap	0	7	0	0
69	Pyruvate formate lyase	T1ZF15	SIR_1079	2	5	0	0

70	Transcription-repair-coupling factor	T1ZC42	trcF	0	0	0	3
71	Chromosome partition protein Smc	T1ZDG1	smc	0	4	2	0
72	Cell division ATP-binding protein FtsE	T1ZF37	ftsE	0	2	2	5
73	50S ribosomal protein L18	T1ZFR9	rplR	2	3	0	0
74	Alkyl hydroperoxide reductase subunit F	T1ZGT3	ahpF	2	4	0	0
75	Uncharacterized protein	T1ZCV9-DECOY		0	0	4	0
76	Valine--tRNA ligase	T1ZFZ4	valS	0	4	2	3
77	Phosphoglycerate kinase	T1ZGG2	pgk	0	2	0	0
78	Beta-N-acetylhexosaminidase	T1ZET7	SIR_1067	2	3	2	0
79	30S ribosomal protein S10	T1ZH81	rpsJ	2	3	0	3
80	GRAM_POS_ANCHORING domain-containing protein	T1ZEJ8	SIR_0758	3	0	0	2
81	Elongation factor G	T1ZDS4	fusA	2	3	0	3
82	Lysine--tRNA ligase	T1ZFC7	lysS	3	2	0	0
83	Putative recombinase	T1ZEH5	SIR_0971	4	2	0	2
84	ABC transporter, substrate-binding protein	T1ZD17	msmE	0	3	0	0
85	Putative conjugal transfer protein	T1ZFW8	SIR_0990	2	6	0	3
86	Pyruvate kinase	T1ZEN4	pyk	0	8	0	0
87	Uncharacterized protein	T1ZCC8	SIR_0176	2	3	0	0

88	DD-transpeptidase	T1ZBB4	SIR_0124	3	3	2	0
89	DUF4366 domain-containing protein	T1ZDN0	SIR_0987	0	2	3	0
90	Type I restriction enzyme R Protein	T1ZDQ2	hsdR	3	4	2	0
91	Putative DNA-entry endonuclease	T1ZDT4	endA	2	3	0	0
92	Putative penicillin binding protein 2B	T1ZDK4	pbp2b	4	0	0	3
93	Threonine--tRNA ligase	T1ZFI3	thrS	3	3	0	0
94	DUF4832 domain-containing protein	T1ZGK3	SIR_1591	0	4	4	0
95	Conjugal transfer protein	T1ZEL6	SIR_1329	4	0	0	0
96	Uncharacterized protein	T1ZDP3	SIR_0613	0	3	6	0
97	Phosphoenolpyruvate-protein phosphotransferase	T1ZG24	ptsI	2	2	0	2
98	Isopentenyl-diphosphate delta-isomerase	T1ZD59	fni	2	5	0	0
99	Peptidyl-prolyl cis-trans isomerase	T1ZE70	ppiA	0	8	0	0
100	Histidine triad protein	T1ZEB0	SIR_0654	0	0	2	2

Table 4.2. X-ray data collection and refinement statistics for LapD2

	LapD2 (selenomethionine)	LapD2 (native)
Data Collection		
Wavelength (Å)	0.9793	0.9793
Space group	$P3_121$	$P3_1$
Cell dimensions		
a, b, c (Å)	45.2, 45.2, 298.4	45.6, 45.6, 298.4
Resolution[a] (Å)	39.20 – 2.42 (2.46 – 2.42)	39.46 – 2.20 (2.24 – 2.20)
Unique reflections	14603 (727)	34956 (1569)
$CC_{1/2}$ [c]	0.999 (0.371)	0.999 (0.474)
R_{merge}[b]	0.128 (4.215)	0.078 (2.118)
R_{pim}[c]	0.036 (1.110)	0.035 (0.840)
$I/\sigma I$	13.0 (0.6)	12.9 (0.8)
Completeness (%)	100 (100)	99.3 (90.6)
Redundancy	21.7 (15.0)	11.1 (6.6)
Refinement		
R_{work}[d]/ R_{free} (%)	-	23.3/26.7
Average B-factors (Å^2)		76.2
Protein	-	76.3
Water/Other	-	56.1/88.5
No. atoms		
Protein	-	3758
Water/Other	-	35/20
Rms deviations		
Bond lengths (Å)	-	0.005
Bond angles (°)	-	0.733
Ramachandran plot (%)		
Total favored[e]	-	98.38
Total allowed	-	1.39
PDB code	-	7UH4

[a]Values in parentheses correspond to the highest resolution shell. [b]$R_{merge} = \Sigma_h\Sigma_j|I_{hj} - \langle I_h \rangle|/\Sigma_h\Sigma_j I_{hj}$, where I_{hj} is the intensity of observation j of reflection h. [c]As defined by Karplus and Diederichs (2012). [d]$R = \Sigma_h||F_o| - |F_c||/\Sigma_h|F_o|$ for all reflections, where F_o and F_c are observed and calculated structure factors, respectively. R_{free} is calculated analogously for the test reflections, randomly selected and excluded from the refinement. [e]As defined by Molprobity (Davis et al., 2004).

Table 4.3. Accession codes and sequence information for LapD2 homologs identified with three iterations of JackHMMER.

Entry	Protein names	Gene names	Organism	Length
F0ISI0	Uncharacterized protein	HMPREF9384_0792	Streptococcus sanguinis SK160	120
A0A829IC82	Uncharacterized protein	SAG0014_09635	Streptococcus agalactiae FSL S3-586	120
A0A427Z096	Uncharacterized protein	D8894_04900	Streptococcus oralis	120
F8DHG2	Uncharacterized protein	HMPREF0833_11762	Streptococcus parasanguinis ATCC 15912	118
A0A8B1YUD9	DUF3958 family protein	J4854_01605	Streptococcus lactarius	118
A0A178KGP4	Uncharacterized protein	A3Q39_01935	Streptococcus sp. CCUG 49591	118
A0A1X1IMY3	Uncharacterized protein	B7710_01130	Streptococcus oralis subsp. oralis	120
A0A3R9HBG1	Uncharacterized protein	D8875_04300	Streptococcus sanguinis	120
A3CR32	Uncharacterized protein	SSA_2275	Streptococcus sanguinis (strain SK36)	121
A0A178KI83	Uncharacterized protein	A3Q39_01965	Streptococcus sp. CCUG 49591	124
S7XHS7	Uncharacterized protein	M059_05495	Streptococcus mitis 18/56	124
A0A139P9I6	Uncharacterized protein	SORDD16_01672	Streptococcus oralis	121
A0A3R9JF83	Uncharacterized protein	D8839_01325	Streptococcus mitis	118

A0A428A3Y3	Uncharacterized protein	D8883_04735	Streptococcus sanguinis	120
A0A427ZT62	Uncharacterized protein	D8886_05325	Streptococcus sanguinis	120
A0A5A7ZT25	Uncharacterized protein	FKX92_00600	Streptococcus sanguinis	129
A0A7H8V963	Uncharacterized protein	FFV08_11455	Streptococcus sanguinis	120
A0A123VUG4	FKBP_N domain-containing protein	ERS132372_01528 ERS132399_02391	Streptococcus suis	128
A0A428A688	Uncharacterized protein	D8879_11740	Streptococcus sanguinis	120
A0A1F0ZSH8	Uncharacterized protein	HMPREF2917_09360	Streptococcus sp. HMSC061E03	118
F3UNP6	Uncharacterized protein	HMPREF9389_0454	Streptococcus sanguinis SK355	121
A0A1X1JWY6	Uncharacterized protein	B7700_09660	Streptococcus mitis	118
A0A345VJJ3	Uncharacterized protein	Sp14A_09740	Streptococcus pluranimalium	129
A0A8B4IQ53	Uncharacterized protein	NCTC3858_00393	Streptococcus uberis	122
A0A0F5MM48	Uncharacterized protein	RN86_02675	Streptococcus gordonii	132
A0A0F2CF76	Uncharacterized protein	TZ86_01640 UA00_00089	Streptococcus gordonii	119
A0A2X3XZG6	Uncharacterized protein	NCTC12278_01112	Streptococcus ferus	131
A0A0E1EH98	Uncharacterized protein	AX245_04160 C4618_11680 C6N07_05900 RDF_1029	Streptococcus agalactiae	118
A0A4T2H8W2	Uncharacterized protein	FAJ36_02910	Streptococcus suis	123

A0A1V0H1D1	Uncharacterized protein	A6J85_03500	Streptococcus gordonii	118
A0A7H8UYG8	Energy transducer TonB	FDP16_01525	Streptococcus sanguinis	120
A0A1E5GHA5	Uncharacterized protein	BCR21_07310	Enterococcus ureasiticus	126
A0A1X1J4E5	Uncharacterized protein	B7708_00960	Streptococcus oralis subsp. dentisani	124
A0A7Z0VFP3	Uncharacterized protein	TH70_0121	Streptococcus agalactiae	123
A0A1E5HGJ1	Uncharacterized protein	BCR24_01620	Enterococcus ureilyticus	118
A0A4P7WQS8	Uncharacterized protein	E8M06_09955 E8M06_09985	Streptococcus suis	123
A0A0U2NRK3	Uncharacterized protein	ATZ35_10685	Enterococcus rotai	118
E6KIR2	Uncharacterized protein	HMPREF8578_0127	Streptococcus oralis ATCC 49296	120
A0A4R5G734	Uncharacterized protein	E0E04_02155	Streptococcus vicugnae	134
A0A6I3PB65	Uncharacterized protein	GMC80_04755 GMC84_06710	Streptococcus parasanguinis	118
A0A7X2UEL6	Uncharacterized protein	NCTC3858_01463	Streptococcus uberis	126
E6KIP9	Uncharacterized protein	HMPREF8578_0114	Streptococcus oralis ATCC 49296	118
A0A540UNN3	FKBP_N domain-containing protein	FH692_10965	Streptococcus suis	128
A0A3L8GE13	Uncharacterized protein	DIY07_08810	Streptococcus iniae	125

			(Streptococcus shiloi)	
A0A7H9FG12	Uncharacterized protein	HRE59_00315	Streptococcus oralis subsp. oralis	118
A0A372KJ05	Uncharacterized protein	DDV21_010945 DDV23_10765	Streptococcus chenjunshii	131
A0A0J6KU02	Uncharacterized protein	VK90_24155	Bacillus sp. LK2	116
A0A427Z4E3	Energy transducer TonB	D8889_08515 FKX92_06260	Streptococcus sanguinis	120
A0A0F5MJX1	Uncharacterized protein	RN86_02700	Streptococcus gordonii	118
A0A7X2UQ75	Uncharacterized protein	NCTC3858_01475	Streptococcus uberis	126
A0A0S3K6Z3	Uncharacterized protein	ATZ33_01285	Enterococcus silesiacus	118
A0A242AUF8	Uncharacterized protein	A5821_000622	Enterococcus sp. 7F3_DIV0205	120
A0A242H4J5	Uncharacterized protein	A5866_002132	Enterococcus sp. 12C11_DIV0727	118
A0A242CWU2	Uncharacterized protein	A5875_003888	Enterococcus sp. 3H8_DIV0648	119
F0IN33	HD domain protein	HMPREF9383_1536	Streptococcus sanguinis SK150	119
A0A427ZN60	Uncharacterized protein	D8886_09175	Streptococcus sanguinis	120
A0A081QRU4	Cell-cycle control medial ring component family protein	SK578_0511	Streptococcus mitis	124
A0A242ATT0	Uncharacterized protein	A5821_000410	Enterococcus sp. 7F3_DIV0205	119
A0A3R9J4D5	Uncharacterized protein	D8860_09785	Streptococcus oralis	118

A0A2X3VDB5	Uncharacterized protein	NCTC11085_00303	Streptococcus sanguinis	120
A0A1X1HW15	Uncharacterized protein	B7714_09145	Streptococcus oralis subsp. oralis	120
A0A0Z8JBB1	Uncharacterized protein	ERS132440_00897	Streptococcus suis	123
A0A242GZP3	Uncharacterized protein	A5866_000650	Enterococcus sp. 12C11_DIV0727	114
A0A841YH39	DUF3958 family protein	HB844_13135	Listeria fleischmannii	118
A0A1E5GX96	Uncharacterized protein	BCR23_04630	Enterococcus quebecensis	115
R2T5H1	Uncharacterized protein	UAY_00975	Enterococcus moraviensis ATCC BAA-383	118
A0A7H8V9W6	Uncharacterized protein	FFV08_11490	Streptococcus sanguinis	120
F0FHF6	Uncharacterized protein	HMPREF9388_2139	Streptococcus sanguinis SK353	121
A0A1X1IPR0	Uncharacterized protein	B7710_00060	Streptococcus oralis subsp. oralis	118
A0A428G5R6	Uncharacterized protein	D8801_04900	Streptococcus oralis	124
A0A0N0KTL2	Uncharacterized protein	AEQ18_02380	Enterococcus sp. RIT-PI-f	116
A0A200JBQ9	Uncharacterized protein	A5889_000138	Enterococcus sp. 9D6_DIV0238	117
A0A7D4GRI0	Uncharacterized protein	FOC63_06870	Streptococcus gallolyticus	134
A0A4T2GM54	Uncharacterized protein	FAJ39_07710	Streptococcus suis	128
A0A242LA88	Uncharacterized protein	A5881_003618	Enterococcus termitis	118

A0A380IM03	Uncharacterized protein	NCTC6175_01411	Streptococcus agalactiae	120
A0A4V6U7E4	Uncharacterized protein	FAJ36_02880	Streptococcus suis	128
A0A3R9HGP9	Uncharacterized protein	D8887_07705	Streptococcus sanguinis	113
A0A7Z7QUJ7	Uncharacterized protein	NCTC8183_01312	Streptococcus agalactiae	133
A0A139NND7	Uncharacterized protein	STRDD11_02626	Streptococcus sp. DD11	120
A0A2L0D3F4	Uncharacterized protein	C0J00_04050	Streptococcus pluranimalium	131
R2T9D5	Uncharacterized protein	UAY_02590	Enterococcus moraviensis ATCC BAA-383	117
A0A0Z8HRE2	Uncharacterized protein	ERS132406_02094	Streptococcus suis	123
A0A0B7GNC7	Uncharacterized protein	SSV_1920	Streptococcus sanguinis	120
F9LWN3	Uncharacterized protein	HMPREF9965_0736	Streptococcus mitis bv. 2 str. SK95	118
A0A1X1JX79	Uncharacterized protein	B7700_09690	Streptococcus mitis	124
A0A1E5H5L3	Uncharacterized protein	BCR24_09885	Enterococcus ureilyticus	122
F0IBB8	Uncharacterized protein	HMPREF9382_2056	Streptococcus sanguinis SK115	120
R3W643	Uncharacterized protein	UC3_02024	Enterococcus phoeniculicola ATCC BAA-412	116
A0A428IHC8	Uncharacterized protein	D8844_06490	Streptococcus oralis	120
A0A2W4BKR6	Uncharacterized protein	CI088_09485	Enterococcus plantarum	115

A0A3R9H620	Uncharacterized protein	D8879_10595	Streptococcus sanguinis	120
A0A1E5GJU3	Uncharacterized protein	BCR25_08215	Enterococcus termitis	115
A0A0Z8I4W5	Uncharacterized protein	ERS132410_02192	Streptococcus suis	123

Table 4.4. Accession codes and sequence information for LapC1 homologs identified with one iteration of JackHMMER.

Entry	Protein names	Gene names	Organism	Length
T1ZH75	Uncharacterized protein	SIR_1491	Streptococcus intermedius B196	91
A0A0E2IQB7	Uncharacterized protein	HMPREF1654_01870	Streptococcus intermedius ATCC 27335	91
A0A139R5L5	TIGR04197 family type VII secretion effector	FOC63_00900 SGADD02_00470 SGADD03_00389	Streptococcus gallolyticus	93
A0A1S5WDW5	Uncharacterized protein	BTR42_08900	Streptococcus gallolyticus subsp. gallolyticus DSM 16831	93
A0A1I7GQI7	Type VII secretion effector, SACOL2603 family	SAMN05660328_102271	Streptococcus gallolyticus	93
F5WVX6	Uncharacterized protein	SGGB_1575	Streptococcus gallolyticus ATCC 43143	93
E8K2Z3	Uncharacterized protein	HMPREF9423_1856	Streptococcus infantis ATCC 700779	92
A0A1H8Z4E7	Type VII secretion effector, SACOL2603 family	SAMN05216346_101162	Streptococcus equinus (Streptococcus bovis)	90
A0A139QYV5	Uncharacterized protein	SGADD02_00817 SGADD03_01202	Streptococcus gallolyticus	90
F9LY30	Uncharacterized protein	HMPREF9965_1762	Streptococcus mitis bv. 2 str. SK95	92
A0A1C3SMV2	Uncharacterized protein	SMA679_0761	Streptococcus macedonicus	90

A0A3R9HJH6	Uncharacterized protein	D8863_08620	Streptococcus oralis	92
A0A1H0MTA7	Type VII secretion effector, SACOL2603 family	SAMN05216347_102469	Streptococcus equinus (Streptococcus bovis)	90
A0A371QFB0	TIGR04197 family type VII secretion effector	DXN33_01140	Streptococcus sp. NM	92
A0A3R9QBN4	Uncharacterized protein	D8786_05750 D8855_04310	Streptococcus mitis	92
A0A1F0BUA5	Type VII secretion protein	HMPREF2613_07245	Streptococcus sp. HMSC070B10	92
A0A501PB50	TIGR04197 family type VII secretion effector	FJN11_06485	Streptococcus symci	92
A0A3R9HQE0	TIGR04197 family type VII secretion effector	D8789_07065 D8849_09150 D8865_10365 JJN14_03035	Streptococcus mitis	92
A0A1E9GAV6	Type VII secretion protein	HMPREF2766_03755	Streptococcus sp. HMSC076C08	92
A0A2G3NUY4	TIGR04197 family type VII secretion effector	CS009_05415 CS010_03220	Streptococcus macedonicus	90
A0A7D4GS34	TIGR04197 family type VII secretion effector	FOC63_08560	Streptococcus gallolyticus	90
A0A1S5WBI2	Uncharacterized protein	BTR42_04595	Streptococcus gallolyticus subsp. gallolyticus DSM 16831	90

A0A1B1ID96	Type VII secretion protein	AXF18_01730	Streptococcus sp. oral taxon 064	92
A0A2I1UMC7	TIGR04197 family type VII secretion effector	CYK17_09995	Streptococcus oralis subsp. dentisani	92
A0A1S0ZA19	Type VII secretion protein	A7T00_33115	Salmonella enterica subsp. enterica serovar Saintpaul	92
A0A380K862	Type VII secretion effector	NCTC13767_01892	Streptococcus gallolyticus	90
A0A1H6SD36	Type VII secretion effector, SACOL2603 family	SAMN05216460_1192	Streptococcus sp. 45	90
A0A3R9FX19	Uncharacterized protein	D8894_04980	Streptococcus oralis	92
A0A1X1J9V7	Type VII secretion effector	B7705_06215	Streptococcus oralis subsp. dentisani	92
A0A428DJZ0	Uncharacterized protein	D8847_09950	Streptococcus mitis	92
A0A3R9J234	Uncharacterized protein	D8847_09775	Streptococcus mitis	92
F5X0A7	Uncharacterized protein	SGGB_0839	Streptococcus gallolyticus ATCC 43143	90
A0A139PV09	Uncharacterized protein	SORDD27_01490	Streptococcus oralis	92
I0Q5G0	Type VII secretion effector, TIGR04197 family	HMPREF1115_1692	Streptococcus oralis SK610	92

A0A1I7FJ84	Type VII secretion effector, SACOL2603 family	SAMN05660328_101420	Streptococcus gallolyticus	90
A0A239RBG6	Type VII secretion effector, SACOL2603 family	SAMN05216470_0920	Streptococcus equinus (Streptococcus bovis)	90
A0A081QNZ9	Uncharacterized protein	SK578_0768 SMIM3I_00648 SMIM3IV_00595	Streptococcus mitis	92
A0A231VWK6	TIGR04197 family type VII secretion effector	CBI42_08510	Streptococcus sp. KR	92
A0A1F0B683	Type VII secretion protein	HMPREF2701_04775	Streptococcus sp. HMSC077D04	92
A0A4V0BUI7	Type VII secretion effector	NCTC5338_01391	Streptococcus australis	92
A0A4V6LQ02	Type VII secretion effector	NCTC10232_01364	Streptococcus oralis	92
A0A2X3W4X4	Type VII secretion effector	NCTC12278_01169	Streptococcus ferus	91
A0A1X1INN9	Type VII secretion effector	B7710_01210	Streptococcus oralis subsp. oralis	92
A0A3R9KT57	Uncharacterized protein	D8788_09675	Streptococcus mitis	92
J5H474	Type VII secretion effector, TIGR04197 family	HMPREF1125_0309	Streptococcus oralis SK304	92

A0A1S1CRP1	Type VII secretion protein	HMPREF2628_07975	Streptococcus sp. HMSC063B03	92
A0A139QMH4	Uncharacterized protein	SORDD24_01549	Streptococcus oralis	92
A0A1X1H983	Type VII secretion effector	B7721_02930	Streptococcus oralis subsp. oralis	92
A0A428IP91	Uncharacterized protein	D8846_06225	Streptococcus oralis	92
A0A1X1HPW8	Type VII secretion effector	B7716_01660	Streptococcus oralis subsp. oralis	92
E9FIV1	Uncharacterized protein	HMPREF0849_01627	Streptococcus sp. C300	92
A0A139Q4A6	Type VII secretion protein	BBP19_06505 SORDD30_01629	Streptococcus oralis	92
A0A1X1GSZ3	Type VII secretion effector	B7712_00855	Streptococcus oralis subsp. oralis	92
A0A139M8G1	Uncharacterized protein	SORDD05_01233	Streptococcus oralis	92
A0A139QLX9	Uncharacterized protein	SORDD24_01677	Streptococcus oralis	92
A0A139PVN7	Uncharacterized protein	D8844_06410 SORDD20_00506	Streptococcus oralis	92
A0A1X1HNL4	Type VII secretion effector	B7718_02130	Streptococcus oralis subsp. oralis	92
A0A428HB07	Uncharacterized protein	D8788_03670	Streptococcus mitis	92
G6C8X2	Uncharacterized protein	HMPREF9184_00751	Streptococcus sp. oral taxon 058 str. F0407	92
A0A4Q2FKS1	TIGR04197 family type VII	DF216_07805	Streptococcus oralis	92

	secretion effector			
J5GN34	Type VII secretion effector, TIGR04197 family	HMPREF1125_2061	Streptococcus oralis SK304	92
A0A1X0X0B5	Type VII secretion protein	ATE37_07430	Streptococcus oralis subsp. tigurinus	92
A0A428CAR2	Uncharacterized protein	D8856_09625	Streptococcus mitis	92
A0A3R9KGB9	Uncharacterized protein	D8854_03060	Streptococcus mitis	92
A0A1X1I2K2	Type VII secretion effector	B7714_02825	Streptococcus oralis subsp. oralis	92
A0A139NUT4	Uncharacterized protein	SORDD14_01568	Streptococcus oralis	92
A0A4Q2FL97	TIGR04197 family type VII secretion effector	DF216_07290	Streptococcus oralis	92
A0A1X1H062	Type VII secretion effector	B7722_01935	Streptococcus oralis subsp. oralis	92
A0A4R5G4Y4	TIGR04197 family type VII secretion effector	E0E04_04080	Streptococcus vicugnae	91
A0A1X1GBF8	Type VII secretion effector	B7727_03280	Streptococcus oralis subsp. tigurinus	93
A0A139NWT6	Uncharacterized protein	SORDD15_01377	Streptococcus oralis	92
A0A135YLC9	Uncharacterized protein	HMPREF3205_02308	Streptococcus pasteurianus	96

A0A1S5WCM9	Uncharacterized protein	BTR42_05645	Streptococcus gallolyticus subsp. gallolyticus DSM 16831	91
A0A7D4GHM0	TIGR04197 family type VII secretion effector	FOC63_09620	Streptococcus gallolyticus	91
A0A7D4K0Q8	TIGR04197 family type VII secretion effector	FOC63_07845	Streptococcus gallolyticus	91
A0A1I7F6C6	Type VII secretion effector, SACOL2603 family	SAMN05660328_101220	Streptococcus gallolyticus	91
A0A1I7FC93	Type VII secretion effector, SACOL2603 family	SAMN05660328_101314	Streptococcus gallolyticus	91
A0A139NQ79	Uncharacterized protein	STRDD11_02464	Streptococcus sp. DD11	89
F3USM5	Uncharacterized protein	HMPREF9389_1833	Streptococcus sanguinis SK355	90
A0A3R9IAM7	TIGR04197 family type VII secretion effector	D8887_08455 FFV08_05580	Streptococcus sanguinis	90
A0A427ZP46	Uncharacterized protein	D8886_07895	Streptococcus sanguinis	90
A0A3R9NTY4	Uncharacterized protein	D8879_08845	Streptococcus sanguinis	90
G5JR52	Uncharacterized protein	STRCR_1677 STRCR_1937	Streptococcus criceti HS-6	92
A0A2A5SDM4	TIGR04197 family type VII	FEZ46_05180 RU88_GL002128	Lactococcus raffinolactis	102

	secretion effector			
A0A0F3H405	Uncharacterized protein	TZ97_00642	Streptococcus parasanguinis	89
A0A6N3CT23	Uncharacterized protein	SPLFYP13_01158	Streptococcus parasanguinis	89
F8DGG8	Uncharacterized protein	HMPREF0833_10386	Streptococcus parasanguinis ATCC 15912	94
A0A359YGK2	Uncharacterized protein	SPADD19_01110	Streptococcus parasanguinis	89
A0A1F1A3X5	Uncharacterized protein	HMPREF2917_04405	Streptococcus sp. HMSC061E03	89
I1ZLJ5	Uncharacterized protein	Spaf_0919	Streptococcus parasanguinis FW213	103
A0A4Q5BT34	TIGR04197 family type VII secretion effector	GMC84_09185 GMC94_02205	Streptococcus parasanguinis	89
I2NMG3	Uncharacterized protein	HMPREF9971_1232	Streptococcus parasanguinis F0449	113
G5JRP1	Uncharacterized protein	STRCR_2050	Streptococcus criceti HS-6	91
A0A1F0AWW4	Uncharacterized protein	HMPREF2686_08175	Streptococcus sp. HMSC057G03	89
V8BGZ4	Uncharacterized protein	HMPREF1195_00404	Streptococcus parasanguinis CC87K	89
A0A428B5A9	Uncharacterized protein	D8866_01720	Streptococcus parasanguinis	89
A0A4Q2FH31	TIGR04197 family type VII secretion effector	DF218_03565	Streptococcus parasanguinis	89
A0A6I3PR01	TIGR04197 family type VII	GMC95_02245	Streptococcus parasanguinis	94

	secretion effector			
E8K4F1	Uncharacterized protein	HMPREF8577_0436	Streptococcus parasanguinis ATCC 903	99
A0A6A0B9J2	Type VII secretion protein	Hs30E_00170	Lactococcus hodotermopsidis	101
E3CE17	Uncharacterized protein	HMPREF9626_1164	Streptococcus parasanguinis F0405	89
A0A4R5G605	TIGR04197 family type VII secretion effector	E0E04_02150	Streptococcus vicugnae	118
A0A7D4GGP0	TIGR04197 family type VII secretion effector	FOC63_06865	Streptococcus gallolyticus	118
A0A139MU55	Uncharacterized protein	STRDD04_00268	Streptococcus sp. DD04	97
R2R504	Type VII secretion effector	UAI_02685	Enterococcus malodoratus ATCC 43197	92
A0A8B1YTF9	TIGR04197 family type VII secretion effector	J4854_05255	Streptococcus lactarius	89
A0A224XFQ7	Uncharacterized protein	RsY01_1995	Lactococcus reticulitermitis	102
A0A0A0DFU7	Uncharacterized protein	SSIN_0557	Streptococcus sinensis	104
A0A242DHL2	Uncharacterized protein	A5875_002996	Enterococcus sp. 3H8_DIV0648	92
A0A7W1YHC9	TIGR04197 family type VII secretion effector	HPK16_15390	Listeria rustica	91

A0A378MC82	Type VII secretion effector	NCTC10815_01240	Listeria grayi (Listeria murrayi)	92
A0A0S3K6Z8	Uncharacterized protein	ATZ33_01280	Enterococcus silesiacus	95
A0A0U2XFK1	Uncharacterized protein	ATZ35_10680	Enterococcus rotai	95
A0A242H2N4	Uncharacterized protein	A5866_002133	Enterococcus sp. 12C11_DIV0727	95
A0A1E5KVA8	Uncharacterized protein	BCR26_15430	Enterococcus rivorum	93
R2TRA5	Type VII secretion effector	UAY_00974	Enterococcus moraviensis ATCC BAA-383	95
D7V0H1	Uncharacterized protein	HMPREF0556_11749	Listeria grayi DSM 20601	95
K8N1C5	Uncharacterized protein	HMPREF9186_00129	Streptococcus sp. F0442	89
A0A242L9F8	Uncharacterized protein	A5881_003619	Enterococcus termitis	95
A0A1E5HGJ2	Uncharacterized protein	BCR24_01625	Enterococcus ureilyticus	95
A0A2R7ZZP2	Uncharacterized protein	CDIMF43_180250 CKN86_07930	Carnobacterium divergens (Lactobacillus divergens)	92
A0A830LAN8	TIGR04197 family type VII secretion effector	CW834_00955	Listeria monocytogenes	97
A0A242ATT5	Uncharacterized protein	A5821_000409	Enterococcus sp. 7F3_DIV0205	95
A0A242CX89	Uncharacterized protein	A5875_003889	Enterococcus sp. 3H8_DIV0648	95
A0A842EF80	TIGR04197 family type VII	HB895_12440 HCB08_04225	Listeria booriae	96

	secretion effector	HCB25_04225 HCB35_09535		
A0A7X0XEW8	TIGR04197 family type VII secretion effector	HCI99_14105 HCJ13_00955	Listeria booriae	96
A0A5E9H6J9	Type VII secretion effector	NCTC13772_01143 NCTC13772_02346	Carnobacterium divergens (Lactobacillus divergens)	92
A0A0J6L2B8	Type VII secretion effector	VK90_21625	Bacillus sp. LK2	99
A0A3R9G5C1	Uncharacterized protein	D8887_07710	Streptococcus sanguinis	92
A0A5E9H653	Type VII secretion effector	NCTC13772_02372	Carnobacterium divergens (Lactobacillus divergens)	92
A0A7X0WR26	TIGR04197 family type VII secretion effector	HB856_08660 HCB51_16600	Listeria booriae	96
A0A081QQI2	Uncharacterized protein	D8845_00760 D8855_02220 D8865_04910 SK578_1302	Streptococcus mitis	102
A0A428IXG9	Uncharacterized protein	D8800_00795	Streptococcus oralis	102
R0P931	Uncharacterized protein	D065_00650	Streptococcus mitis 13/39	102
A0A0B7GL02	Putative type VII secretion effector	SSV_1220	Streptococcus sanguinis	97
A0A1E5KUF8	Uncharacterized protein	BCR26_04480	Enterococcus rivorum	93
A0A7Z8G2U5	Uncharacterized protein	CKN67_07395	Carnobacterium divergens	92

			(Lactobacillus divergens)	
A0A8B5GW87	Uncharacterized protein	CKN75_08770	Carnobacterium divergens (Lactobacillus divergens)	92
A0A0S3KD07	Uncharacterized protein	ATZ33_12420	Enterococcus silesiacus	117
R2QLU5	Type VII secretion effector	UAY_03088	Enterococcus moraviensis ATCC BAA-383	120
A0A7I0FCU7	Uncharacterized protein	CKN77_09500	Carnobacterium divergens (Lactobacillus divergens)	92
A0A242AQ39	Uncharacterized protein	A5821_003000	Enterococcus sp. 7F3_DIV0205	120
A0A7X9QZ50	TIGR04197 family type VII secretion effector	HF881_01535	Streptococcus sp. WB01_FAA12	102
W7C7M5	Uncharacterized protein	MFLO_05320	Listeria floridensis FSL S10-1187	96
A0A2R8A462	Uncharacterized protein	CDIMF43_50002 CKN69_02300 CKN86_04715	Carnobacterium divergens (Lactobacillus divergens)	92
A0A200JBQ5	Uncharacterized protein	A5889_000137	Enterococcus sp. 9D6_DIV0238	93
A0A6L6HCB9	TIGR04197 family type VII secretion effector	GIX45_16890	Erwinia sp. CPCC 100877	93
A0A0J6L0K0	Type VII secretion effector	VK90_24150	Bacillus sp. LK2	99
F0IBB7	Uncharacterized protein	HMPREF9382_2055	Streptococcus sanguinis SK115	90

A0A346NBA4	TIGR04197 family type VII secretion effector	DDV21_004010 DDV21_004700 DDV23_11140	Streptococcus chenjunshii	93
A0A4R6ZPZ6	Type VII secretion effector (TIGR04197 family)	DFP96_102255	Listeria rocourtiae	96
A0A842AZ36	TIGR04197 family type VII secretion effector	HCJ13_15535	Listeria booriae	102
A0A2C1R825	TIGR04197 family type VII secretion effector	CON44_02325	Bacillus cereus	99
A0A5F0MN81	Uncharacterized protein	CKN67_04115 CKN75_04550	Carnobacterium divergens (Lactobacillus divergens)	92
A0A2W3Z748	TIGR04197 family type VII secretion effector	CI088_09490	Enterococcus plantarum	96
A0A0N0KSY6	Type VII secretion effector	AEQ18_02375	Enterococcus sp. RIT-PI-f	88
A0A7X0T610	TIGR04197 family type VII secretion effector	HB853_09795	Listeria welshimeri	88
A0AFZ6	Uncharacterized protein	lwe0510	Listeria welshimeri serovar 6b ATCC 35897	88
C5NV31	Uncharacterized protein	GEMHA0001_1408	Gemella haemolysans ATCC 10379	97

A0A2L0D3S0	TIGR04197 family type VII secretion effector	C0J00_04045	Streptococcus pluranimalium	119
A0A2X3XH13	Type VII secretion effector	NCTC11085_01347	Streptococcus sanguinis	97
F3UAG5	Uncharacterized protein	HMPREF9393_0458	Streptococcus sanguinis SK1056	97
A0A7H8UYP6	TIGR04197 family type VII secretion effector	FDP16_01520	Streptococcus sanguinis	90
A0A427Z4K1	Uncharacterized protein	D8889_08520	Streptococcus sanguinis	90
A0A0J6L9L6	Type VII secretion effector	VK90_07905	Bacillus sp. LK2	99
F0IN34	Uncharacterized protein	HMPREF9383_1537	Streptococcus sanguinis SK150	104
C5NV27	Uncharacterized protein	GEMHA0001_1404	Gemella haemolysans ATCC 10379	97
A0A6I3IT94	TIGR04197 family type VII secretion effector	GGH90_02870	Streptococcus sp. zg-36	86
A0A6I3I620	TIGR04197 family type VII secretion effector	GGG87_02865	Streptococcus sp. zg-86	86
A0A6I4RB72	TIGR04197 family type VII secretion effector	GGH11_02895	Streptococcus sp. zg-70	102
A0A1E5GIS0	Type VII secretion effector	BCR25_08220	Enterococcus termitis	96

W7C2R9	Uncharacterized protein	MFLO_13765	Listeria floridensis FSL S10-1187	96
A0A2C6WMQ9	TIGR04197 family type VII secretion effector	BTJ66_11860	Staphylococcus edaphicus	91
A0A5A7ZNX9	TIGR04197 family type VII secretion effector	FKX92_06255	Streptococcus sanguinis	90
A0A1E5GX99	Type VII secretion effector	BCR23_04625	Enterococcus quebecensis	96
A0A2X3V3S3	Type VII secretion effector	D8883_04730 NCTC11085_00302	Streptococcus sanguinis	90
A0A2N6SD26	TIGR04197 family type VII secretion effector	CJ218_07575	Gemella sanguinis	94
A0A1E5H6D9	Uncharacterized protein	BCR24_09880	Enterococcus ureilyticus	92
A0A7Z7QU85	Type VII secretion effector	NCTC8183_01311	Streptococcus agalactiae	126
F3UDH1	Uncharacterized protein	HMPREF9393_1578	Streptococcus sanguinis SK1056	90
J4X2D6	Type VII secretion effector, TIGR04197 family	HMPREF1150_0118	Streptococcus sp. AS14	90
A0A0F5MK39	Uncharacterized protein	RN86_02680	Streptococcus gordonii	116
A0A428AH08	Uncharacterized protein	D8875_04305	Streptococcus sanguinis	90
A0A2I1Z9Q6	TIGR04197 family type VII	CYK23_08645	Streptococcus salivarius	90

	secretion effector			
A0A841YI15	TIGR04197 family type VII secretion effector	HB844_13830	Listeria fleischmannii	97
A0A2N6SD59	TIGR04197 family type VII secretion effector	CJ218_07595	Gemella sanguinis	94
A0A2V3VWP8	Type VII secretion effector (TIGR04197 family)	DFR56_108171	Pseudogracilibacillus auburnensis	88
A0A841YHX4	TIGR04197 family type VII secretion effector	HB844_13140	Listeria fleischmannii	96
A0A7X1CAH5	TIGR04197 family type VII secretion effector	HCJ38_14380	Listeria immobilis	97
A0A1J4HAR3	Type VII secretion protein	HMPREF3241_05535	Staphylococcus sp. HMSC34G04	91
A0A3D8TTD4	Uncharacterized protein	UR08_00425	Listeria kieliensis	97
F0FPX7	Uncharacterized protein	HMPREF9392_0404	Streptococcus sanguinis SK678	90
F2CGJ8	Uncharacterized protein	HMPREF9391_1993	Streptococcus sanguinis SK408	90
F0ISH9	Uncharacterized protein	HMPREF9384_0791	Streptococcus sanguinis SK160	90
G5JNA7	Uncharacterized protein	STRCR_0144	Streptococcus criceti HS-6	93
A0A1E5GH45	Uncharacterized protein	BCR21_07315	Enterococcus ureasiticus	92

A0A7I0BHX0	TIGR04197 family type VII secretion effector	E1N03_11860	Staphylococcus epidermidis	91
A0A829M3W1	Type VII secretion protein	M453_0212855	Staphylococcus epidermidis CIM40	91
R2SNU8	Type VII secretion effector	UAY_02591	Enterococcus moraviensis ATCC BAA-383	96
A0A0B7GN05	Putative type VII secretion effector	SSV_1921	Streptococcus sanguinis	90
W7B2A2	Uncharacterized protein (Fragment)	MAQA_04586	Listeria aquatica FSL S10-1188	82
A0A1H9PSA2	Type VII secretion effector, SACOL2603 family	SAMN04488559_10172	Isobaculum melis	94
W7B6K0	Uncharacterized protein (Fragment)	MAQA_04296	Listeria aquatica FSL S10-1188	83
A0A841YE47	TIGR04197 family type VII secretion effector	HB844_07260	Listeria fleischmannii	90
A0AK27	Uncharacterized protein	lwe1941	Listeria welshimeri serovar 6b ATCC 35897	97
A0A242AUE1	Uncharacterized protein	A5821_000621	Enterococcus sp. 7F3_DIV0205	92
A0A7X0Y3T3	TIGR04197 family type VII secretion effector	HCA69_08900	Listeria grandensis	90
A0A7X1C884	TIGR04197 family type VII	HCJ38_03345	Listeria immobilis	97

	secretion effector			
W7B9C0	Uncharacterized protein	MAQA_15976	Listeria aquatica FSL S10-1188	97
A0A172Q5Q7	Uncharacterized protein	A0O21_01495	Streptococcus pantholopis	90
A0A7X0XCD6	TIGR04197 family type VII secretion effector	HCI99_06655	Listeria booriae	90
A0A239X809	Type VII secretion effector	SAMEA4504048_01597	Streptococcus acidominimus	105
V6Z4V5	Uncharacterized protein	SAG0136_11275	Streptococcus agalactiae LMG 14747	105
A0A540UVH0	TIGR04197 family type VII secretion effector	FH692_06345	Streptococcus suis	106
W7CDF0	Uncharacterized protein	MFLO_01075	Listeria floridensis FSL S10-1187	97
A0AKF5	Uncharacterized protein	lwe2069	Listeria welshimeri serovar 6b ATCC 35897	97
A0A7I0AJ13	TIGR04197 family type VII secretion effector	E1N03_09545	Staphylococcus epidermidis	91
A0A7X1C121	TIGR04197 family type VII secretion effector	HB856_09015	Listeria booriae	96
A0A2K4FCE9	TIGR04197 family type VII secretion effector	CD039_08645	Staphylococcus argensis	91

Q8DZR6	Uncharacterized protein	SAG1032	Streptococcus agalactiae serotype V ATCC BAA-611	85
A0A1F0CEK0	Uncharacterized protein	HMPREF2570_04395	Streptococcus sp. HMSC069D09	85
J8J5K7	Uncharacterized protein	IIO_06123	Bacillus cereus VD115	91
A0A1E5L0N0	Uncharacterized protein	BCR26_07815	Enterococcus rivorum	103
C0MDX1	Uncharacterized protein	SZO_07980	Streptococcus equi subsp. zooepidemicus (strain H70)	104
A0A076Z409	TIGR04197 family type VII secretion effector (Type VII secretion effector)	C4618_05905 D5F95_10620 DK41_05465 NCTC6175_01412 NCTC8185_02368	Streptococcus agalactiae	116
A0A829IEV4	Uncharacterized protein	SAG0014_09640	Streptococcus agalactiae FSL S3-586	116
Q8E5G5	Uncharacterized protein	gbs1067	Streptococcus agalactiae serotype III (strain NEM316)	116
A0A243G320	Type VII secretion effector	BK774_26435	Bacillus thuringiensis	91
A0A428IGV6	Uncharacterized protein	D8844_06495	Streptococcus oralis	121
A0A2S7RWC9	TIGR04197 family type VII secretion effector	CUS89_04340	Enterococcus mundtii	93

Table 4.5. Accession codes and sequence information for LapD1 homologs identified with one iteration of JackHMMER.

Entry	Protein names	Gene names	Organism	Length
A0A1F0ZSZ0	Type VII secretion protein	HMPREF2917_09355	Streptococcus sp. HMSC061E03	117
A0A359YHE7	Uncharacterized protein	SPADD19_01412	Streptococcus parasanguinis	117
I1ZK44	Uncharacterized protein	Spaf_0401	Streptococcus parasanguinis FW213	117
A0A2I1TT29	TIGR04197 family type VII secretion effector	CYK20_05490	Streptococcus parasanguinis	117
A0A6I3PAZ6	TIGR04197 family type VII secretion effector	GMC80_04760 GMC84_06705	Streptococcus parasanguinis	117
A0A1V0H196	TIGR04197 family type VII secretion effector	A6J85_03505	Streptococcus gordonii	118
A0A0F5MM43	Type VII secretion protein	RN86_02705	Streptococcus gordonii	118
S7XKY2	Type VII secretion protein	M059_05530	Streptococcus mitis 18/56	118
A0A1X1L326	Type VII secretion effector	B7692_08470 B7696_07565 B7700_09665	Streptococcus mitis	118
A0A178KGQ9	Type VII secretion protein	A3Q39_01930	Streptococcus sp. CCUG 49591	118
A0A414PGR1	TIGR04197 family type VII secretion effector	DW666_08555	Streptococcus parasanguinis	117
F8DHG1	Uncharacterized protein	HMPREF0833_11761	Streptococcus parasanguinis ATCC 15912	117

A0A3R9LZL4	Uncharacterized protein	D8803_08265	Streptococcus oralis	119
A0A428EFW5	Uncharacterized protein	D8839_01320	Streptococcus mitis	119
E6KIQ0	Uncharacterized protein	HMPREF8578_0115	Streptococcus oralis ATCC 49296	119
A0A8B1YMV5	TIGR04197 family type VII secretion effector	J4854_01600	Streptococcus lactarius	117
F9LWN2	Uncharacterized protein	HMPREF9965_0735	Streptococcus mitis bv. 2 str. SK95	119
A0A7H9FG17	TIGR04197 family type VII secretion effector	HRE59_00320	Streptococcus oralis subsp. oralis	119
A0A3R9PR96	Uncharacterized protein	D8860_09790	Streptococcus oralis	119
A0A1X1IPJ3	Type VII secretion effector	B7710_00065	Streptococcus oralis subsp. oralis	119
A0A139PJZ1	Uncharacterized protein	SORDD21_01112	Streptococcus oralis	119
A0A3L8GDQ6	TIGR04197 family type VII secretion effector	DIY07_08815	Streptococcus iniae (Streptococcus shiloi)	116
A0A178KI70	Type VII secretion protein	A3Q39_01960	Streptococcus sp. CCUG 49591	121
A0A1B1IDA9	Type VII secretion protein	AXF18_01820	Streptococcus sp. oral taxon 064	117
A0A427ZT45	Uncharacterized protein	D8882_08140	Streptococcus sanguinis	128
A3CR33	Uncharacterized protein	SSA_2276	Streptococcus sanguinis SK36	128
A0A3R9JBV3	Uncharacterized protein	D8860_05090	Streptococcus oralis	117
K0ZUT2	Uncharacterized protein	GMD4S_06157	Streptococcus sp. GMD4S	117

A0A3R9FWZ7	Uncharacterized protein	D8894_04895	Streptococcus oralis	117
K1A200	Uncharacterized protein	GMD6S_07863	Streptococcus sp. GMD6S	117
E6KIR3	Uncharacterized protein	HMPREF8578_0128	Streptococcus oralis ATCC 49296	117
A0A1X1IMY5	Type VII secretion effector	B7710_01125	Streptococcus oralis subsp. oralis	117
I0Q2A4	Type VII secretion effector, TIGR04197 family	HMPREF1115_1417	Streptococcus oralis SK610	117
F3UNP7	Uncharacterized protein	HMPREF9389_0455	Streptococcus sanguinis SK355	128
A0A1X1HVT4	Type VII secretion effector	B7714_09150	Streptococcus oralis subsp. oralis	117
S7XHE0	Type VII secretion protein	M059_05500	Streptococcus mitis 18/56	121
A0A1X1KD41	Type VII secretion effector	B7692_08440 B7696_07595	Streptococcus mitis	121
A0A428IGV6	Uncharacterized protein	D8844_06495	Streptococcus oralis	121
E3CF41	Uncharacterized protein	HMPREF9626_1803	Streptococcus parasanguinis F0405	117
A0A1X1JX30	Type VII secretion effector	B7700_09695	Streptococcus mitis	121
F0FHF7	Uncharacterized protein	HMPREF9388_2140	Streptococcus sanguinis SK353	128
A0A1X1J482	Type VII secretion effector	B7708_00965	Streptococcus oralis subsp. dentisani	121
A0A3R9KBB5	Uncharacterized protein	D8801_04895	Streptococcus oralis	121

A0A076Z409	TIGR04197 family type VII secretion effector (Type VII secretion effector)	C4618_05905 D5F95_10620 DK41_05465 NCTC6175_01412 NCTC8185_02368	Streptococcus agalactiae	116
Q8E5G5	Uncharacterized protein	gbs1067	Streptococcus agalactiae serotype III strain NEM316	116
A0A829IEV4	Uncharacterized protein	SAG0014_09640	Streptococcus agalactiae FSL S3-586	116
A0A0E1EMX7	TIGR04197 family type VII secretion effector	AX245_04155 C4618_11685 C6N07_05895 RDF_1030	Streptococcus agalactiae	111
A0A837KW31	Uncharacterized protein	WA04_10840	Streptococcus agalactiae	116
A0A4R5G605	TIGR04197 family type VII secretion effector	E0E04_02150	Streptococcus vicugnae	118
A0A7D4GGP0	TIGR04197 family type VII secretion effector	FOC63_06865	Streptococcus gallolyticus	118
A0A139N5A4	Uncharacterized protein	SCRDD08_00137	Streptococcus cristatus	121
A0A4T2GKS9	TIGR04197 family type VII secretion effector	FAJ39_07705	Streptococcus suis	109
A0A7X2UFL6	Type VII secretion effector	NCTC3858_01464	Streptococcus uberis	111
A0A7Z0VGH5	Uncharacterized protein	TH70_0120	Streptococcus agalactiae	111
A0A8B4IN87	Type VII secretion effector	NCTC3858_00392	Streptococcus uberis	108
A0A7Z7QU85	Type VII secretion effector	NCTC8183_01311	Streptococcus agalactiae	126

Q8DZR6	Uncharacterized protein	SAG1032	Streptococcus agalactiae serotype V ATCC BAA-611	85
A0A1F0CEK0	Uncharacterized protein	HMPREF2570_04395	Streptococcus sp. HMSC069D09	85
A0A1E5KUF8	Uncharacterized protein	BCR26_04480	Enterococcus rivorum	93
A0A3R9NTY4	Uncharacterized protein	D8879_08845	Streptococcus sanguinis	90
A0A1E5KVA8	Uncharacterized protein	BCR26_15430	Enterococcus rivorum	93
A0A0F5MK39	Uncharacterized protein	RN86_02680	Streptococcus gordonii	116
A0A7H8V643	TIGR04197 family type VII secretion effector	FFV08_03635	Streptococcus sanguinis	125
F3USM5	Uncharacterized protein	HMPREF9389_1833	Streptococcus sanguinis SK355	90
A0A3R9IAM7	TIGR04197 family type VII secretion effector	D8887_08455 FFV08_05580	Streptococcus sanguinis	90
A0A427ZP46	Uncharacterized protein	D8886_07895	Streptococcus sanguinis	90
A0A5A7ZT92	TIGR04197 family type VII secretion effector	FKX92_00595	Streptococcus sanguinis	124
A0A139NQ79	Uncharacterized protein	STRDD11_02464	Streptococcus sp. DD11	89
A0A540UNN4	TIGR04197 family type VII secretion effector	FH692_10960	Streptococcus suis	134
A0A0Z8X7W8	Type VII secretion effector	ERS132372_01527 ERS132399_02390	Streptococcus suis	111

A0A116LSC7	TIGR04197 family type VII secretion effector (Type VII secretion effector)	ERS132406_02093 ERS132410_02193 FAJ36_02915	Streptococcus suis	108
A0A4P7WT47	TIGR04197 family type VII secretion effector	E8M06_09960	Streptococcus suis	108
A0A0Z8DGM0	TIGR04197 family type VII secretion effector (Type VII secretion effector)	E8M06_09990 ERS132392_00702 JZY07_10375	Streptococcus suis	108
A0A0S3K715	Type VII secretion effector	ATZ33_01365	Enterococcus silesiacus	104
A0A1E5HGI4	Type VII secretion effector	BCR24_01530	Enterococcus ureilyticus	104
A0A0B7GN05	Putative type VII secretion effector	SSV_1921	Streptococcus sanguinis	90
A0A239X809	Type VII secretion effector	SAMEA4504048_01597	Streptococcus acidominimus	105
V6Z4V5	Uncharacterized protein	SAG0136_11275	Streptococcus agalactiae LMG 14747	105
A0A540UVH0	TIGR04197 family type VII secretion effector	FH692_06345	Streptococcus suis	106
A0A242AX92	Uncharacterized protein	A5821_001500	Enterococcus sp. 7F3_DIV0205	99
A0A139QYV5	Uncharacterized protein	SGADD02_00817 SGADD03_01202	Streptococcus gallolyticus	90
A0A380K862	Type VII secretion effector	NCTC13767_01892	Streptococcus gallolyticus	90

A0A1E5GK70	Type VII secretion effector	BCR25_06445	Enterococcus termitis	104
F0IN34	Uncharacterized protein	HMPREF9383_1537	Streptococcus sanguinis SK150	104
A0A2G3NUY4	TIGR04197 family type VII secretion effector	CS009_05415 CS010_03220	Streptococcus macedonicus	90
A0A3R9G5C1	Uncharacterized protein	D8887_07710	Streptococcus sanguinis	92
A0A0A0DFU7	Uncharacterized protein	SSIN_0557	Streptococcus sinensis	104
A0A2L0D3S0	TIGR04197 family type VII secretion effector	C0J00_04045	Streptococcus pluranimalium	119
A0A4T2H474	TIGR04197 family type VII secretion effector	FAJ36_02885	Streptococcus suis	134
F5X0A7	Uncharacterized protein	SGGB_0839	Streptococcus gallolyticus ATCC 43143	90
A0A7D4GS34	TIGR04197 family type VII secretion effector	FOC63_08560	Streptococcus gallolyticus	90
A0A1S5WBI2	Uncharacterized protein	BTR42_04595	Streptococcus gallolyticus subsp. gallolyticus DSM 16831	90
A0A359YGK2	Uncharacterized protein	SPADD19_01110	Streptococcus parasanguinis	89
A0A242H2M5	Uncharacterized protein	A5866_002123	Enterococcus sp. 12C11_DIV0727	103
A0A7H8UYP6	TIGR04197 family type VII secretion effector	FDP16_01520	Streptococcus sanguinis	90

A0A0U2NRL1	Type VII secretion effector TIGR04197 family type VII secretion effector	ATZ35_10775	Enterococcus rotai	103
A0A4Q2FH31		DF218_03565	Streptococcus parasanguinis	89
A0A1I7FJ84	Type VII secretion effector, SACOL2603 family	SAMN05660328_101420	Streptococcus gallolyticus	90
A0A1E5GE80	Type VII secretion effector	BCR21_11960	Enterococcus ureasiticus	103
A0A428AH08	Uncharacterized protein	D8875_04305	Streptococcus sanguinis	90
A0A2I1Z9Q6	TIGR04197 family type VII secretion effector	CYK23_08645	Streptococcus salivarius	90
F3UDH1	Uncharacterized protein	HMPREF9393_1578	Streptococcus sanguinis SK1056	90
I2NMG3	Uncharacterized protein	HMPREF9971_1232	Streptococcus parasanguinis F0449	113
A0A242LA78	Uncharacterized protein	A5881_003608	Enterococcus termitis	104
A0A6I3PR01	TIGR04197 family type VII secretion effector	GMC95_02245	Streptococcus parasanguinis	94
A0A0E2IQB7	Uncharacterized protein	HMPREF1654_01870	Streptococcus intermedius ATCC 27335	91
R2QN21	Type VII secretion effector	UAY_02986	Enterococcus moraviensis ATCC BAA-383	98
A0A139MU55	Uncharacterized protein	STRDD04_00268	Streptococcus sp. DD04	97
A0A6N3CT23	Uncharacterized protein	SPLFYP13_01158	Streptococcus parasanguinis	89

A0A1F1A3X5	Uncharacterized protein	HMPREF2917_04405	Streptococcus sp. HMSC061E03	89
A0A427Z4K1	Uncharacterized protein	D8889_08520	Streptococcus sanguinis	90
A0A1F0AWW4	Uncharacterized protein	HMPREF2686_08175	Streptococcus sp. HMSC057G03	89
J4X2D6	Type VII secretion effector, TIGR04197 family	HMPREF1150_0118	Streptococcus sp. AS14	90
A0A0B7GL02	Putative type VII secretion effector	SSV_1220	Streptococcus sanguinis	97
A0A4Q5BT34	TIGR04197 family type VII secretion effector	GMC84_09185 GMC94_02205	Streptococcus parasanguinis	89
F8DGG8	Uncharacterized protein	HMPREF0833_10386	Streptococcus parasanguinis ATCC 15912	94
E8K4F1	Uncharacterized protein	HMPREF8577_0436	Streptococcus parasanguinis ATCC 903	99

Methods

Bacterial strains, plasmids, and growth conditions

S. intermedius strains used in this study were generated from either the B196 or GC1825 wild-type strains and genomic DNA isolated from these strains was used for molecular cloning. *E. coli* XL1-blue was used for molecular cloning and plasmid maintenance. *E. coli* BL21 (DE3) CodonPlus and B834 (DE3) were used for protein expression of native and selenomethionine substituted proteins, respectively. The complete list of bacterial strains generated for this study can be found in Table S4.1. *E. coli* overexpression was performed using the IPTG-inducible pETDuet-1 and pET29b vectors, while pDL277 was used for constitutive gene expression in *S. intermedius*. PCR amplification of genes of interest for this study was done with Phusion polymerase (NEB). For pET vector cloning the PCR amplicons were digested with restriction endonucleases NdeI/XhoI for pET29b/pETduet-1 MCS2 or BamHI/SalI for pETduet-1 MCS1. DNA ligation was then done using T4 DNA ligase. These constructs were cloned with N- or C-terminal His-6 tags to facilitate affinity purification as required. Cloning into the pDL277 vector was done with restriction endonucleases BamHI/SalI followed by ligation with T4 DNA ligase. In this case, *S. intermedius* genes were fused with the P96 promoter of *Streptococcus pneumoniae* by splicing by overlap extension (SOE) PCR as previously described (Whitney et al., 2017). A complete list of the plasmids used in this study can be found in Table S4.2. *E. coli* was grown in lysogeny broth at 37°C at 225rpm. 50ug/mL kanamycin and 150ug/mL carbenicillin was added to the media when growing strains with the pET29b and pETduet-1 vectors, respectively. *S. intermedius*

strains were grown in Todd Hewitt broth supplemented with 0.5% yeast extract at 37°C and 5% CO_2 without shaking. 50ug/mL of spectinomycin for *S. intermedius* or 100ug/mL of spectinomycin for *E. coli* was added to media when growing strains with the pDL277 plasmid. For all *S. intermedius* experiments, strains were first grown on solid media before being inoculated into liquid culture to ensure consistent growth between strains.

DNA manipulation

S. intermedius B196 and GC1825 genomic DNA was prepared by using InstaGene Matrix (Bio-Rad) to extract and purify DNA from 2 mL of cells pelleted from an overnight culture. Primers used in this study were synthesized by Integrated DNA Technology (IDT). Molecular cloning was performed using Phusion polymerase, appropriate restriction enzymes, and T4 DNA ligase (NEB). All Sanger sequencing was performed by Genewiz/Azenta Life Sciences.

Transformation of *S. intermedius*

S. intermedius B196 and GC1825 strains were back diluted 1:10 from an overnight culture, grown to $OD_{600} = 0.5$, and supplemented with 5uL of 0.1mg/mL competence stimulating peptide (DSRIRMGFDFSKLFGK, synthesized by Genscript). Cultures were then incubated at 37°C and 5% CO_2 without shaking for 45 minutes (GC1825) or two hours (B196). Approximately 100ng of plasmid or linear DNA was then added and the cultures were again incubated for three hours (1 hour for GC1825). 100uL of these cultures were then plated on Todd Hewitt plates supplemented with 0.5%

yeast extract and either 50 ug/mL spectinomycin to select for pDL277 transformants or 250 ug/mL kanamycin for allelic replacement mutants.

Gene deletion in *S. intermedius* by allelic replacement

Our *S. intermedius* gene deletion protocol was previously described in (Klein et al., 2021). In brief, deletion constructs were made using SOE PCR to fuse a spectinomycin promoter to a kanamycin resistance cassette flanked by two 1000bp fragments of DNA that are immediately adjacent to the target gene. These constructs were cloned into pETduet-1 with the final plasmid designation being pETduet-1::5'geneflank_SpecProm_*kanR*_3'geneflank. Plasmids were then digested with BamHI and NotI and the deletion fragment was gel extracted (Monarch DNA gel extraction kit, NEB). 100ng of purified deletion fragment was then added to competent *S. intermedius* cells and mutants were selected for by plating on Todd Hewitt agar with 0.5% yeast extract and 250 ug/mL kanamycin. All gene deletions were confirmed by colony PCR.

Secretion assays

20mL cultures of *S. intermedius* were grown overnight to an OD_{600}=1.0. Cell and supernatant fractions were then separated by centrifugation at 4000g for 15 minutes and cell fractions were washed once in PBS pH 7.4 before being resuspended in 100 uL of PBS. 100 uL of Laemmli buffer was added and samples were boiled for 10 minutes. Supernatant fractions were incubated at 4°C overnight after adding trichloroacetic acid to a final concentration of 10%. Precipitated proteins were then centrifuged at 35,000g for

30 minutes and the resulting pellets were washed once with cold acetone. The pellets were then centrifuged at 35,000g for an additional 30 minutes and the acetone was decanted off. Any remaining acetone was left to evaporate off in a fume hood. The dry pellets were then resuspended in minimal Laemmli buffer diluted with urea (300uL 4X Laemmli, 600uL 8M urea) and boiled for 10 minutes. Both the cell and secreted samples were analysed using SDS PAGE gels run with a tris-tricine based running buffer (see below) and Western blot analysis.

Antibody generation

A custom polyclonal antibody for the TelD protein was generated for this study by Genscript. The LXG domain of TelD (amino acids 1-203) with a C-terminal His_6 tag was expressed and purified by affinity and size exclusion chromatography as described (see "protein expression and purification") except with PBS pH 7.4 in place of Tris-HCl pH 8.0. In total, 10mg of protein was shipped to Genscript for antibody production. Generation of the a-TelC and a-EsxA antibodies have been described previously (Whitney et al., 2017; Klein et al., 2021).

SDS-PAGE, SYPRO red staining and Western blotting

SDS-PAGE gels run for this study were done using a tris-tricine buffer system (200mM Tris, 100mM Tricine, 0.1% SDS, pH 8.3) to better resolve low molecular weight proteins (<20kDa) (Schägger, 2006). Protein visualization on SDS-PAGE gels was done with the SYPRO Red protein gel stain (Invitrogen). The gel was rinsed briefly

in DI water before being stained for one hour with 1:5000 SYPRO Red (Invitrogen) diluted in 10% (v/v) acetic acid. The gel was then destained for 15 minutes in 7.5% (v/v) acetic acid before being imaged on a Chemidoc imaging system (Bio-Rad). For western blots, the resolved proteins were transferred to a nitrocellulose membrane by wet transfer (100V, 30 minutes). Nitrocellulose membranes were then blocked with 5% skim milk dissolved in TBS-T for 30 minutes with light agitation followed by addition of primary antibody (titer 1:5000) to the blocking buffer and further incubation for 1 hour. Blots were washed for five minutes three times with TBS-T then incubated in TBS-T with an HRP-conjugated anti-rabbit secondary antibody (titer 1:5000) for 45 minutes. After three additional five-minute washes, the blots were developed using Clarity Max Western ECL reagent (Bio-Rad) and imaged with a ChemiDoc XRS+ (Bio-Rad).

Co-immunoprecipitation in *Streptococcus intermedius*

Co-immunoprecipitation assays were performed on VSV-G tagged TelC in a Δ*telC-tipC2* background (ΔSIR_1486-1489) and VSV-G tagged WxgC in a Δ*wxgC* background (ΔSIR_1491). In both experiments, strains lacking SIR1486-1489 or SIR1491 but containing empty pDL277 were used as negative controls. 50mL cultures of *S. intermedius* were grown to an OD of 0.5 and centrifuged at 5000g for 15 minutes to harvest cells. The pellets were then resuspended and incubated in lysis buffer (20mM Tris-HCl pH 7, 150mM NaCl, 10% glycerol, 5 mg/mL lysozyme, 100U/mL mutanolysin, 1 mM PMSF) and incubated at 37°C for 30 minutes. Cells were lysed by sonication (three, thirty second pulses at 30 amps) and the cell pellets were removed by

centrifugation at 30,000g for 30 minutes at 4°C. The supernatants were then transferred to fresh 2 mL Eppendorf tubes and incubated with 50uL of anti-VSV-G beads overnight at 4°C with gentle agitation. The beads were harvested by centrifugation at low speed (<100g) and washed thrice with 10 mLs of wash buffer (20mM Tris-HCl pH 7, 150mM NaCl, 10% glycerol). An additional three wash steps were performed with 50 mM ammonium bicarbonate. The beads were then covered in a minimal amount of ammonium bicarbonate buffer and the bound protein was digested with 10 ng/ul of sequencing grade trypsin for four hours at 37°C. The buffer was then harvested, and the beads were washed with an additional 50uL of ammonium bicarbonate buffer to remove any remaining peptides. The peptide samples were then incubated with 1 mM tris(2-carboxyethyl)phosphine for one hour at 37°C to reduce any disulphide bonds. Iodoacetamide was added to a final concentration of 10 mM and the samples were incubated in the dark at room temperature for 30 minutes. This reaction was quenched with 12 mM N-acetylcysteine. The peptides were purified using Pierce C18 spin columns (Thermo Scientific). LC-MS/MS analysis of the purified peptides was done at the Sick Kids Proteomics, Analytics, Robotics, and Chemical Biology Centre (SPARC) at The Hospital for Sick Children.

TelD toxicity assay

E. coli XL1 blue was transformed with either the pSCRhaB2 plasmid encoding the *telD* toxin gene or an empty vector control. For the toxicity plating assay, these strains were OD matched and serially diluted (1:10) then plated on LB plates containing 200

ug/mL of trimethoprim with and without 0.1% L-rhamnose. The plates were incubated at 37°C overnight and then imaged using an iPhone 11 (Apple). Growth curves were generated by back diluting overnight cultures 1:100 into fresh LB media supplemented with 200 ug/mL trimethoprim and 15 ug/mL gentamicin in a 96-well plate. The cell cultures were allowed to grow at 37°C with shaking for 1.5 hours at which point toxin expression was induced by adding L-rhamnose to a final concentration of 0.1% and immunity protein expression was induced by adding IPTG to a final concentration of 0.1 mM. The OD of the cultures was measured with a Synergy 4 Microplate Reader (Biotek Instruments).

Protein expression and purification

All native proteins were expressed in *E. coli* BL21(DE3) CodonPlus whereas selenomethionine-labeled LapD1 was expressed in *E. coli* B834 (DE3). In general, protein expression strains were grown in LB in a shaking incubator at 37°C to an OD_{600}=0.5. Temperature was then lowered to 18°C and protein expression was induced with 1mM IPTG followed by overnight protein expression (approximately 18 hours). Cells were then centrifuged and lysed by sonication (four pulses, 30% amplitude, 30 seconds) in lysis buffer (20mM Tris-HCl pH 8.0, 300mM NaCl, 10mM imidazole). Cellular debris was cleared from the lysate by centrifugation at 35,000g for 30 minutes and the lysate was run over Ni-NTA resin using a gravity flow column on the benchtop. Resin was then washed three times with 20mL lysis/wash buffer and protein was eluted in 4mL of elution buffer (20mM Tris-HCl pH, 8.0, 300mM NaCl, 400mM imidazole).

Eluted protein was further purified by size exclusion chromatography using a HiLoad 16/600 Superdex 200 connected to an ÄKTAexplorer (Cytiva). Selenomethionine-labeled protein was similarly expressed using *E. coli* B834 (DE3) except that the cells were grown in SelenoMethionine Media (Molecular Dimensions) supplemented with 40 mg/L of L-selenomethionine.

Protein crystallization

Native and selenomethionine-labeled LapD2 was concentrated to 10 mg/ml and screened for crystallization conditions using the MCSG1-4 crystallization suites (Anatrace) and the hanging drop vapour diffusion method. After one week, trapezoid shaped crystals formed in 0.2M lithium sulfate, 0.1M Tris-HCl, pH 8.0, 30% (w/v) PEG 4000. Crystals were cryoprotected using a buffer identical to the crystallization buffer but supplemented with 20% ethylene glycol.

X-ray data collection, structure determination and model refinement

X-ray data were collected with the Structure Biology Center sector 19-ID at the Advanced Photon Source. Diffraction of both selenomethionine-incorporated and native protein crystals were measured at a temperature of 100 K using a 0.3s exposure and 0.5 degree of rotation over 450°. Native and selenomethionine-incorporated crystals diffracted to resolutions of 2.20 Å and 2.42 Å, respectively, and the diffraction images were collected on a dectris Pilatus 3 X 6M detector with an X-ray wavelength near the selenium edge of 12.66 keV (0.97926 Å). Diffraction data were processed using the

HKL3000 suite (Minor et al., 2006). The structure of LapD2 was determined by SAD phasing with data from selenomethionine-containing protein crystal using SHELX C/D/E (Sheldrick, 2010), mlphare and dm (Winn et al., 2011), and initial automatic protein model building with Buccaneer (Cowtan, 2006), all implemented in the HKL3000 software package (Minor et al., 2006). The initial model of the structure of the homodimer was completed manually by using Coot (Emsley et al., 2010) and briefly refined using refmac (Murshudov et al., 2011). Using this dimeric structure from the SAD phasing as the search model, molecular replacement was applied with the native data using molrep implemented in HKL3000. The structure was then refined iteratively using Coot for manual adjustment and Phenix (phenix.refine) (Afonine et al., 2012) for restrained refinement until R_{work} and R_{free} values converged to 0.23 and 0.26, respectively. The final refined structure contained two copies of homodimeric LapD1 with each dimer formed through a disulfide bond. The stereochemistry of the structure was assessed using PROCHECK (Laskowski, 2001) and a Ramachandran plot and was validated using the PDB validation server. X-ray data and refinement statistics are listed in Table 4.1. All structural figures were generated using UCSF ChimeraX (Goddard et al., 2018).

Protein structure prediction and analysis

Surface hydrophobicity (Testa et al., 1996), conservation mapping (Pei & Grishin, 2001), structural alignments were visualized using ChimeraX's built in functions with default parameters (Pettersen et al., 2021). DALI pairwise was used to calculate reported RMSD values (Holm, 2020). 2D protein structure predictions were generated by

PSIPRED 4.0 on the UCL PSIPRED Workbench (Buchan & Jones, 2019) (http://bioinf.cs.ucl.ac.uk/psipred/). 3D Protein structure predictions were performed by AlphaFold v2.0.0 running on our local server with default parameters (Jumper et al., 2021). Multimer predictions were calculated using ColabFold using default parameters.

Sequence analysis, conservation mapping and sequence logos

Homologous sequences to LapC1, LapC2, LapD1 and LapD2 were identified using JackHMMER (HmmerWeb version 2.41.2) searches of the UniprotKB database, restricted to the phylum Firmicutes, iterating until at least 100 sequences were obtained (Finn et al., 2015). Accessions were downloaded and full sequences of active entries were subsequently retrieved from Uniprot. Duplicate sequences were removed and the remaining aligned using MAFFT (scoring matrix: BLOSUM30) (Katoh & Standley, 2013) implemented in Geneious Prime 2022.1.0 (www.geneious.com). Final sequence lists used for HMM logo generation can be found in Tables S4.3, S4.4, and S4.5. HMMs were generated and initially visualized by uploading multiple sequence alignments to the Skylign webserver (www.skylign.org) and set to "create HMM – remove mostly empty columns" (Wheeler et al., 2014). The resulting matrices were downloaded as tabular text, formatted and then visualized using Logomaker (Tareen & Kinney, 2020). Sequence alignments depicted in Supplemental Figures S4.3, S4.4 and S4.6 were generated using M-Coffee on the T-Coffee webserver (https://tcoffee.crg.eu) (Moretti et al., 2007) and visualized with the ESPript 3.0 webserver (Robert & Gouet, 2014) (https://espript.ibcp.fr/ESPript/ESPript/).

Data availability

The data supporting Chapter IV can be found entirely within this thesis. Structure files and information pertaining to the structure of LapD2 are indexed in the protein data bank (PDB: 7UH4). For access to strains and plasmids used in this chapter please contact Dr. John Whitney.

Chapter V – Conclusions and future directions

Overview

Through my doctoral work, I have attempted to advance our understanding of several aspects of T7SSb structure and function. Using X-ray crystallography, I have elucidated the experimental structures of several proteins associated with this system. This approach, in conjunction with a substantial amount of protein biochemistry, has ultimately revealed several novel details of T7SSb function. I started my graduate studies by examining an immunity protein that is protective against a T7SSb-secreted antibacterial toxin of *S. intermedius*. Through this work, I sought to better understand how this immunity protein interacts with and protects against its cognate toxin. Next, I sought to understand the function of EsaA, which previous to my work, was the most poorly understood T7SSb apparatus protein in terms of its overall architecture. The hypothesis driving this work was that EsaA forms a conduit through which effector secretion occurs. Although our work did not outright prove this hypothesis, it was supportive of the idea, and we learned much about EsaA's structure and topology. The final focus of my graduate work was on T7SSb effector recognition. My work has begun to suggest a model where the LXG domain of T7SSb effectors requires direct interaction with two small α-helical chaperones to form a pre-secretion complex which is then recognized by the T7SSb apparatus. Although future studies will be needed to fully understand the exact determinants of LXG effector recognition, we have shown minimally the necessity of the pre-secretion complex and a conserved FxxxD motif found in Lap1 chaperones. The rest of this section will contain a summary of each of the three chapters as well as the greater context of the field and future directions.

Chapter II summary and discussion

Current understandings of T7SSb immunity proteins

The T7SSb is an antibacterial weapon used by Firmicutes bacteria to inhibit the growth of competitors (Klein et al., 2020). One of the tenets of these bacterial competition systems is that bacteria-targeting effectors are co-transcribed with immunity proteins. Immunity proteins are essential for blocking self-intoxication and intoxication from sister cells (Carr et al., 2000; Klein et al., 2018; Ting et al., 2018). In this regard, the T7SSb is no different from the various antibacterial systems of Gram-negative bacteria and all T7SSb toxins that have been characterized to date have a corresponding immunity gene downstream of the effector gene. Furthermore, many of these immunity determinants have been proven to inhibit toxicity in cells although the exact mechanism of this inhibition is not entirely understood.

In Chapter II, I attempted to deepen our understanding of the interaction between effectors and immunity proteins by studying TelC-TipC1 as a model toxin-immunity pair. We found that TipC1 is a membrane protein that faces the IWZ of *S. intermedius*. The directionality of TipC1 is critical to its function as the TelC toxin is a lipid II phosphatase that is active specifically in the IWZ. *tipC* genes are generally found in *telC*-containing operons, but interestingly, there can often be multiple homologs of *tipC* in these operons and *tipC* genes can also be found in poly-immunity loci (Klein et al., 2018). This latter finding is reminiscent of research on the T6SS in that T6SS immunity genes also frequently cluster in poly-immunity loci and these clusters have been shown to be part of mobile genetic elements that function to spread genes (such as immunity determinants)

throughout broad bacterial populations (Ross et al., 2019). Although it is not known if T7SSb genes can be mobilized, it is an intriguing possibility and requires future work.

I showed that, as a general principle, only one of the multiple TipC homologs (TipC1) encoded by *telC* operons bind to and inhibits the cognate TelC toxin. I used crystallography and mutagenesis to probe which parts of TipC1 contribute to TelC binding. TipC1 forms a mixed α/β fold with a concave face and a convex face. By mapping the conservation between the TelC-interacting TipC1 and its non-interacting homolog TipC2, I determined that the concave face of the protein was more likely to facilitate binding to TelC as it was in this region that TipC1 and TipC2 showed stark differences. Indeed, by probing various conserved residues in this face with mutagenesis, I determined that TipC1 interacts with TelC via the concave surface and mutation of these conserved residues will lead to a decrease in TelC binding and inhibition (Klein et al., 2018).

Since the publication of my work on TipC1, several papers focusing on T7SSb-secreted toxins have corroborated some of my findings on T7SSb immunity proteins. A study by Ulhuq et al. characterized the first LXG effector from *S. aureus*. The effector, called type seven dependent protein A (TspA), is a membrane depolarizing toxin that is bacteriostatic when expressed in *E. coli* (Ulhuq et al., 2020). Using a zebrafish hindbrain model, they showed that toxins like TspA and EsaD yield an advantage to *S. aureus in vivo* over related strains that lack immunity determinants. *tspA* is encoded alongside an immunity gene called *tsaI*. Similar to our findings with *tipC*, *tspA*-encoding strains of *S. aureus* often harbour multiple copies of *tsaI*. Intriguingly, some of these *tsaI* genes seem

to be protective against the toxicity of TspA homologs from other strains suggesting that bacteria can accumulate immunity determinants to protect against toxins that they do not encode. Although we previously showed that Firmicutes often encode multiple immunity homologs, this work provides evidence that these determinants may indeed be protective against toxin homologs from competitor strains (Ulhuq et al., 2020).

A later study on the T7SSb toxin-immunity pairs of *B. subtilis* confirmed that, similar to *Staphylococcus* and *Streptococcus*, LXG effectors contribute to interbacterial killing and that this killing can be inhibited by specific immunity proteins (Kobayashi, 2021). Indeed, intraspecies competition in *B. subtilis* seems to be quite potent as strains can encode up to nine toxin-immunity pairs. Although, this work did not develop our understanding of how immunity proteins function, it did exemplify their importance as immunity deficient strains of *B. subtilis* were rendered defenceless in culture and in biofilm competition assays (Kobayashi, 2021).

Recently, *Listeria monocytogenes* has become a focus of T7SSb research, at least from a bioinformatics standpoint. Bowran and Palmer were the first to show that there are seven EssC variants within *L. monocytogenes* genomes and these variants dictate the downstream repertoire of effector and immunity genes (Bowran & Palmer, 2021). This study suggests that most *L. monocytogenes* T7SSb operons include an LXG toxin and multiple immunity genes. Indeed, these genes seem to be of various families, suggestive of a single operon that encodes immunity for several different toxins. The authors find that some probable immunity proteins fall into conserved families including SUKH-1, DUF1851, DUF1871, immunity protein 74, immunity protein 70, and the cysteine rich

CPCC superfamily of proteins, amongst various unknown domains. This work was suggestive of a role for T7SSb in interbacterial competition and showed that immunity genes of seemingly different functions cluster together in *L. monocytogenes* (Bowran & Palmer, 2021).

In an unpublished preprint, Garrett et al. suggests a plausible mechanism that could lead to the high degree of variability between immunity homologs (Garrett et al., 2022). They find that homologs of the *esaD* inhibiting protein *esaG* have three blocks of high sequence homology and these three sites facilitate extensive recombination among homologous genes. In the case of *esaD* homologs, it seems that this recombination leads to loss of *esaD* genes. Similarly, *tipC* homologs also have two blocks of high sequence homology, and the researchers suggest that this homology leads to a high degree of recombination in *Streptococcus*. The researchers conclude that this high degree of recombination in immunity genes likely drives the evolution of the T7SSb in *Staphylococcus* and *Streptococcus* (Garrett et al., 2022).

Future directions

In general, research into the function of immunity proteins comes through structure-function studies of toxin-immunity pairs. Prototypical effector-immunity studies have been highly impactful in T6SS research and as the T7SSb field matures, it is likely that these studies will become more prominent (Whitney et al., 2015; Ting et al., 2018; Ahmad et al., 2019). In Chapter IV we use TelD, a novel LXG effector from *S. intermedius* GC1825 as a model for our research on chaperones. Although we proved that

TelD is a secreted antibacterial toxin, there is still much to learn about its mechanism of action (Klein et al., 2022). Furthermore, the TelD-specific immunity protein, which we call TipD represents a possible avenue for future research into T7SSb immunity proteins. Since the TelD toxin domain and TipD protein express only sparingly in an *E. coli* overexpression system (unpublished finding) it is unlikely that TelD-TipD represents a viable model for understanding effector immunity interactions from a structural perspective. Instead, more soluble toxin-immunity pairs, especially the many novel pairs from *Bacillus*, *Listeria*, and more diverse *Streptococci*, could provide better models for understanding the structure-function dynamics of effector-immunity interactions. To date, all characterized T7SSb immunity proteins inhibit toxicity by binding directly to their cognate toxin and, in general, this is true for other interbacterial antagonism systems such as T6SS and CDI. However, the recently discovered T6SS immunity protein Tri1 from *Serratia proteamaculans* was found to inhibit toxicity through both direct binding and catalyzing the opposite enzymatic reaction of its cognate toxin (Ting et al., 2018). Effector-immunity studies have been some of the most impactful works of the T6SS field as it was recently demonstrated that T6SS-containing cells could be used to target and remove specific bacteria from mixed populations (Ting et al., 2020). Furthermore, a modified T6SS DNA deaminase toxin was recently implicated as a possible method for targeted mutation of mitochondrial DNA with implications in both mitochondrial research and medicine (Mok et al., 2020; de Moraes et al., 2021). It is not yet clear if research into T7SSb toxin-immunity pairs will be as fruitful as T6SS pairs but since such

exciting applications are being explored in the context of T6SS, it is likely that this trend will continue into T7SS research.

Chapter III summary and discussion

Current understandings of the large T7SSb apparatus protein EsaA

In Chapter III, I discuss my structural and biochemical work on the poorly characterized T7SSb apparatus protein EsaA. In this work, I show that EsaA, despite its initial designation as an "accessory" protein of the T7SSb, is indeed essential for the secretion of both LXG and WXG effectors (Klein et al., 2021). Through subcellular fractionation and a maleimide dye-based labelling assay, I show that EsaA is a membrane protein with a large soluble domain that faces the extracellular side of the cell. The crystal structure of *Sg*EsaA suggests that this large soluble domain forms a dimer that extends at least 200Å beyond the membrane, although, the full-length of EsaA likely extends much further since we were unable to crystallize the domain as a whole. I go on to show that the EsaA dimers observed *in crystallo* were relevant *in vitro* and *in vivo*. EsaA's propensity to dimerize is consistent with the other T7SSb components that have been studied structurally in isolation (Klein et al., 2021).

Since the publication of my findings on EsaA, another group submitted a preprint containing similar structural analysis for the soluble domain of EsaA from *S. aureus* (*Sa*EsaA). In this work, Mietrach et al. present a 3.8Å structure of a protease-resistant fragment of *Sa*EsaA and find that it also forms an extended dimer (Mietrach et al.,

2020a). Similar to our analysis of *Si*EsaA, they suggest that the length and topology of EsaA is suggestive of a role in transporting effectors through the thick peptidoglycan layer of Firmicutes bacteria. The soluble domains of *Sg*EsaA and *Sa*EsaA both have two α-helical domains followed by a β-sheet domain (Klein et al., 2021; Mietrach et al., 2020a). Although a direct one-to-one comparison of the two proteins is not possible since the *Sa*EsaA structure is not published, the two structures look somewhat distinct from one another besides their similar secondary structure. While *Sg*EsaA is very linear, the structure of *Sa*EsaA has a distinct kink between the first and second α-helical domains. It is possible that this difference in the two structures is an artifact of crystal packing or it may represent unique differences between streptococcal and staphylococcal EsaA. One of the more surprising findings of the unpublished preprint is that the β-domain of *Sa*EsaA shares structural similarity with bacterial lectins and human integrins (Mietrach et al., 2020a). Both of these families of proteins have a role in cell-to-cell adhesion and, since this outermost domain of EsaA is likely surface exposed, it is suggestive of the possibility that EsaA may mediate interbacterial adhesion as well. The authors further suggest that, in conjunction with EsaA facilitating effector export, the protein also has a direct role in the growth inhibition of competitors. To support this notion, the researchers present data that suggests a strain lacking EssC inhibits competitor growth more similarly to wild-type cells than an *esaA* deletion strain. They further present data that shows that the purified extracellular portion of EsaA can damage membranes (Mietrach et al., 2020a). It is important to note that I could not replicate these findings using biofilm assays and *E. coli* growth curves (data not shown). Although my research on *Sg*EsaA/*Si*EsaA does not

corroborate the cell-to-cell adhesion and membrane damaging phenotypes suggested by Mietrach et al., it is possible that we missed these findings in our work or that EsaA proteins from *Streptococcus* and *Staphylococcus* behave differently. Minimally, the current research is in agreement that EsaA is a membrane protein with a large extracellular domain that homodimerizes and is required for the export of T7SSb effectors. Future research will be needed to determine if this protein also facilitates cell-to-cell adhesion and membrane damage and if this damage leads to effector import into targeted cells.

Future directions

The work presented in Chapter III, in conjunction with the work done by Mietrach et al., suggests a critical role for EsaA in type VIIb secretion. Building on our understanding of the exact role that this protein plays will probably require a more holistic understanding of the structure and function of the T7SSb, and this understanding will likely come through cryo-EM studies on the T7SSb apparatus. Recently, several cryo-EM studies of the T7SSa apparatus have elucidated the structure of the apparatus and facilitated the understanding of various mechanistic details (Famelis et al., 2019; Poweleit et al., 2019; Bunduc et al., 2021; Beckham et al., 2021). The collection of available T7SSa macrostructures have led to the consensus that the central pore of the apparatus is formed by hexameric EccC and therefore it is probable that the EccC homolog, EssC, also forms this central pore for the T7SSb. EccC also directly interacts with effectors, suggesting a role in both effector recruitment and export (Rosenberg et al.,

2015; Mietrach et al., 2020b). It would therefore also be logical to assume that T7SSb effectors are similarly recruited by EssC but an unpublished preprint recently asserted that EssB, rather than EssC, directly interacts with LXG proteins (Tassinari et al., 2020). This new data defines a larger role for EssB than was previously appreciated and may suggest a multi-step mechanism for effector recruitment. The recent cryo-EM maps of T7SSa also showed that the EccC pore, through which effectors are likely exported, extends into the periplasmic space through EccB which is the only protein that has a periplasmic domain (Bunduc et al., 2021). This finding aligns with our hypothesis for EsaA, which is that EsaA is the protein that forms a conduit through which effector secretion occurs for T7SSb apparatuses (Klein et al., 2021). It is possible that the stark difference in cell envelope architecture between Actinobacteria and Firmicutes has led to the evolution of two divergent T7SSs. In the case of Mycobacteria, which have a thin peptidoglycan layer and a mycomembrane, the T7SSa can secrete an arsenal of mostly small single-domain effectors into the periplasm. Some of these effectors diffuse through the cell wall and oligomerize into outer-membrane pores that then facilitate the export of other effectors into the extracellular space. In the case of Firmicutes, which have no outer membrane and a thick peptidoglycan layer, the T7SSb requires a large cell wall-spanning conduit formed by EsaA to facilitate export of the large toxin-containing LXG effectors. Since no macrostructure of the T7SSb apparatus has yet been solved, these ideas are speculative and can probably only be unequivocally proven through cryo-EM.

Ultimately, I am hopeful that my EsaA crystal structure can be used as a high-resolution

model that is docked into lower resolution maps of the T7SSb macrostructure to enable accurate structure determination of an intact T7SSb apparatus.

Chapter IV summary and discussion

Current understandings of T7SSb effector recognition and chaperones

My research in Chapter IV focuses on the concept of effector recognition, which is the process by which the T7SSb apparatus recognizes and recruits its effectors for export. In this work, I characterize a chaperone-co-chaperone pair that directly interact with a specific LXG effector (Klein et al., 2022). The chaperone pair are part of two conserved DUF families, DUF3130 and DUF3958, which we rename LXG-associated α-helical protein 1 and 2 (Lap1 and Lap2), respectively. LXG operon synteny is highly conserved with *lap1* and *lap2* genes being typically encoded immediately upstream of the LXG effector gene. Lap1 and Lap2 both directly interact with the N-terminal LXG domain of T7SSb effectors to form a heterotrimeric complex that we refer to as the pre-secretion complex. Lap1 and Lap2 are required for the secretion of their cognate LXG effector and in some cases, they also appear to stabilize intrinsically unstable toxin effectors (Whitney et al., 2017; Klein et al., 2022). Since Lap proteins are not secreted along with their LXG effector, my data suggests that they function as secretion chaperones rather than co-secreted effectors. Structural analysis of Lap1 and Lap2 indicates that these small proteins are α-helical in nature and are reminiscent of canonical WXG100 effectors. While Lap2 proteins show very little sequence conservation in

general, Lap1 proteins possess a conserved FxxxD motif at their C-terminus. This FxxxD motif is highly similar to the YxxxD/E motif of T7SSa effectors, which is a well-known determinant of secretion and is necessary for the export of T7SSa effector pairs. Through experiments on FxxxD mutant Lap1 variants, we find that this motif is critical for the secretion of the corresponding LXG effector despite having no influence on heterotrimer formation (Klein et al., 2022). Ultimately, this work is the first in-depth exploration of the determinants that underlie LXG effector secretion, however, more work is needed to fully understand the complexities and unique requirements of secretion for this important family of toxins.

The best studied example of a T7SSb effector chaperone pair is that of EsaD:EsaE. EsaD is an effector of *S. aureus* that possesses a C-terminal nuclease domain that has been linked to both interbacterial competition and virulence (Cao et al., 2016; Ohr et al., 2017). Secretion of EsaD requires the EsaE chaperone which specifically interacts with the N-terminal domain of EsaD (Cao et al., 2016; Anderson et al., 2017). EsaE was also shown to interact with the T7SSb motor ATPase EssC, which suggests a role in guiding the EsaD effector to the T7SSb apparatus (Cao et al., 2016). There are some distinct differences between the EsaE chaperone and the Lap1/2 chaperone-co-chaperone pair described in my work. First, the toxin proteins with which they interact are of different classes of effectors and it is therefore possible that EsaD-type T7SSb effectors are recruited via a different molecular mechanism than that of the canonical LXG effectors. Second, Lap1 and Lap2 have been shown or predicted to be helix-turn-helix proteins similar to canonical WXG100 proteins. Although no structure of EsaE has

yet been solved, the protein is predicted to form a globular mixed α/β fold (data not shown). This structural difference may suggest a functional difference between the two recruitment mechanisms. It is important to note that we cannot currently rule out the possibility that LXG effectors also require a globular chaperone in addition to Lap1 and Lap2, nor can we rule out the possibility that non-LXG effectors, such as EsaD, require α-helical chaperones. Finally, unlike Lap1 and Lap2, EsaE has been shown to directly interact with EssC (Cao et al., 2016). Directly linking LXG effectors to the T7SSb apparatus via the Lap1-Lap2 chaperone-co-chaperone pair is a critical part of this project's future directions (see below).

Although the discovery that pairs of small α-helical proteins could act as a chaperone-co-chaperone pairs by forming a heterotrimeric complex with a cognate LXG effector is a novel finding, the hetero-oligomerization of α-helical proteins is a well-established phenomenon for the T7SS field (Poulsen et al., 2014; Strong et al., 2006; Renshaw et al., 2005). The most widespread family of T7SS effectors, the WXG proteins, exist as obligate hetero- or homodimers and are secreted as a pair (Brodin et al., 2005; Renshaw et al., 2005). PE and PPE effectors are similarly secreted as a heterodimeric pair in all tested instances (Strong et al., 2006; Bottai & Brosch, 2009). Although there is no solved structure of an LXG domain, its interaction with Lap1 and Lap2 along with its predicted α-helical structure suggests that LXG domains also exist as α-helical bundles that interact with other α-helical proteins to facilitate their secretion by the T7SS. While the α-helical nature of T7SS effectors is well-established, the significance of this ubiquitous pattern is not known, although the propensity of some of these effectors to

oligomerize into α-helical membrane pores is an intriguing observation (Piton et al., 2020; Tak et al., 2021). Another interesting parallel between T7SSa and T7SSb effector recognition is in the role of secretion motifs. Data on T7SSa effectors now suggests that recruitment of WXG and PE/PPE effectors requires a bipartite secretion signal. This signal consists of both the WXG motif in the turn region of one α-helical effector (e.g. EsxA) and the unstructured tail and YxxxD/E motif of its heterodimeric partner (e.g.EsxB) (Champion et al., 2006, Daleke et al., 2012a). Crystal structures of both WXG and PE/PPE pairs show that these two motifs exist in close proximity with one another and may therefore form a three-dimensional recognition signal that is recognized by the EccC translocase (Daleke et al., 2012a; Korotkova et al., 2014). Our characterization of Lap1 found that these proteins encode a FxxxD motif in their C-terminal helix (Klein et al., 2022). Although we have shown that this motif is necessary for LXG effector export, future research is required to elucidate if this motif, in conjunction with the LXG motif of its cognate LXG effector, forms a similar bipartite signal that is recognized by the T7SSb apparatus.

Future directions

In Chapter IV, I propose a three-step mechanism for LXG secretion in which the required steps are: (1) formation of the heterotrimeric pre-secretion complex, (2) recognition of the pre-secretion complex by EssC, and (3) effector export by the T7SSb apparatus (Klein et al., 2022). Thus far, our data are supportive of the first step of this mechanism as we have shown that LXG domains indeed interact with Lap1 and Lap2

proteins and that this chaperone-co-chaperone pair is necessary for effector secretion. In the short term, more work is required to better understand the dynamics of this complex. A co-crystal structure of the LXG-Lap1-Lap2 heterotrimer is needed to explain how these three proteins interact and could help explain why two different chaperone protein families are required for the export of a single effector. For example, it is not yet known whether Lap1 and Lap2 bind each other before interacting with the LXG domain or if each of the chaperones interact separately with different regions of the LXG domain. Furthermore, a co-crystal structure would yield useful information on the location of the FxxxD motif of Lap1 in the context of the entire pre-secretion complex and would likely inform on regions of the LXG domain and Lap2 that contribute to a three-dimensional T7SSb effector recognition motif.

After more detailed structural analyses of the LXG-Lap1-Lap2 pre-secretion complex are done, the next step in understanding T7SSb secretion is in determining how the system recognizes LXG effectors. In this regard, determination of recognition motifs on the effectors and chaperones is only half of the story. The precedent from T7SSa suggests that the ATPase EssC is the most likely apparatus protein to recruit effectors (Rosenberg et al., 2015; Mietrach et al., 2020b). Despite this assumption, Tassinari et al. recently demonstrated that EssB interacts with LXG proteins *in vitro* (Tassinari et al., 2020). This finding may suggest that the T7SSa and the T7SSb recruit effectors through different mechanisms or it may suggest the existence of an effector "handoff" from one structural component of the T7SSb to another. Regardless of which protein(s) bind to effectors and how, the work presented in Chapter IV provides thought provoking insights

into how to test effector recognition *in vitro* and *in vivo*. Ultimately, studies on effector recognition will lead to questions of how the effectors are then secreted through the apparatus. With regards to our hypothesis that EsaA acts as a T7SSb protein export conduit, more definitive insights into secretion mechanism will likely be best understood through cryo-EM studies on the T7SSb as a whole. Whether or not these studies can also provide information on how effectors then enter competitor or host cells will also be critical for a full mechanistic understanding of T7SSb function.

Concluding remarks

During my doctoral work, I have attempted to further our current understanding of the structure and function of the T7SSb primarily through X-ray crystallography and protein biochemistry. Using TipC1 as a model, I explored how T7SSb immunity proteins interact with and protect against antibacterial toxins (Klein et al., 2018). By solving the first structure of the T7SSb apparatus component EsaA, I generated a hypothesis for how the T7SSb can transport effectors through the cell membrane and cell wall in a single step (Klein et al., 2021). Finally, through the discovery and characterization of Lap1 and Lap2 as a model, I studied how this novel chaperone-co-chaperone pair interacts with the LXG domain of T7SSb effectors in the first required step of effector secretion (Klein et al., 2022). Despite these advances, my work also leaves many questions to be answered by current and future members of the Whitney lab as well as the T7SSb field at large. Do all T7SSb immunity proteins inhibit toxin activity through a direct protein-protein

interaction? What is the full set of T7SSb toxin activities and do all of these activities require an immunity determinant? Would this also be true of effectors that target host cells rather than competitor bacteria? What is the macrostructure of the T7SSb and does this structure suggest that T7SSa and T7SSb function similarly? What is the precise role of EsaA in T7SSb secretion and how does its structure support this role? How do Lap1 and Lap2 interact with the LXG domain of T7SSb effectors and how does this interaction support effector recognition? Which component of T7SSb recognizes LXG effectors and is this different from effector recognition of T7SSa? These questions represent several important gaps in our understanding of T7SSb biology. Although much more work is needed to answer these and other questions that remain unanswered, I hope that the research conducted during my PhD will provide a useful starting point from which these ideas can be explored.

Appendix

Chapter II supplement

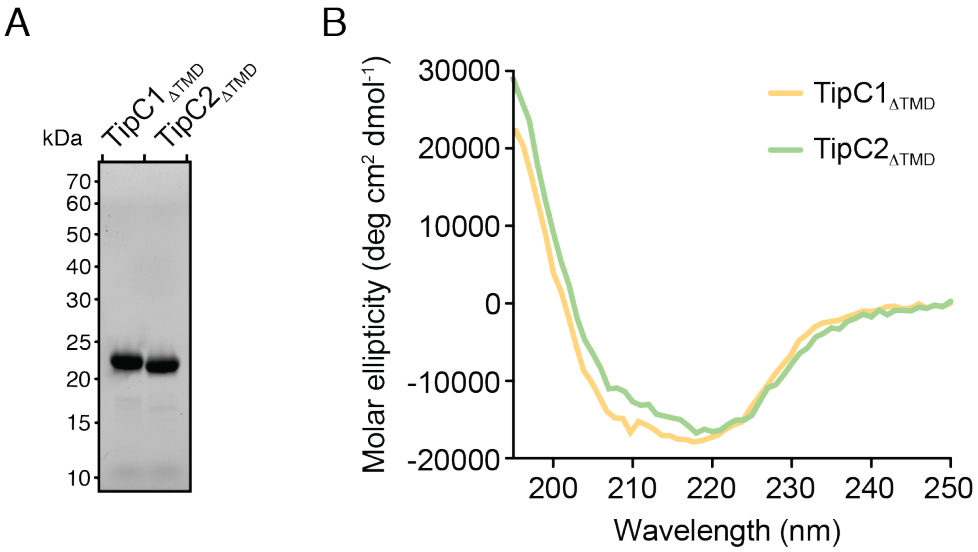


Figure S2.1: TipC1$_{\Delta TMD}$ and TipC2$_{\Delta TMD}$ are comprised of highly similar secondary structure elements. (A) SDS-PAGE analysis of purified TipC1$_{\Delta TMD}$ and TipC2$_{\Delta TMD}$ used for circular dichroism analysis. (B) Far-UV circular dichroism spectra of TipC1$_{\Delta TMD}$ and TipC2$_{\Delta TMD}$.

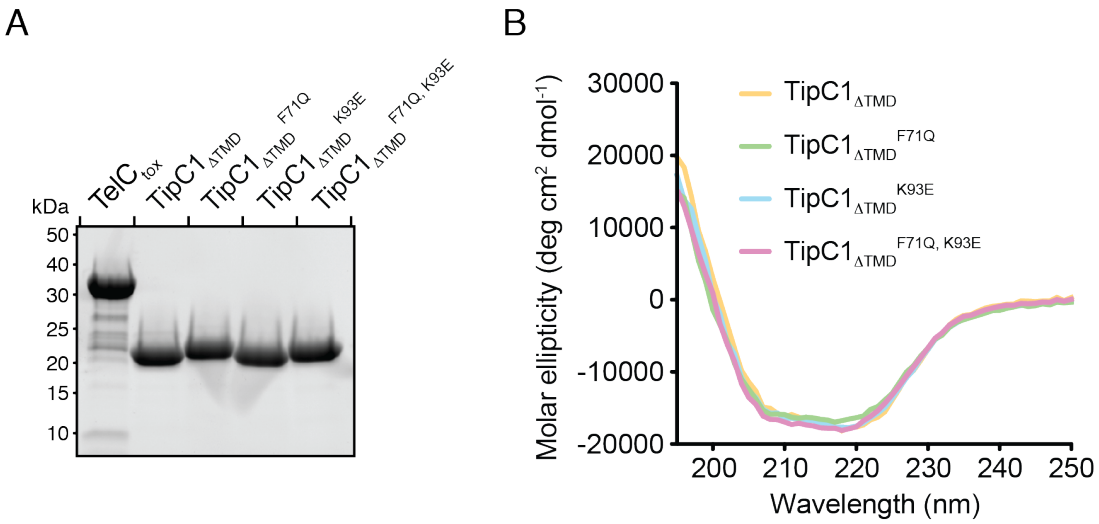


Figure S2.2: TipC1$_{\Delta TMD}$ and the indicated TipC1$_{\Delta TMD}$ site-specific variants are comprised of highly similar secondary structure elements. (A) SDS-PAGE analysis of purified TelC$_{tox}$, TipC1$_{\Delta TMD}$ and the indicated TipC1$_{\Delta TMD}$ site-specific variants used for lipid II phosphatase assays. (B) Far-UV circular dichroism spectra of TipC1$_{\Delta TMD}$ and the indicated TipC1$_{\Delta TMD}$ site-specific variants.

Table S2.1: Strains used in chapter II

Organism	Genotype	Description	Reference
S. intermedius B196	wild-type		Olson et al., 2013
	ΔSIR_0175 ::kan^R	*essC* deletion strain	Whitney et al., 2017
	ΔSIR_01486 ΔSIR_01487 ΔSIR_01488 ::kan^R	*tipC1*, SIR_1487, *tipC2* deletion strain	This study
	ΔSIR_01486 ΔSIR_01487 ΔSIR_01488 ΔSIR_01489 ::kan^R	*telC*, *tipC1*, SIR_1487, *tipC2* deletion strain	This study
S. gallolyticus ATCC 43143	wild-type		Schlegel et al., 2003
E. coli XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lacI*^{q} Z*ΔM15* Tn*10* (Tet^R)]	Cloning strain	Agilent
E. coli DH5α	F$^-$ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG* Φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *hsdR17*(r_K^- m_K^+), λ–	Cloning strain	Novagen
E. coli BTH101	F$^-$, *cya-99*, *araD139*, *galE15*, *galK16*, *rpsL1* (Str^R), *hsdR2*, *mcrA1*, *mcrB1*	Bacterial two-hybrid strain	Euromedex
E. coli BL21 (DE3) CodonPlus	F$^-$ *ompT gal dcm lon* *hsdS$_B$*(r_B^- m_B^-) λ(DE3) pLysS(Cm^R)	Protein expression strain	Novagen

Table S2.2: Plasmids used in chapter II

Plasmid	Relevant features	Reference
pDL277	*Streptococcus-E. coli* shuttle vector, $Spec^R$	Aspiras et al., 2000
pKNT25	B2H expression vector with *plac*, Kan^R, C-terminal fusion to T25 fragment of CyaA	Euromedex
pUT18C	B2H expression vector with *plac*, Amp^R, C-terminal fusion to T18 fragment of CyaA	Euromedex
pETDuet-1	Co-expression vector with *lacI*, T7 promoter, N-terminal His_6 tag in MCS-1, Amp^R	Novagen
pET29b	Expression vector with *lacI*, T7 promoter, C-terminal His_6 tag, Kan^R	Novagen
pDL277::P96_ss-SIR_1489_202-552	*S. intermedius* expression vector for residues 202-552 of TelC fused to a sec signal sequence (ss-$TelC_{tox}$)	Whitney et al., 2017
pDL277::P96_ss-SIR1489_202-552_D401A	*S. intermedius* expression vector for ss-$TelC_{tox}^{D401A}$	Whitney et al., 2017
pDL277::P96_ss-SIR1489_202-552–SIR1488	*S. intermedius* expression vector for ss-$TelC_{tox}$ and TipC1	Whitney et al., 2017
pDL277::P96_ss-SIR1489_202-552–SIR1486	*S. intermedius* expression vector for ss-$TelC_{tox}$ and TipC2	This study
pDL277::P96_ss-SIR1489_202-552–SIR1488_F71Q	*S. intermedius* expression vector for ss-$TelC_{tox}$ and $TipC1^{F71Q}$	This study
pDL277::P96_ss-SIR1489_202-552–SIR1488_K93E	*S. intermedius* expression vector for ss-$TelC_{tox}$ and $TipC1^{K93E}$	This study
pDL277::P96_ss-SIR1489_202-552–SIR1488_F71Q_K93E	*S. intermedius* expression vector for ss-$TelC_{tox}$ and $TipC1^{F71Q, K93E}$	This study
pDL277::P96_SIR_1488-V	*S. intermedius* expression vector for TipC1 fused to a C-terminal VSV-G epitope tag	This study
pDL277::P96_SIR_1488_23-204-V	*S. intermedius* expression vector for residues 23-304 of TipC1 ($TipC1_{\Delta TMD}$) fused to a C-terminal VSV-G epitope tag	This study

pDL277::P96_SIR_1157-V	*S. intermedius* expression vector for SodA fused to a C-terminal VSV-G epitope tag	This study
pDL277::P96_SIR_1047-V	*S. intermedius* expression vector for LsrS fused to a C-terminal VSV-G epitope tag	This study
pDL277::P96_SIR_1489	*S. intermedius* expression vector for TelC	This study
pDL277::P96_ss-SIR_1489_SIR1488	*S. intermedius* expression vector for ss-TelC and TipC1	This study
pKNT25::sgTelC	B2H expression vector for TelC from *S. gallolyticus*	This study
pUT18C::sgTipC1	B2H expression vector for TipC1 from *S. gallolyticus*	This study
pUT18C::sgTipC2	B2H expression vector for TipC2 from *S. gallolyticus*	This study
pUT18C::sgTipC3	B2H expression vector for TipC3 from *S. gallolyticus*	This study
pUT18C::sgTipC4	B2H expression vector for TipC4 from *S. gallolyticus*	This study
pETDuet-1::SIR_1489_202-552	*E. coli* expression vector for $TelC_{tox}$	Whitney et al., 2017
pET29b::SIR_1488_23-204-V	*E. coli* expression vector for $TipC1_{\Delta TMD}$ fused to a C-terminal VSV-G epitope tag ($TipC1_{\Delta TMD}$-V)	This study
pET29b::SIR_1486_23-203-V	*E. coli* expression vector for $TipC2_{\Delta TMD}$ fused to a C-terminal VSV-G epitope tag	This study
pETDuet-1::SIR_1486_23-203	*E. coli* expression vector for $TipC2_{\Delta TMD}$ fused to an N-terminal His_6-tag	This study
pET29b::SIR_1488_23-204_D60R-V	*E. coli* expression vector for $TipC1_{\Delta TMD}$-V D60R variant	This study
pET29b::SIR_1488_23-204_F71Q-V	*E. coli* expression vector for $TipC1_{\Delta TMD}$-V F71Q variant	This study
pET29b::SIR_1488_23-204_S81Q-V	*E. coli* expression vector for $TipC1_{\Delta TMD}$-V S81Q variant	This study
pET29b::SIR_1488_23-204_K93E-V	*E. coli* expression vector for $TipC1_{\Delta TMD}$-V K93E variant	This study
pET29b::SIR_1488_23-204_S100Q-V	*E. coli* expression vector for $TipC1_{\Delta TMD}$-V S100Q variant	This study
pET29b::SIR_1488_23-204_S112Q-V	*E. coli* expression vector for $TipC1_{\Delta TMD}$-V S112Q variant	This study

pET29b::SIR_1488_23-204_S114Q-V	*E. coli* expression vector for TipC1$_{\Delta TMD}$-V S114Q variant	This study
pET29b::SIR_1488_23-204_K160E-V	*E. coli* expression vector for TipC1$_{\Delta TMD}$-V K160E variant	This study
pET29b::SIR_1488_23-204_K168E-V	*E. coli* expression vector for TipC1$_{\Delta TMD}$-V K168E variant	This study
pET29b::SIR_1488_23-204_K185E-V	*E. coli* expression vector for TipC1$_{\Delta TMD}$-V K185E variant	This study
pET29b:: SIR_1488_23-203	*E. coli* expression vector for TipC2$_{\Delta TMD}$ fused to a C-terminal His$_6$-tag	This study
pET29b:: SIR_1488_23-204	*E. coli* expression vector for TipC1$_{\Delta TMD}$ fused to a C-terminal His$_6$-tag	This study
pET29b:: SIR_1488_23-204_F71Q	*E. coli* expression vector for TipC1$_{\Delta TMD}$-his$_6$ F71Q variant	This study
pET29b:: SIR_1488_23-204_K93E	*E. coli* expression vector for TipC1$_{\Delta TMD}$-his$_6$ K93E variant	This study
pET29b:: SIR_1488_23-204_F71Q	*E. coli* expression vector for TipC1$_{\Delta TMD}$-his6 F71Q, K93E variant	This study

Chapter III supplement

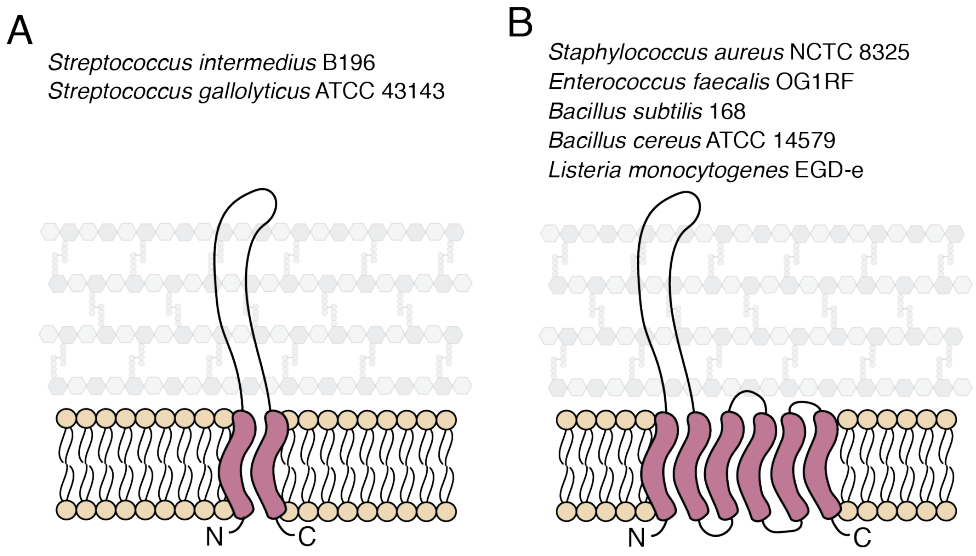


Figure S3.1: Schematic depicting the two common predicted membrane topologies of EsaA. (A-B) EsaA proteins typically have one N-terminal and one C-terminal TMD (A) or one N-terminal and five C-terminal TMDs (B). TMDs predicted by TMHMM are depicted in red (Krogh et al., 2001). Several representative strains of Firmicutes bacteria are listed for each topology.

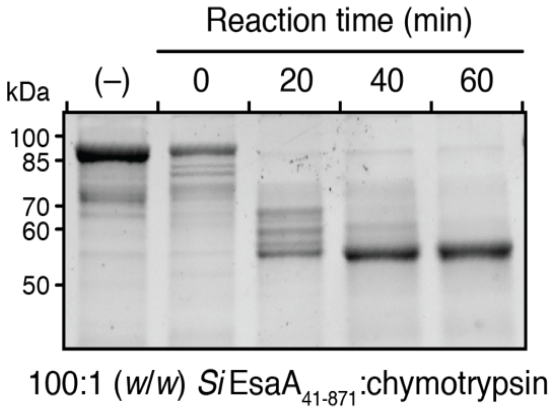
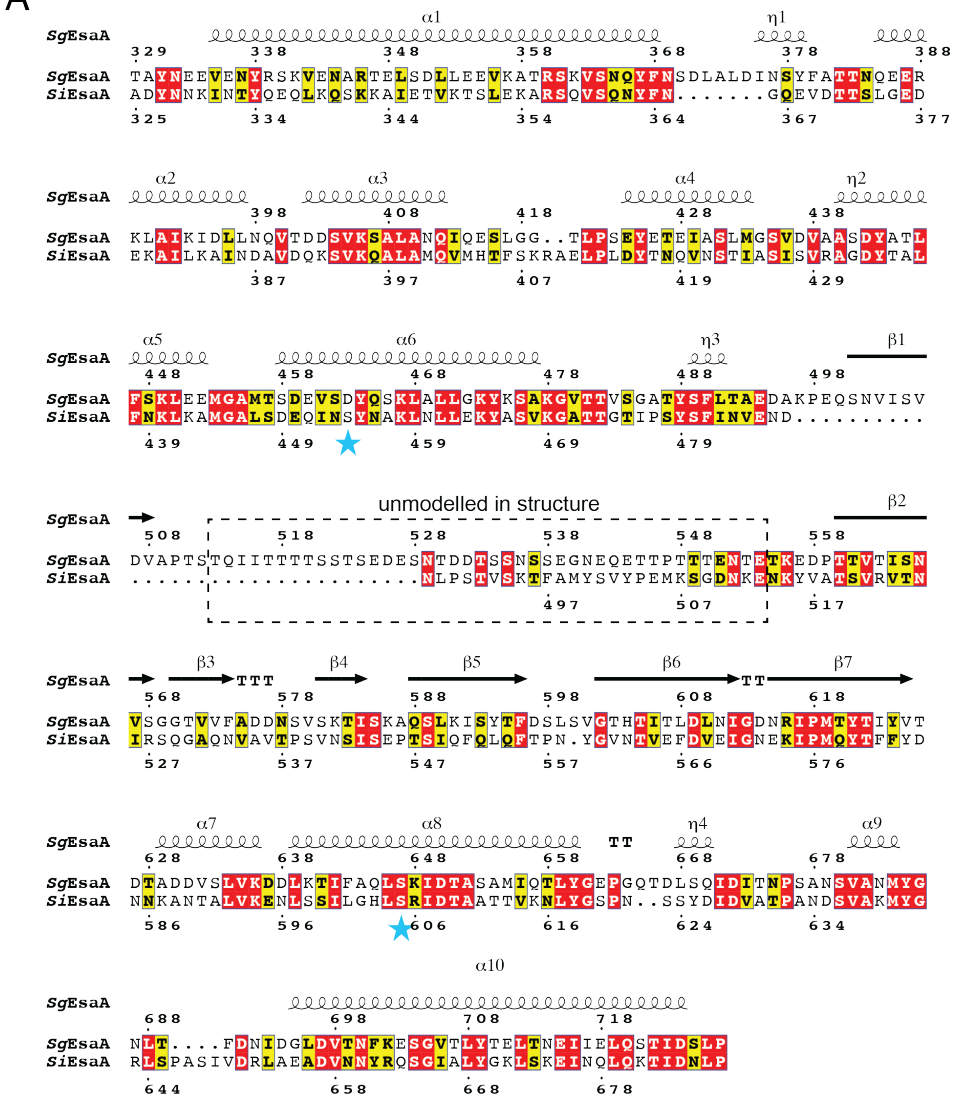


Figure S3.2: Digestion of *Si*EsaA$_{41\text{-}871}$ with chymotrypsin results in a stable truncation of approximately 55kDa. A 1:100 (*w*/*w*) chymotrypsin: *Si*EsaA$_{41\text{-}871}$ digestion was conducted over one hour with samples being taken every 20 minutes. The (–) condition indicates untreated *Si*EsaA$_{41\text{-}871}$. *Si*EsaA$_{41\text{-}871}$ has a predicted molecular weight of 92.5kDa and the amino acid sequence of the 55kDa truncation of *Si*EsaA$_{41\text{-}871}$ was confirmed by liquid chromatography-tandem mass spectrometry.

A


SgEsaA
SiEsaA
α1
η1
α2
α3
α4
η2
α5
α6
η3
β1
unmodelled in structure
β2
β3
β4
β5
β6
β7
α7
α8
η4
α9
α10

B

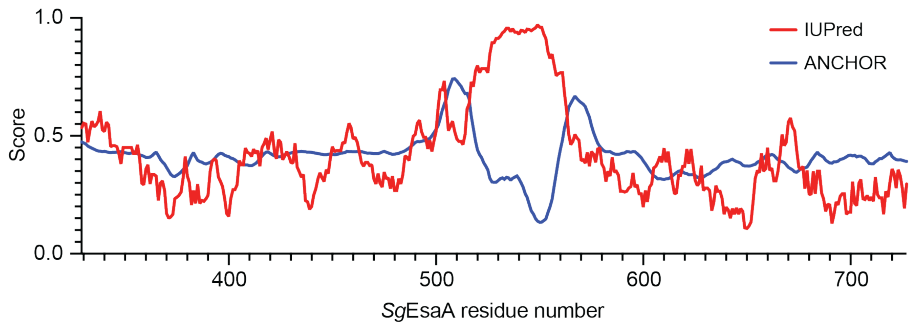

IUPred
ANCHOR
Score
1.0
0.5
0.0
400
500
600
700
SgEsaA residue number

Figure S3.3: The principal difference between *Sg*EsaA and *Si*EsaA is the length of the unmodeled β1-β2 loop. (A) Alignment of *Sg*EsaA$_{329\text{-}727}$ and *Si*EsaA$_{325\text{-}687}$ depicting the secondary structure derived from the *Sg*EsaA$_{329\text{-}727}$ structure was generated using ESPript3 (Robert and Gouet, 2014). The secondary structure of *Sg*EsaA$_{329\text{-}727}$ is shown as squiggles and arrows for α-helices and β-strands, respectively. Residues that are identical between the two sequences are highlighted red whereas similar residues are highlighted yellow. Blue stars denote residues mutated to cysteine for the topology mapping experiment in Figure 3.2. (B) The IUPred2A web server was used to predict disorder across the primary sequence of *Sg*EsaA$_{329\text{-}727}$ (Erdos and Dosztanyi, 2020; Meszaros et al., 2018). The IUPred2 algorithm predicts intrinsically disordered protein regions, while ANCHOR2 predicts disordered regions that are likely stabilized upon interaction with a partner protein. The highest scoring region occurs roughly between residues 498-576, which includes both the unmodelled region in the *Sg*EsaA$_{329\text{-}727}$ structure and the large gap in the alignment between *Sg*EsaA and *Si*EsaA.

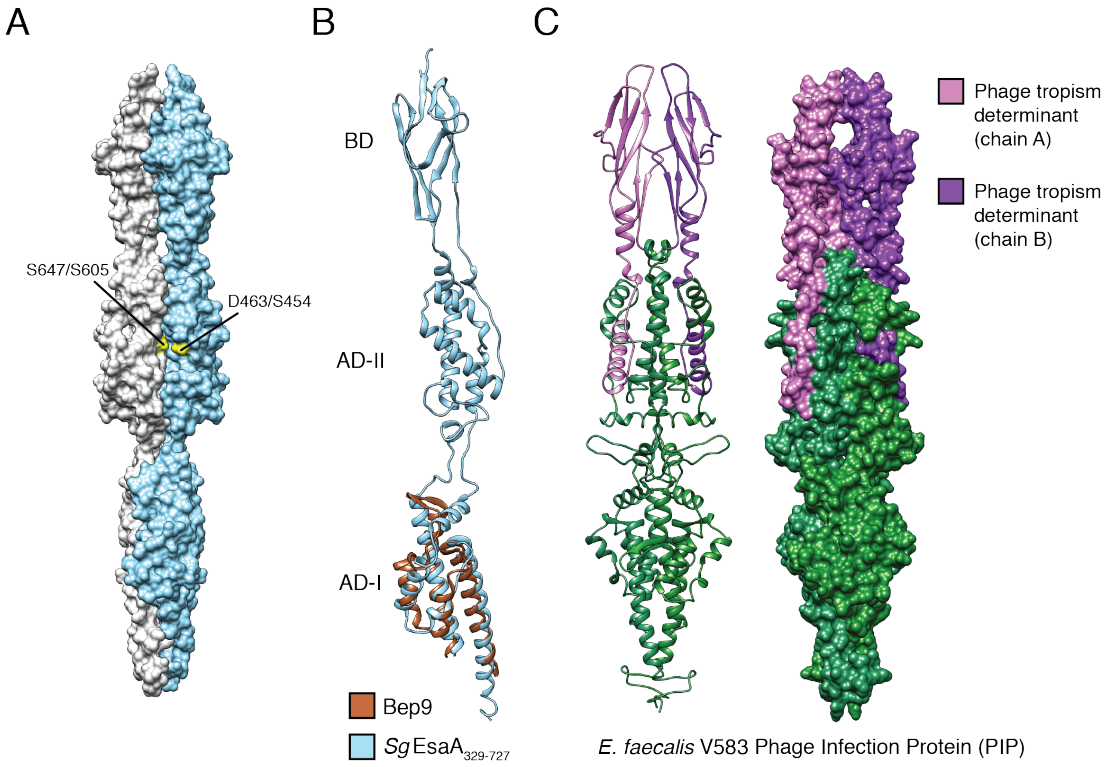


Figure S3.4: Additional structural analyses of *Sg*EsaA$_{329\text{-}727}$. (A) Location of two cysteine mutation sites used for the membrane topology experiment described in Figure 3.2. The wild-type residues for *Sg*EsaA (left) and *Si*EsaA (right) are indicated. (B) The BID domain of Bep9 from *Bartonella clarridgeiae* resembles the AD-I domain of *Sg*EsaA$_{329\text{-}727}$. Bep9 (PDB code 4YK2) is the highest scoring structural homologue for *Sg*EsaA$_{329\text{-}727}$ as determined by DALILITE (Z-score, 8.5; Cα root mean squared deviation of 3.5Å over 100 aligned residues). The structures of Bep9 (orange) and *Sg*EsaA$_{329\text{-}727}$ (blue) were superimposed using UCSF Chimera. (C-D) The structure of *Sg*EsaA$_{329\text{-}727}$ allows for homology modelling of *E. faecalis* V583 PIP. (C) Ribbon and (D) surface diagrams of the structure of the *E. faecalis* V583 PIP protein were generated by the one-to-one threading algorithm of Phyre^2 (Kelley et al., 2015). The phage tropism region,

coloured pink and purple, encompasses the BDs and a short segment of the AD-II domains of the *Sg*EsaA$_{329\text{-}727}$ homodimer.

Table S3.1: Strains used in chapter III

Organism	Genotype	Description	Reference
S. intermedius B196	Wild-type		Olson et al., 2013
S. intermedius B196 *ΔSIR_0175::kanR*	*ΔSIR_0175::kanR*	*essC* deletion	Whitney et al., 2017
S. intermedius B196 *ΔSIR_0176::kanR*	*ΔSIR_0176::kanR*	*esaA deletion*	This study
S. gallolyticus ATCC 43143	Wild-type		Schlegel et al., 2003
E. coli XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lacI*^{q} Z*ΔM15* Tn*10* (Tet^{R})]	Cloning strain	Agilent
E. coli BL21 (DE3) CodonPlus	F^{-} *ompT gal dcm lon hsdS$_{B}$*(r$_{B}^{-}$ m$_{B}^{-}$) λ(DE3) pLysS(Cm^{R})	Protein expression strain	Novagen
E. coli B834 (DE3)	F^{-} *ompT gal dcm hsdS$_{B}$*(r$_{B}^{-}$ m$_{B}^{-}$) λ(DE3) *met*	Protein expression methionine auxotroph.	Novagen

Table S3.2: Plasmids used in chapter III

Plasmid	Relevant features	Reference
pBAV1K	Bryksin et al., 2010	NA
pDL277	Aspiras et al., 2000	NA
pDL277::p96_*esaA*_VSV-G	This study	NA
pDL277::p96_*esaA*_V8C_VSV-G	This study	NA
pDL277::p96_*esaA*_V150C_VSV-G	This study	NA
pDL277::p96_*esaA*_F302C_VSV-G	This study	NA
pDL277::p96_*esaA*_S454C_VSV-G	This study	NA
pDL277::p96_*esaA*_S605C_VSV-G	This study	NA
pDL277::p96_*esaA*_V762C_VSV-G	This study	NA
pDL277::p96_*esaA*_F909C_VSV-G	This study	NA
pDL277::p96_*esaA*_N586C_VSV-G	This study	NA
pDL277::p96_*esaA*_T612C_VSV-G	This study	NA
pDL277::p96_*esaA*_L644C_VSV-G	This study	NA
pET29b	Novagen	#69872-3
pET29b::Si*esaA*	This study	NA
pET29b::Si*esaA*_30-871	This study	NA
pET29b::Si*esaA*_234-790	This study	NA
pET29b::Si*esaA*_328-685	This study	NA
pET29b::Si*esaA*_328-685_N586C	This study	NA
pET29b::Si*esaA*_328-685_T612C	This study	NA
pET29b::Si*esaA*_328-685_L644C	This study	NA
pET29b::Sg*esaA*_235-829	This study	NA
pET29b::Sg*esaA*_332-725	This study	NA
pET29b::Sg*esaA*_332-725_T628C	This study	NA
pET29b::Sg*esaA*_332-725_A654C	This study	NA
pET29b::Sg*esaA*_332-725_L688C	This study	NA
pETDuet-1	Novagen	NA
pETDuet-1::5'*esaA*flank_3'*esaA*flank	This study	NA
pETDuet-1:: 5'EsaAflank_SpecPromoter_*kanR*_ 3'EsaAflank	This study	NA
pETDuet-1::SpecPromoter_*kanR*	This study	NA

Chapter IV supplement

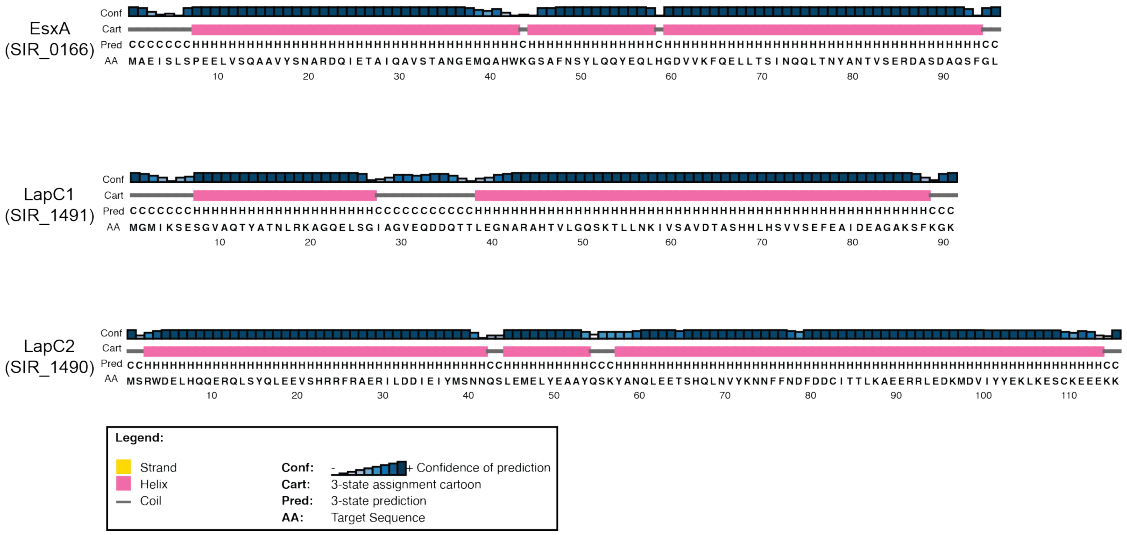


Figure S4.1: Secondary structure predictions for EsxA, LapC1 and LapC2 from *Streptococcus intermedius* B196. Graphical output from PSIPRED 4.0 analyses of EsxA, LapC1 and LapC2. Per-residue secondary structure predictions and confidence scores are indicated above each amino acid in the sequence.

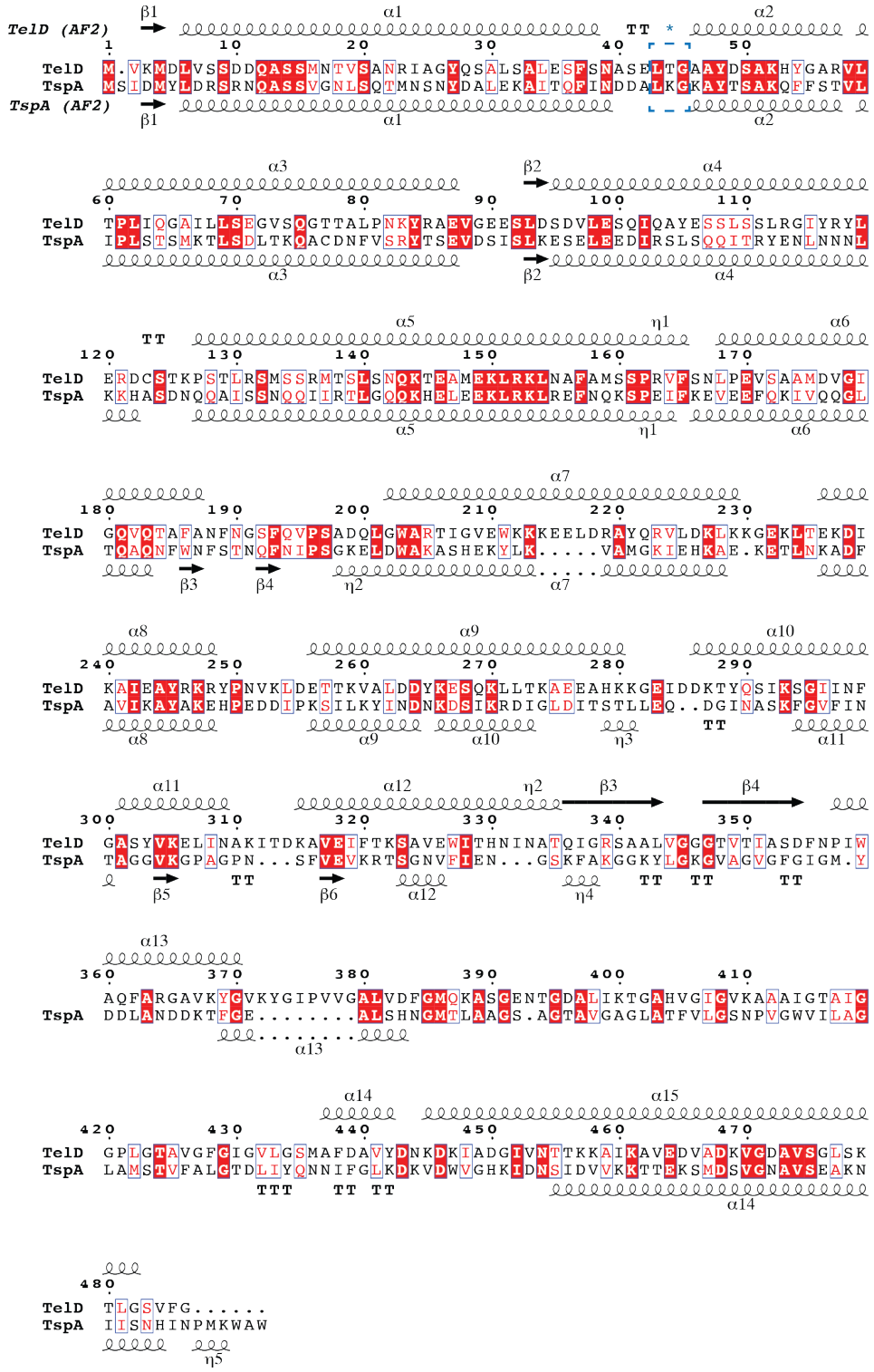

TelD (AF2)
TspA (AF2)
TelD
TspA

Figure S4.2: Sequence and predicted secondary structure alignment of TelD and TspA. Secondary structure assignments are based on AlphaFold 2 predicted tertiary structures. Overall pairwise sequence identity is 23.8%. TelD and TspA have highest levels of sequence homology within their predicted N-terminal LXG domains. Dashed blue box indicates each effector's LXG motif.

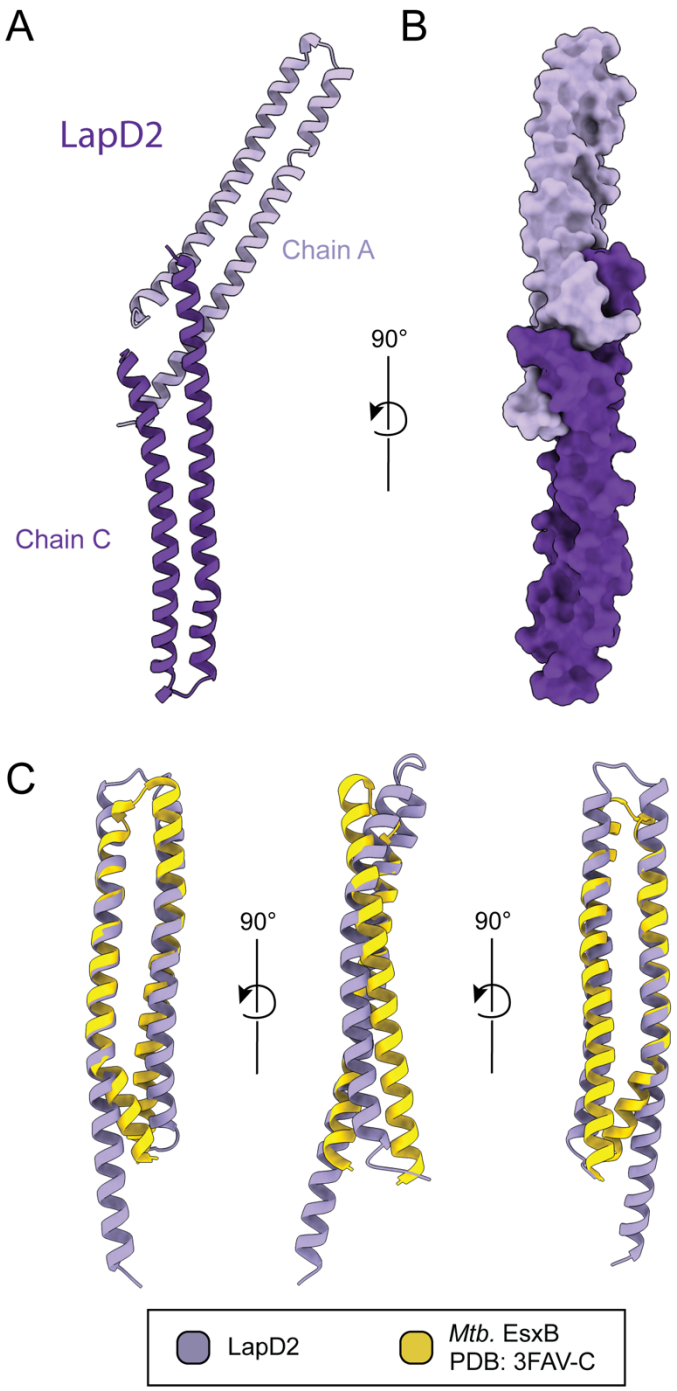

A
B
LapD2
Chain A
90°
Chain C
C
90°
90°
LapD2
Mtb. EsxB
PDB: 3FAV-C

Figure S4.3: Toe-to-Toe packing arrangement of LapD2 and structural alignment of LapD2 to *M. tuberculosis* EsxB. (A-B) LapD2 chains A and C interact with one another in a toe-to-toe manner that involves both N- and C-termini. (C) Structural alignment of LapD2 with *M. tuberculosis* EsxB (PDB code 3FAV) shown in ribbon representation.

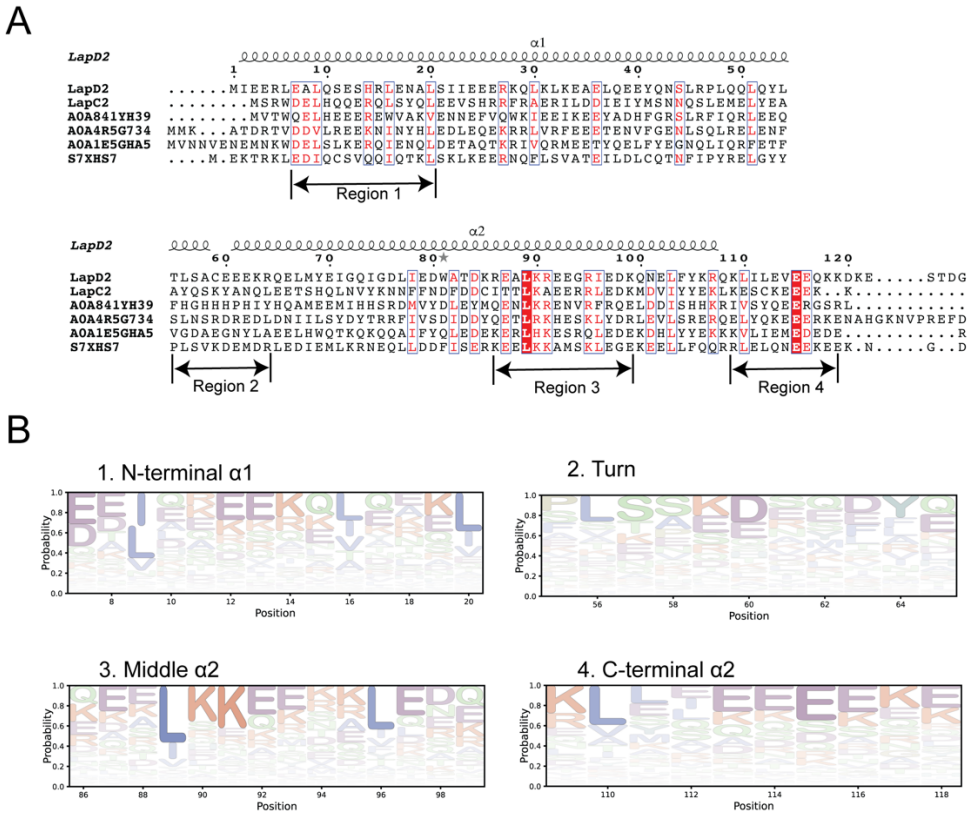


Figure S4.4: Sequence alignment of LapD2, LapC2 and four DUF3958 homologs and sequence logo representation of regions exhibiting modest sequence conservation. (A) Multiple sequence alignment of LapD2 with four randomly selected homologs identified by JackHMMER (UniprotKB accessions listed), and LapC2. (B) Normalized HMM logos generated from the entire JackHMMER sequence hit table reveal a high degree of sequence variability across the group, even in the most conserved regions of the protein (3 and 4).

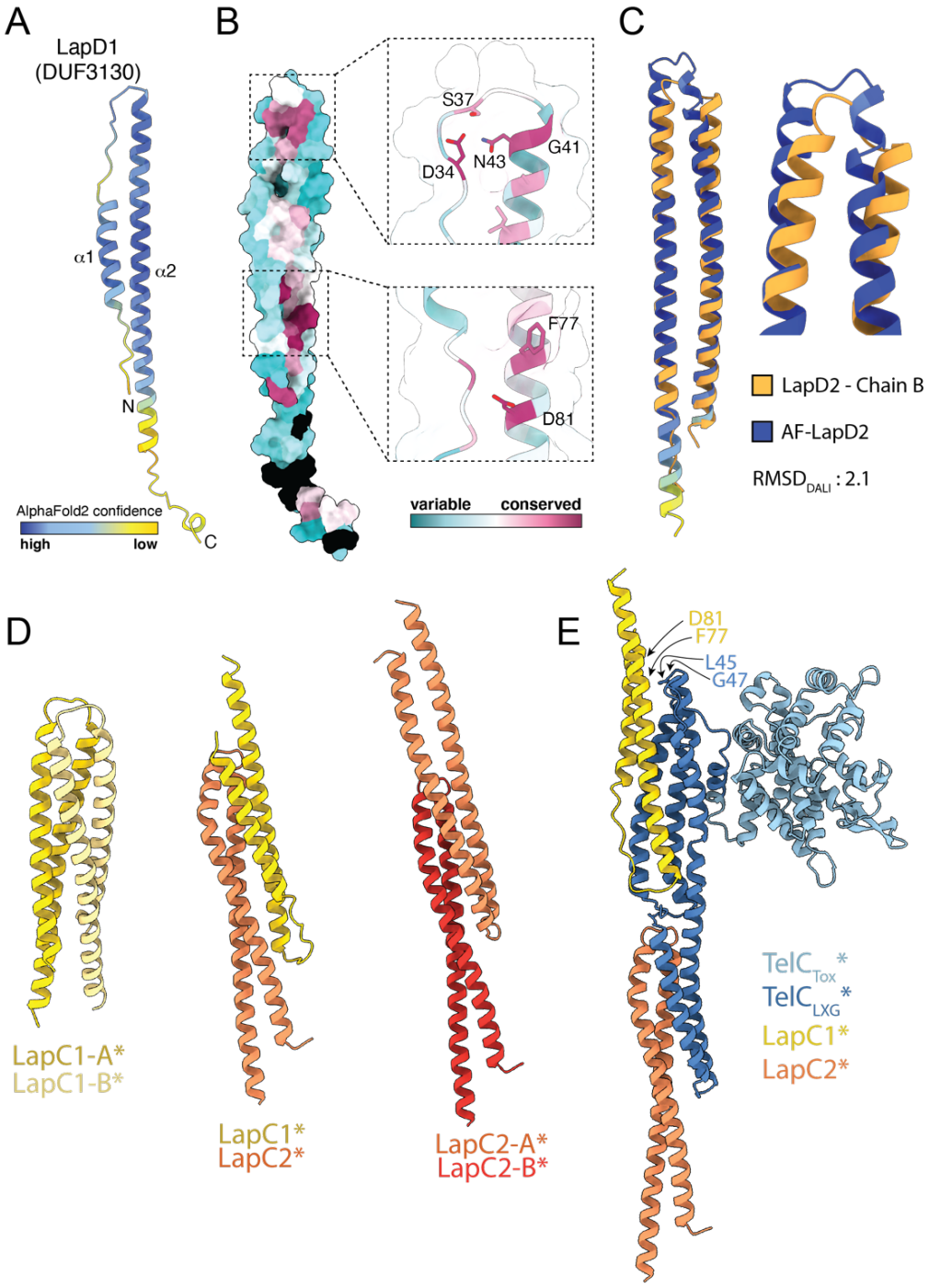


Figure S4.5: AlphaFold2 predicted structure and sequence conservation mapping of LapD1 and comparison of the LapD2 crystal structure to its AlphaFold2 model. (A) AlphaFold2 model of LapD1 coloured by confidence score. (B) Surface representation of

DUF3130 sequence conservation mapped onto the LapD1 predicted structure. (C) LapD2 crystal structure (gold) aligned to the AlphaFold2 predicted model (coloured by confidence score, as in panel A). (D) AlphaFold-multimer models of LapC1 and LapC2 in hypothetical homodimeric (left and right panels, respectively) and heterodimeric (middle) arrangements. (E) AlphaFold-multimer model of the hypothetical TelC-LapC1-LapC2 heterotrimeric complex. The LapC1 FxxxD and the TelC LxG motifs are both highlighted.

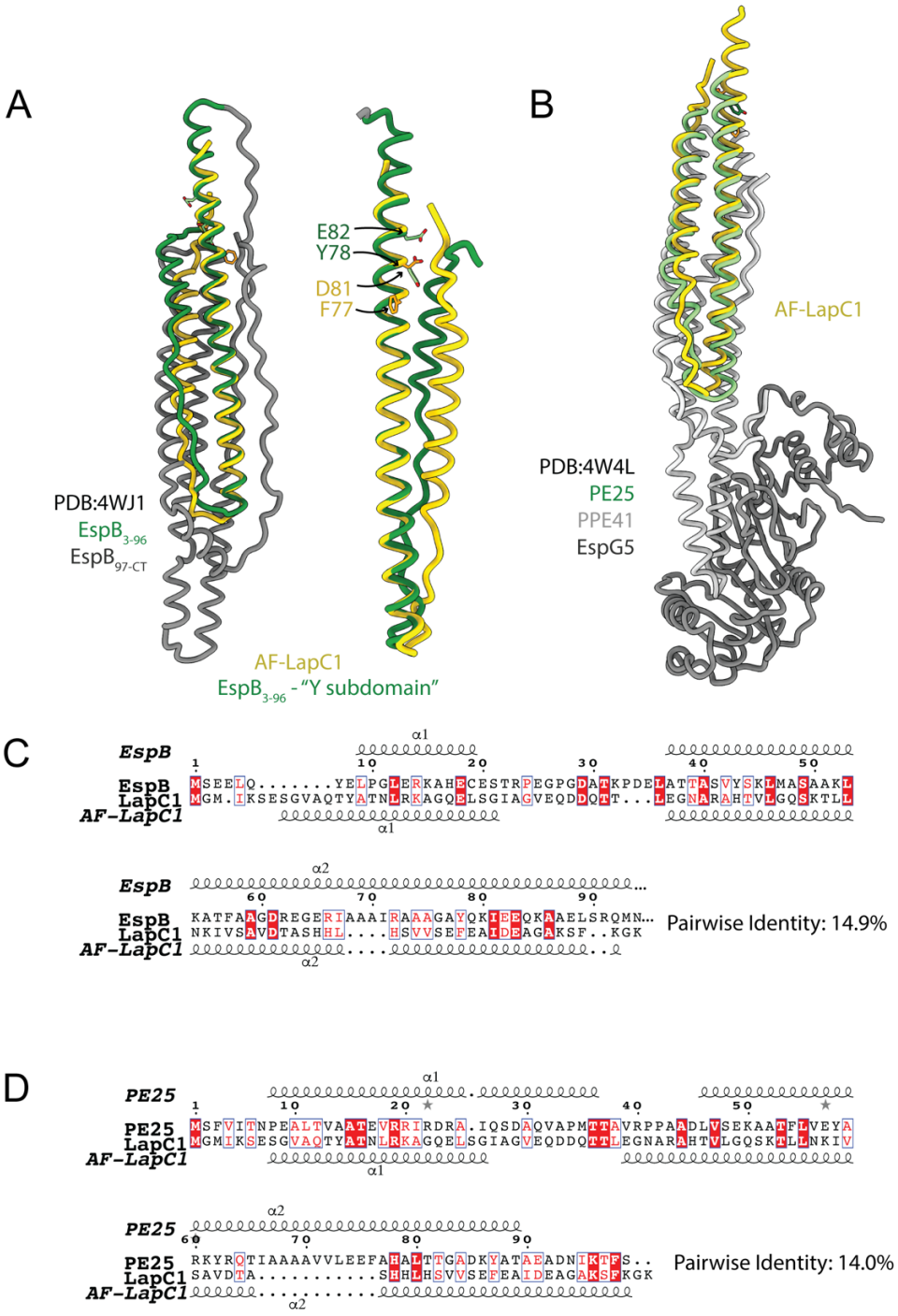


Figure S4.6: AlphaFold2 predicted structure of LapC1 aligned to crystal structures of the Type VIIa substrates EspB and PE25. (A) Predicted structure of LapC1 (yellow) aligned to the Y-subdomain (dark green) of EspB from *M. tuberculosis* (PDB ID: 4WJ1)

reveals a conserved FxxxD motif found in a similar location as the YxxxD/E export motif required for EspB secretion. (B) Structural alignment of predicted LapC1 structure to PE25 when in complex with its cognate PPE41 protein and EspG5 chaperone (PDB ID: 4W4L). (C-D) Pairwise sequence alignments of LapC1 to $EspB_{1-96}$ (C) and PE25 (D).

Table S4.1: Strains used in chapter IV

Organism	Genotype	Description	Reference
S. intermedius B196	Wildtype		Olson et al., 2013
	ΔSIR_1490::kan^R	*lapC2* deletion strain	This study
	ΔSIR_0175::kan^R	*essC* deletion strain	Whitney et al., 2017
	ΔSIR_1489-1486::kan^R	*telC-tipC2* deletion strain	Klein et al., 2018
S. intermedius GC1825	Wildtype		This study
	ΔGC1825_00253::kan^R	*essB* deletion strain	This study
	ΔGC1825_00255::kan^R	*lapD1* deletion strain	This study
	ΔGC1825_00256::kan^R	*lapD2* deletion strain	This study
E. coli XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [*F' proAB lacI^q* Z Δ M15 Tn*10* (Tet^R)]	Cloning strain.	Agilent
E. coli BL21 (DE3) CodonPlus	F$^-$ *ompT gal dcm lon hsdS*$_B$(r_B^- m_B^-) λ(DE3) pLysS(Cm^R)	Protein expression strain.	Novagen
E. coli B834 (DE3)	F$^-$ *ompT gal dcm hsdS*$_B$(r_B^- m_B^-) λ(DE3) *met*	Protein expression methionine auxotroph.	Novagen

Table S4.2: Plasmids used in chapter IV

Plasmid	Relevant features	Reference
pDL277	*Streptococcus-E. coli* shuttle vector, $Spec^R$	Aspiras et al., 2000
pETDuet-1	Co-expression vector with *lacI*, T7 promoter, N-terminal His_6 tag in MCS1, Amp^R	Novagen
pET29b	Expression vector with *lacI*, T7 promoter, C-terminal His_6 tag, Kan^R	Novagen
pSCrhaB2	Expression vector with *PrhaB*, Tmp^R	Cardona & Valvano, 2005
pPSV39	Expression vector with *lacI*, *lacUV5* promoter, Gm^R	Silverman et al., 2013
pDL277::P96_*lapC1*_VSV-G	*S. intermedius* expression vector for LapC1, C-terminal VSV-G tag	This study
pDL277::P96_*lapC2*_VSV-G	*S. intermedius* expression vector for LapC2, C-terminal VSV-G tag	This study
pDL277::P96_*telC*_VSV-G	*S. intermedius* expression vector for TelC, C-terminal VSV-G tag	Klein et al., 2018
pDL277::P96_*essB*$_{GC1825}$_VSV-G	*S. intermedius* expression vector for EssB from strain GC1825, C-terminal VSV-G tag	This study
pDL277::P96_lapD1_VSV-G	*S. intermedius* expression vector for LapD1, C-terminal VSV-G tag	This study
pDL277::P96_*lapD1*$_{F77A}$_VSV-G	*S. intermedius* expression vector for LapD1 with an F77A mutation, C-terminal VSV-G tag	This study
pDL277::P96_*lapD1*$_{D81A}$_VSV-G	*S. intermedius* expression vector for LapD1 with an D81A mutation, C-terminal VSV-G tag	This study
pDL277::P96_*lapD2*_VSV-G	*S. intermedius* expression vector for LapD2, C-terminal VSV-G tag	This study
pDL277::P96_*lapD2*$_{C59S}$_VSV-G	*S. intermedius* expression vector for LapD2 with an C59S mutation, C-terminal VSV-G tag	This study

pETDuet-1::*telC*$_{LXG}$_His$_6$::*lapC1*	*E. coli* co-expression vector for the LXG domain of TelC with LapC1, C-terminal His$_6$ on TelC	This study
pETDuet-1::*telD*$_{LXG}$_His$_6$::*lapD1*	*E. coli* co-expression vector for the LXG domain of TelD with LapD1, C-terminal His$_6$ on TelD	This study
pETDuet-1::*telD*$_{LXG}$_His$_6$::*lapD1*$_{F77A}$	*E. coli* co-expression vector for the LXG domain of TelD with LapD1$_{F77A}$, C-terminal His$_6$ on TelD	This study
pETDuet-1::*telD*$_{LXG}$_His$_6$::*lapD1*$_{D81A}$	*E. coli* co-expression vector for the LXG domain of TelD with LapD1$_{D81A}$, C-terminal His$_6$ on TelD	This study
pET29b::*lapC2*	*E. coli* expression vector for LapC2	This study
pET29b::*lapD2*	*E. coli* expression vector for LapD2	This study
pET29b::*lapD2*$_{C59S}$	*E. coli* expression vector for LapD2$_{C59S}$	This study
pET29b::*lapD2*_His$_6$	*E. coli* expression vector for LapD2, C-terminal His$_6$ tag	This study
pET29b::*lapD2*$_{C59S}$_His$_6$	*E. coli* expression vector for LapD2$_{C59S}$, C-terminal His$_6$ tag	This study
pSCrhaB2::*telD*	Rhamnose inducible expression of TelD	This study
pPSV39::*tipD*	IPTG inducible expression of TipD	This study

References

Abdallah, A. M., Bestebroer, J., Savage, N. D., de Punder, K., van Zon, M., Wilson, L., . . . Peters, P. J. (2011). Mycobacterial secretion systems ESX-1 and ESX-5 play distinct roles in host cell death and inflammasome activation. *J Immunol*, *187*(9), 4744-4753. https://doi.org/10.4049/jimmunol.1101457

Abdallah, A. M., Gey van Pittius, N. C., Champion, P. A., Cox, J., Luirink, J., Vandenbroucke-Grauls, C. M., . . . Bitter, W. (2007). Type VII secretion--mycobacteria show the way. *Nat Rev Microbiol*, *5*(11), 883-891. https://doi.org/10.1038/nrmicro1773

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., . . . Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*, *66*(Pt 2), 213-221. https://doi.org/10.1107/S0907444909052925

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., . . . Adams, P. D. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr*, *68*(Pt 4), 352-367. https://doi.org/10.1107/S0907444912001308

Ahmad, S., Tsang, K. K., Sachar, K., Quentin, D., Tashin, T. M., Bullen, N. P., . . . Whitney, J. C. (2020). Structural basis for effector transmembrane domain recognition by type VI secretion system chaperones. *Elife*, *9*. https://doi.org/10.7554/eLife.62816

Ahmad, S., Wang, B., Walker, M. D., Tran, H. R., Stogios, P. J., Savchenko, A., . . . Whitney, J. C. (2019). An interbacterial toxin inhibits target cell growth by synthesizing (p)ppApp. *Nature*, *575*(7784), 674-678. https://doi.org/10.1038/s41586-019-1735-9

Ahmed, M. M., Aboshanab, K. M., Ragab, Y. M., Missiakas, D. M., & Aly, K. A. (2018). The transmembrane domain of the Staphylococcus aureus ESAT-6 component EssB mediates interaction with the integral membrane protein EsaA, facilitating

partially regulated secretion in a heterologous host. *Arch Microbiol*, *200*(7), 1075-1086. https://doi.org/10.1007/s00203-018-1519-x

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol*, *215*(3), 403-410. https://doi.org/10.1016/S0022-2836(05)80360-2

Aly, K. A., Anderson, M., Ohr, R. J., & Missiakas, D. (2017). Isolation of a Membrane Protein Complex for Type VII Secretion in Staphylococcus aureus. *J Bacteriol*, *199*(23). https://doi.org/10.1128/JB.00482-17

Anderson, M., Aly, K. A., Chen, Y. H., & Missiakas, D. (2013). Secretion of atypical protein substrates by the ESAT-6 secretion system of Staphylococcus aureus. *Mol Microbiol*, *90*(4), 734-743. https://doi.org/10.1111/mmi.12395

Anderson, M., Ohr, R. J., Aly, K. A., Nocadello, S., Kim, H. K., Schneewind, C. E., . . . Missiakas, D. (2017). EssE Promotes Staphylococcus aureus ESS-Dependent Protein Secretion To Modify Host Immune Responses during Infection. *J Bacteriol*, *199*(1). https://doi.org/10.1128/JB.00527-16

Aoki, S. K., Diner, E. J., de Roodenbeke, C. T., Burgess, B. R., Poole, S. J., Braaten, B. A., . . . Low, D. A. (2010). A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature*, *468*(7322), 439-442. https://doi.org/10.1038/nature09490

Aoki, S. K., Pamma, R., Hernday, A. D., Bickham, J. E., Braaten, B. A., & Low, D. A. (2005). Contact-dependent inhibition of growth in Escherichia coli. *Science*, *309*(5738), 1245-1248. https://doi.org/10.1126/science.1115109

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., & Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res*, *44*(W1), W344-350. https://doi.org/10.1093/nar/gkw408

Aspiras, M. B., Kazmerzak, K. M., Kolenbrander, P. E., McNab, R., Hardegen, N., & Jenkinson, H. F. (2000). Expression of green fluorescent protein in Streptococcus gordonii DL1 and its use as a species-specific marker in coadhesion with

Streptococcus oralis 34 in saliva-conditioned biofilms in vitro. *Appl Environ Microbiol*, *66*(9), 4074-4083. https://doi.org/10.1128/AEM.66.9.4074-4083.2000

Ates, L. S. (2020). New insights into the mycobacterial PE and PPE proteins provide a framework for future research. *Mol Microbiol*, *113*(1), 4-21. https://doi.org/10.1111/mmi.14409

Ates, L. S., & Brosch, R. (2017). Discovery of the type VII ESX-1 secretion needle? *Mol Microbiol*, *103*(1), 7-12. https://doi.org/10.1111/mmi.13579

Ates, L. S., van der Woude, A. D., Bestebroer, J., van Stempvoort, G., Musters, R. J., Garcia-Vallejo, J. J., . . . Bitter, W. (2016). The ESX-5 System of Pathogenic Mycobacteria Is Involved In Capsule Integrity and Virulence through Its Substrate PPE10. *PLoS Pathog*, *12*(6), e1005696. https://doi.org/10.1371/journal.ppat.1005696

Bayer-Santos, E., Cenens, W., Matsuyama, B. Y., Oka, G. U., Di Sessa, G., Mininel, I. D. V., . . . Farah, C. S. (2019). The opportunistic pathogen Stenotrophomonas maltophilia utilizes a type IV secretion system for interbacterial killing. *PLoS Pathog*, *15*(9), e1007651. https://doi.org/10.1371/journal.ppat.1007651

Beck, C. M., Morse, R. P., Cunningham, D. A., Iniguez, A., Low, D. A., Goulding, C. W., & Hayes, C. S. (2014). CdiA from Enterobacter cloacae delivers a toxic ribosomal RNase into target bacteria. *Structure*, *22*(5), 707-718. https://doi.org/10.1016/j.str.2014.02.012

Beckham, K. S., Ciccarelli, L., Bunduc, C. M., Mertens, H. D., Ummels, R., Lugmayr, W., . . . Houben, E. N. (2017). Structure of the mycobacterial ESX-5 type VII secretion system membrane complex by single-particle analysis. *Nat Microbiol*, *2*, 17047. https://doi.org/10.1038/nmicrobiol.2017.47

Beckham, K. S. H., Ritter, C., Chojnowski, G., Ziemianowicz, D. S., Mullapudi, E., Rettel, M., . . . Wilmanns, M. (2021). Structure of the mycobacterial ESX-5 type VII secretion system pore complex. *Sci Adv*, *7*(26). https://doi.org/10.1126/sciadv.abg9923

Biswas, S., & Biswas, I. (2014). A conserved streptococcal membrane protein, LsrS, exhibits a receptor-like function for lantibiotics. *J Bacteriol*, *196*(8), 1578-1587. https://doi.org/10.1128/JB.00028-14

Bitter, W., Houben, E. N., Bottai, D., Brodin, P., Brown, E. J., Cox, J. S., . . . Brosch, R. (2009). Systematic genetic nomenclature for type VII secretion systems. *PLoS Pathog*, *5*(10), e1000507. https://doi.org/10.1371/journal.ppat.1000507

Bottai, D., & Brosch, R. (2009). Mycobacterial PE, PPE and ESX clusters: novel insights into the secretion of these most unusual protein families. *Mol Microbiol*, *73*(3), 325-328. https://doi.org/10.1111/j.1365-2958.2009.06784.x

Bowman, L., & Palmer, T. (2021). The Type VII Secretion System of. *Annu Rev Microbiol*, *75*, 471-494. https://doi.org/10.1146/annurev-micro-012721-123600

Bowran, K., & Palmer, T. (2021). Extreme genetic diversity in the type VII secretion system of. *Microbiology (Reading)*, *167*(3). https://doi.org/10.1099/mic.0.001034

Brodin, P., de Jonge, M. I., Majlessi, L., Leclerc, C., Nilges, M., Cole, S. T., & Brosch, R. (2005). Functional analysis of early secreted antigenic target-6, the dominant T-cell antigen of Mycobacterium tuberculosis, reveals key residues involved in secretion, complex formation, virulence, and immunogenicity. *J Biol Chem*, *280*(40), 33953-33959. https://doi.org/10.1074/jbc.M503515200

Bryksin, A. V., & Matsumura, I. (2010). Rational design of a plasmid origin that replicates efficiently in both gram-positive and gram-negative bacteria. *PLoS One*, *5*(10), e13244. https://doi.org/10.1371/journal.pone.0013244

Buchan, D. W. A., & Jones, D. T. (2019). The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res*, *47*(W1), W402-W407. https://doi.org/10.1093/nar/gkz297

Bunduc, C. M., Fahrenkamp, D., Wald, J., Ummels, R., Bitter, W., Houben, E. N. G., & Marlovits, T. C. (2021). Structure and dynamics of a mycobacterial type VII secretion system. *Nature*, *593*(7859), 445-448. https://doi.org/10.1038/s41586-021-03517-z

Burkinshaw, B. J., Liang, X., Wong, M., Le, A. N. H., Lam, L., & Dong, T. G. (2018). A type VI secretion system effector delivery mechanism dependent on PAAR and a chaperone-co-chaperone complex. *Nat Microbiol*, *3*(5), 632-640. https://doi.org/10.1038/s41564-018-0144-4

Burts, M. L., DeDent, A. C., & Missiakas, D. M. (2008). EsaC substrate for the ESAT-6 secretion pathway and its role in persistent infections of Staphylococcus aureus. *Mol Microbiol*, *69*(3), 736-746. https://doi.org/10.1111/j.1365-2958.2008.06324.x

Burts, M. L., Williams, W. A., DeBord, K., & Missiakas, D. M. (2005). EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of Staphylococcus aureus infections. *Proc Natl Acad Sci U S A*, *102*(4), 1169-1174. https://doi.org/10.1073/pnas.0405620102

Cao, Z., Casabona, M. G., Kneuper, H., Chalmers, J. D., & Palmer, T. (2016). The type VII secretion system of Staphylococcus aureus secretes a nuclease toxin that targets competitor bacteria. *Nat Microbiol*, *2*, 16183. https://doi.org/10.1038/nmicrobiol.2016.183

Cardona, S. T., & Valvano, M. A. (2005). An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in Burkholderia cenocepacia. *Plasmid*, *54*(3), 219-228. https://doi.org/10.1016/j.plasmid.2005.03.004

Carr, S., Walker, D., James, R., Kleanthous, C., & Hemmings, A. M. (2000). Inhibition of a ribosome-inactivating ribonuclease: the crystal structure of the cytotoxic domain of colicin E3 in complex with its immunity protein. *Structure*, *8*(9), 949-960. https://doi.org/10.1016/s0969-2126(00)00186-6

Casabona, M. G., Buchanan, G., Zoltner, M., Harkins, C. P., Holden, M. T. G., & Palmer, T. (2017). Functional analysis of the EsaB component of the. *Microbiology (Reading)*, *163*(12), 1851-1863. https://doi.org/10.1099/mic.0.000580

Champion, P. A., Stanley, S. A., Champion, M. M., Brown, E. J., & Cox, J. S. (2006). C-terminal signal sequence promotes virulence factor secretion in Mycobacterium

tuberculosis. *Science*, *313*(5793), 1632-1636. https://doi.org/10.1126/science.1131167

Chapot-Chartier, M. P., Rul, F., Nardi, M., & Gripon, J. C. (1994). Gene cloning and characterization of PepC, a cysteine aminopeptidase from Streptococcus thermophilus, with sequence similarity to the eucaryotic bleomycin hydrolase. *Eur J Biochem*, *224*(2), 497-506. https://doi.org/10.1111/j.1432-1033.1994.00497.x

Chatterjee, A., Willett, J. L. E., Dunny, G. M., & Duerkop, B. A. (2021). Phage infection and sub-lethal antibiotic exposure mediate Enterococcus faecalis type VII secretion system dependent inhibition of bystander bacteria. *PLoS Genet*, *17*(1), e1009204. https://doi.org/10.1371/journal.pgen.1009204

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., . . . Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*, *66*(Pt 1), 12-21. https://doi.org/10.1107/S0907444909042073

Chou, S., Bui, N. K., Russell, A. B., Lexa, K. W., Gardiner, T. E., LeRoux, M., . . . Mougous, J. D. (2012). Structure of a peptidoglycan amidase effector targeted to Gram-negative bacteria by the type VI secretion system. *Cell Rep*, *1*(6), 656-664. https://doi.org/10.1016/j.celrep.2012.05.016

Christie, P. J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S., & Cascales, E. (2005). Biogenesis, architecture, and function of bacterial type IV secretion systems. *Annu Rev Microbiol*, *59*, 451-485. https://doi.org/10.1146/annurev.micro.58.030603.123630

Christie, P. J., Whitaker, N., & González-Rivera, C. (2014). Mechanism and structure of the bacterial type IV secretion systems. *Biochim Biophys Acta*, *1843*(8), 1578-1591. https://doi.org/10.1016/j.bbamcr.2013.12.019

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., . . . Barrell, B. G. (1998). Deciphering the biology of Mycobacterium tuberculosis from the

complete genome sequence. *Nature*, *393*(6685), 537-544. https://doi.org/10.1038/31159

Colin, R., Ni, B., Laganenka, L., & Sourjik, V. (2021). Multiple functions of flagellar motility and chemotaxis in bacterial physiology. *FEMS Microbiol Rev*, *45*(6). https://doi.org/10.1093/femsre/fuab038

Conrad, W. H., Osman, M. M., Shanahan, J. K., Chu, F., Takaki, K. K., Cameron, J., . . . Ramakrishnan, L. (2017). Mycobacterial ESX-1 secretion system mediates host cell lysis through bacterium contact-dependent gross membrane disruptions. *Proc Natl Acad Sci U S A*, *114*(6), 1371-1376. https://doi.org/10.1073/pnas.1620133114

Cotter, P. D., Ross, R. P., & Hill, C. (2013). Bacteriocins - a viable alternative to antibiotics? *Nat Rev Microbiol*, *11*(2), 95-105. https://doi.org/10.1038/nrmicro2937

Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr*, *62*(Pt 9), 1002-1011. https://doi.org/10.1107/S0907444906022116

Cox, G., & Wright, G. D. (2013). Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions. *Int J Med Microbiol*, *303*(6-7), 287-292. https://doi.org/10.1016/j.ijmm.2013.02.009

Crosskey, T. D., Beckham, K. S. H., & Wilmanns, M. (2020). The ATPases of the mycobacterial type VII secretion system: Structural and mechanistic insights into secretion. *Prog Biophys Mol Biol*, *152*, 25-34. https://doi.org/10.1016/j.pbiomolbio.2019.11.008

Crump, K. E., Bainbridge, B., Brusko, S., Turner, L. S., Ge, X., Stone, V., . . . Kitten, T. (2014). The relationship of the lipoprotein SsaB, manganese and superoxide dismutase in Streptococcus sanguinis virulence for endocarditis. *Mol Microbiol*, *92*(6), 1243-1259. https://doi.org/10.1111/mmi.12625

Daleke, M. H., Ummels, R., Bawono, P., Heringa, J., Vandenbroucke-Grauls, C. M., Luirink, J., & Bitter, W. (2012a). General secretion signal for the mycobacterial

type VII secretion pathway. *Proc Natl Acad Sci U S A*, *109*(28), 11342-11347. https://doi.org/10.1073/pnas.1119453109

Daleke, M. H., van der Woude, A. D., Parret, A. H., Ummels, R., de Groot, A. M., Watson, D., . . . Houben, E. N. (2012b). Specific chaperones for the type VII protein secretion pathway. *J Biol Chem*, *287*(38), 31939-31947. https://doi.org/10.1074/jbc.M112.397596

Damen, M. P. M., Phan, T. H., Ummels, R., Rubio-Canalejas, A., Bitter, W., & Houben, E. N. G. (2020). Modification of a PE/PPE substrate pair reroutes an Esx substrate pair from the mycobacterial ESX-1 type VII secretion system to the ESX-5 system. *J Biol Chem*, *295*(18), 5960-5969. https://doi.org/10.1074/jbc.RA119.011682

Davis, I. W., Murray, L. W., Richardson, J. S., & Richardson, D. C. (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res*, *32*(Web Server issue), W615-619. https://doi.org/10.1093/nar/gkh398

de Jonge, M. I., Pehau-Arnaudet, G., Fretz, M. M., Romain, F., Bottai, D., Brodin, P., . . . Brosch, R. (2007). ESAT-6 from Mycobacterium tuberculosis dissociates from its putative chaperone CFP-10 under acidic conditions and exhibits membrane-lysing activity. *J Bacteriol*, *189*(16), 6028-6034. https://doi.org/10.1128/JB.00469-07

de Moraes, M. H., Hsu, F., Huang, D., Bosch, D. E., Zeng, J., Radey, M. C., . . . Mougous, J. D. (2021). An interbacterial DNA deaminase toxin directly mutagenizes surviving target populations. *Elife*, *10*. https://doi.org/10.7554/eLife.62967

Demain, A. L. (1999). Pharmaceutically active secondary metabolites of microorganisms. *Appl Microbiol Biotechnol*, *52*(4), 455-463. https://doi.org/10.1007/s002530051546

Demchick, P., & Koch, A. L. (1996). The permeability of the wall fabric of Escherichia coli and Bacillus subtilis. *J Bacteriol*, *178*(3), 768-773. https://doi.org/10.1128/jb.178.3.768-773.1996

Di Luca, M., Bottai, D., Batoni, G., Orgeur, M., Aulicino, A., Counoupas, C., . . . Esin, S. (2012). The ESX-5 associated eccB-EccC locus is essential for Mycobacterium tuberculosis viability. *PLoS One*, *7*(12), e52059. https://doi.org/10.1371/journal.pone.0052059

Dreisbach, A., Hempel, K., Buist, G., Hecker, M., Becher, D., & van Dijl, J. M. (2010). Profiling the surfacome of Staphylococcus aureus. *Proteomics*, *10*(17), 3082-3096. https://doi.org/10.1002/pmic.201000062

Duerkop, B. A., Huo, W., Bhardwaj, P., Palmer, K. L., & Hooper, L. V. (2016). Molecular Basis for Lytic Bacteriophage Resistance in Enterococci. *mBio*, *7*(4). https://doi.org/10.1128/mBio.01304-16

Ekiert, D. C., & Cox, J. S. (2014). Structure of a PE-PPE-EspG complex from Mycobacterium tuberculosis reveals molecular specificity of ESX protein secretion. *Proc Natl Acad Sci U S A*, *111*(41), 14758-14763. https://doi.org/10.1073/pnas.1409345111

El Ghachi, M., Bouhss, A., Barreteau, H., Touzé, T., Auger, G., Blanot, D., & Mengin-Lecreulx, D. (2006). Colicin M exerts its bacteriolytic effect via enzymatic degradation of undecaprenyl phosphate-linked peptidoglycan precursors. *J Biol Chem*, *281*(32), 22761-22772. https://doi.org/10.1074/jbc.M602834200

Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, *60*(Pt 12 Pt 1), 2126-2132. https://doi.org/10.1107/S0907444904019158

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*, *66*(Pt 4), 486-501. https://doi.org/10.1107/S0907444910007493

Erdős, G., & Dosztányi, Z. (2020). Analyzing Protein Disorder with IUPred2A. *Curr Protoc Bioinformatics*, *70*(1), e99. https://doi.org/10.1002/cpbi.99

Famelis, N., Rivera-Calzada, A., Degliesposti, G., Wingender, M., Mietrach, N., Skehel, J. M., . . . Geibel, S. (2019). Architecture of the mycobacterial type VII secretion system. *Nature*, *576*(7786), 321-325. https://doi.org/10.1038/s41586-019-1633-1

Fenton, A. K., Manuse, S., Flores-Kim, J., Garcia, P. S., Mercy, C., Grangeasse, C., . . . Rudner, D. Z. (2018). Phosphorylation-dependent activation of the cell wall synthase PBP2a in. *Proc Natl Acad Sci U S A*, *115*(11), 2812-2817. https://doi.org/10.1073/pnas.1715218115

Finn, R. D., Clements, J., Arndt, W., Miller, B. L., Wheeler, T. J., Schreiber, F., . . . Eddy, S. R. (2015). HMMER web server: 2015 update. *Nucleic Acids Res*, *43*(W1), W30-38. https://doi.org/10.1093/nar/gkv397

Gao, L. Y., Guo, S., McLaughlin, B., Morisaki, H., Engel, J. N., & Brown, E. J. (2004). A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion. *Mol Microbiol*, *53*(6), 1677-1693. https://doi.org/10.1111/j.1365-2958.2004.04261.x

García-Bayona, L., Guo, M. S., & Laub, M. T. (2017). Contact-dependent killing by. *Elife*, *6*. https://doi.org/10.7554/eLife.24869

Garrett, S. R., Mariano, G., Dicks, J., & Palmer, T. (2022). Homologous recombination between tandem paralogues drives evolution of a subset of Type VII secretion system immunity genes in firmicute bacteria. *bioRxiv*, 2022.2001.2007.475358. https://doi.org/10.1101/2022.01.07.475358

Gey van Pittius, N. C., Sampson, S. L., Lee, H., Kim, Y., van Helden, P. D., & Warren, R. M. (2006). Evolution and expansion of the Mycobacterium tuberculosis PE and PPE multigene families and their association with the duplication of the ESAT-6 (esx) gene cluster regions. *BMC Evol Biol*, *6*, 95. https://doi.org/10.1186/1471-2148-6-95

Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., & Ferrin, T. E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci*, *27*(1), 14-25. https://doi.org/10.1002/pro.3235

Gray, T. A., Clark, R. R., Boucher, N., Lapierre, P., Smith, C., & Derbyshire, K. M. (2016). Intercellular communication and conjugation are mediated by ESX

secretion systems in mycobacteria. *Science*, *354*(6310), 347-350. https://doi.org/10.1126/science.aag0828

Green, E. R., & Mecsas, J. (2016). Bacterial Secretion Systems: An Overview. *Microbiol Spectr*, *4*(1). https://doi.org/10.1128/microbiolspec.VMBF-0012-2015

Grossman, A. S., Mauer, T. J., Forest, K. T., & Goodrich-Blair, H. (2021). A Widespread Bacterial Secretion System with Diverse Substrates. *mBio*, *12*(4), e0195621. https://doi.org/10.1128/mBio.01956-21

Guilhelmelli, F., Vilela, N., Albuquerque, P., Derengowski, L. a. S., Silva-Pereira, I., & Kyaw, C. M. (2013). Antibiotic development challenges: the various mechanisms of action of antimicrobial peptides and of bacterial resistance. *Front Microbiol*, *4*, 353. https://doi.org/10.3389/fmicb.2013.00353

Gérard, F., Brooks, M. A., Barreteau, H., Touzé, T., Graille, M., Bouhss, A., . . . Mengin-Lecreulx, D. (2011). X-ray structure and site-directed mutagenesis analysis of the Escherichia coli colicin M immunity protein. *J Bacteriol*, *193*(1), 205-214. https://doi.org/10.1128/JB.01119-10

Hasegawa, N., Sekizuka, T., Sugi, Y., Kawakami, N., Ogasawara, Y., Kato, K., . . . Kuroda, M. (2017). Characterization of the Pathogenicity of Streptococcus intermedius TYG1620 Isolated from a Human Brain Abscess Based on the Complete Genome Sequence with Transcriptome Analysis and Transposon Mutagenesis in a Murine Subcutaneous Abscess Model. *Infect Immun*, *85*(2). https://doi.org/10.1128/IAI.00886-16

Hayes, C. S., Koskiniemi, S., Ruhe, Z. C., Poole, S. J., & Low, D. A. (2014). Mechanisms and biological roles of contact-dependent growth inhibition systems. *Cold Spring Harb Perspect Med*, *4*(2). https://doi.org/10.1101/cshperspect.a010025

Hendrickson, W. A., Horton, J. R., & LeMaster, D. M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J*, *9*(5), 1665-1672. https://doi.org/10.1002/j.1460-2075.1990.tb08287.x

Holberger, L. E., Garza-Sánchez, F., Lamoureux, J., Low, D. A., & Hayes, C. S. (2012). A novel family of toxin/antitoxin proteins in Bacillus species. *FEBS Lett*, *586*(2), 132-136. https://doi.org/10.1016/j.febslet.2011.12.020

Holm, L. (2020). DALI and the persistence of protein shape. *Protein Sci*, *29*(1), 128-140. https://doi.org/10.1002/pro.3749

Holm, L., & Laakso, L. M. (2016). Dali server update. *Nucleic Acids Res*, *44*(W1), W351-355. https://doi.org/10.1093/nar/gkw357

Houben, D., Demangel, C., van Ingen, J., Perez, J., Baldeón, L., Abdallah, A. M., . . . Peters, P. J. (2012). ESX-1-mediated translocation to the cytosol controls virulence of mycobacteria. *Cell Microbiol*, *14*(8), 1287-1298. https://doi.org/10.1111/j.1462-5822.2012.01799.x

Houben, E. N., Bestebroer, J., Ummels, R., Wilson, L., Piersma, S. R., Jiménez, C. R., . . . Bitter, W. (2012). Composition of the type VII secretion system membrane complex. *Mol Microbiol*, *86*(2), 472-484. https://doi.org/10.1111/j.1365-2958.2012.08206.x

Izquierdo Lafuente, B., Ummels, R., Kuijl, C., Bitter, W., & Speer, A. (2021). Mycobacterium tuberculosis Toxin CpnT Is an ESX-5 Substrate and Requires Three Type VII Secretion Systems for Intracellular Secretion. *mBio*, *12*(2). https://doi.org/10.1128/mBio.02983-20

Jana, B., Fridman, C. M., Bosis, E., & Salomon, D. (2019). A modular effector with a DNase domain and a marker for T6SS substrates. *Nat Commun*, *10*(1), 3595. https://doi.org/10.1038/s41467-019-11546-6

Jeon, H. J., Noda, M., Matoba, Y., Kumagai, T., & Sugiyama, M. (2009). Crystal structure and mutagenic analysis of a bacteriocin immunity protein, Mun-im. *Biochem Biophys Res Commun*, *378*(3), 574-578. https://doi.org/10.1016/j.bbrc.2008.11.093

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. https://doi.org/10.1038/s41586-021-03819-2

Jäger, F., Kneuper, H., & Palmer, T. (2018). EssC is a specificity determinant for Staphylococcus aureus type VII secretion. *Microbiology (Reading)*, *164*(5), 816-820. https://doi.org/10.1099/mic.0.000650

Kanonenberg, K., Spitz, O., Erenburg, I. N., Beer, T., & Schmitt, L. (2018). Type I secretion system-it takes three and a substrate. *FEMS Microbiol Lett*, *365*(11). https://doi.org/10.1093/femsle/fny094

Karplus, P. A., & Diederichs, K. (2012). Linking crystallographic model and data quality. *Science*, *336*(6084), 1030-1033. https://doi.org/10.1126/science.1218231

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*, *30*(4), 772-780. https://doi.org/10.1093/molbev/mst010

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*, *10*(6), 845-858. https://doi.org/10.1038/nprot.2015.053

Klein, T. A., Ahmad, S., & Whitney, J. C. (2020). Contact-Dependent Interbacterial Antagonism Mediated by Protein Secretion Machines. *Trends Microbiol*, *28*(5), 387-400. https://doi.org/10.1016/j.tim.2020.01.003

Klein, T. A., Grebenc, D. W., Gandhi, S. Y., Shah, V. S., Kim, Y., & Whitney, J. C. (2021). Structure of the Extracellular Region of the Bacterial Type VIIb Secretion System Subunit EsaA. *Structure*, *29*(2), 177-185.e176. https://doi.org/10.1016/j.str.2020.11.002

Klein, T. A., Grebenc, D. W., Shah, P. Y., McArthur, O. D., Surette, M. G., Kim, Y., & Whitney, J. C. (2022). A conserved chaperone-co-chaperone pair are required for LXG toxin export by the bacterial type VIIb secretion system. *mBio (in revision)*.

Klein, T. A., Pazos, M., Surette, M. G., Vollmer, W., & Whitney, J. C. (2018). Molecular Basis for Immunity Protein Recognition of a Type VII Secretion System Exported Antibacterial Toxin. *J Mol Biol*, *430*(21), 4344-4358. https://doi.org/10.1016/j.jmb.2018.08.027

Kneuper, H., Cao, Z. P., Twomey, K. B., Zoltner, M., Jäger, F., Cargill, J. S., . . . Palmer, T. (2014). Heterogeneity in ess transcriptional organization and variable contribution of the Ess/Type VII protein secretion system to virulence across closely related Staphylocccus aureus strains. *Mol Microbiol*, *93*(5), 928-943. https://doi.org/10.1111/mmi.12707

Kobayashi, K. (2021). Diverse LXG toxin and antitoxin systems specifically mediate intraspecies competition in Bacillus subtilis biofilms. *PLoS Genet*, *17*(7), e1009682. https://doi.org/10.1371/journal.pgen.1009682

Korea, C. G., Balsamo, G., Pezzicoli, A., Merakou, C., Tavarini, S., Bagnoli, F., . . . Unnikrishnan, M. (2014). Staphylococcal Esx proteins modulate apoptosis and release of intracellular Staphylococcus aureus during infection in epithelial cells. *Infect Immun*, *82*(10), 4144-4153. https://doi.org/10.1128/IAI.01576-14

Korotkova, N., Freire, D., Phan, T. H., Ummels, R., Creekmore, C. C., Evans, T. J., . . . Korotkov, K. V. (2014). Structure of the Mycobacterium tuberculosis type VII secretion system chaperone EspG5 in complex with PE25-PPE41 dimer. *Mol Microbiol*, *94*(2), 367-382. https://doi.org/10.1111/mmi.12770

Korotkova, N., Piton, J., Wagner, J. M., Boy-Röttger, S., Japaridze, A., Evans, T. J., . . . Korotkov, K. V. (2015). Structure of EspB, a secreted substrate of the ESX-1 secretion system of Mycobacterium tuberculosis. *J Struct Biol*, *191*(2), 236-244. https://doi.org/10.1016/j.jsb.2015.06.003

Krissinel, E., & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, *372*(3), 774-797. https://doi.org/10.1016/j.jmb.2007.05.022

Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*, *305*(3), 567-580. https://doi.org/10.1006/jmbi.2000.4315

Laencina, L., Dubois, V., Le Moigne, V., Viljoen, A., Majlessi, L., Pritchard, J., . . . Girard-Misguich, F. (2018). Identification of genes required for. *Proc Natl Acad Sci U S A*, *115*(5), E1002-E1011. https://doi.org/10.1073/pnas.1713195115

Laskowski, R. A. (2001). PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res*, *29*(1), 221-222. https://doi.org/10.1093/nar/29.1.221

Lauber, F., Deme, J. C., Lea, S. M., & Berks, B. C. (2018). Type 9 secretion system structures reveal a new protein transport mechanism. *Nature*, *564*(7734), 77-82. https://doi.org/10.1038/s41586-018-0693-y

Le, N. H., Pinedo, V., Lopez, J., Cava, F., & Feldman, M. F. (2021). Killing of Gram-negative and Gram-positive bacteria by a bifunctional cell wall-targeting T6SS effector. *Proc Natl Acad Sci U S A*, *118*(40). https://doi.org/10.1073/pnas.2106555118

Leiman, P. G., Basler, M., Ramagopal, U. A., Bonanno, J. B., Sauder, J. M., Pukatzki, S., . . . Mekalanos, J. J. (2009). Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc Natl Acad Sci U S A*, *106*(11), 4154-4159. https://doi.org/10.1073/pnas.0813360106

Li, M., Le Trong, I., Carl, M. A., Larson, E. T., Chou, S., De Leon, J. A., . . . Mougous, J. D. (2012). Structural basis for type VI secretion effector recognition by a cognate immunity protein. *PLoS Pathog*, *8*(4), e1002613. https://doi.org/10.1371/journal.ppat.1002613

Limoli, D. H., Jones, C. J., & Wozniak, D. J. (2015). Bacterial Extracellular Polysaccharides in Biofilm Formation and Function. *Microbiol Spectr*, *3*(3). https://doi.org/10.1128/microbiolspec.MB-0011-2014

Llosa, M., Gomis-Rüth, F. X., Coll, M., & de la Cruz Fd, F. (2002). Bacterial conjugation: a two-step mechanism for DNA transport. *Mol Microbiol*, *45*(1), 1-8. https://doi.org/10.1046/j.1365-2958.2002.03014.x

Lo Sapio, M., Hilleringmann, M., Barocchi, M. A., & Moschioni, M. (2012). A novel strategy to over-express and purify homologous proteins from Streptococcus

pneumoniae. *J Biotechnol*, *157*(2), 279-286. https://doi.org/10.1016/j.jbiotec.2011.11.011

Lou, Y., Rybniker, J., Sala, C., & Cole, S. T. (2017). EspC forms a filamentous structure in the cell envelope of Mycobacterium tuberculosis and impacts ESX-1 secretion. *Mol Microbiol*, *103*(1), 26-38. https://doi.org/10.1111/mmi.13575

Macey, M. G., Whiley, R. A., Miller, L., & Nagamune, H. (2001). Effect on polymorphonuclear cell function of a human-specific cytotoxin, intermedilysin, expressed by Streptococcus intermedius. *Infect Immun*, *69*(10), 6102-6109. https://doi.org/10.1128/IAI.69.10.6102-6109.2001

Mahairas, G. G., Sabo, P. J., Hickey, M. J., Singh, D. C., & Stover, C. K. (1996). Molecular analysis of genetic differences between Mycobacterium bovis BCG and virulent M. bovis. *J Bacteriol*, *178*(5), 1274-1282. https://doi.org/10.1128/jb.178.5.1274-1282.1996

Mariano, G., Trunk, K., Williams, D. J., Monlezun, L., Strahl, H., Pitt, S. J., & Coulthurst, S. J. (2019). A family of Type VI secretion system effector proteins that form ion-selective pores. *Nat Commun*, *10*(1), 5484. https://doi.org/10.1038/s41467-019-13439-0

Mietrach, N., Damian-Aparicio, D., Isupov, M., Krupka, M., Lopez, D., & Geibel, S. (2020a). The conserved core component EsaA mediates bacterial killing by the type VIIb secretion system. *Research Square (preprint)*.

Mietrach, N., Damián-Aparicio, D., Mielich-Süss, B., Lopez, D., & Geibel, S. (2020b). Substrate Interaction with the EssC Coupling Protein of the Type VIIb Secretion System. *J Bacteriol*, *202*(7). https://doi.org/10.1128/JB.00646-19

Mietrach, N., Schlosser, A., & Geibel, S. (2019). An extracellular domain of the EsaA membrane component of the type VIIb secretion system: expression, purification and crystallization. *Acta Crystallogr F Struct Biol Commun*, *75*(Pt 12), 725-730. https://doi.org/10.1107/S2053230X1901495X

Miller, M. B., & Bassler, B. L. (2001). Quorum sensing in bacteria. *Annu Rev Microbiol*, *55*, 165-199. https://doi.org/10.1146/annurev.micro.55.1.165

Minor, W., Cymborowski, M., Otwinowski, Z., & Chruszcz, M. (2006). HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr*, *62*(Pt 8), 859-866. https://doi.org/10.1107/S0907444906019949

Mok, B. Y., de Moraes, M. H., Zeng, J., Bosch, D. E., Kotrys, A. V., Raguram, A., . . . Liu, D. R. (2020). A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. *Nature*, *583*(7817), 631-637. https://doi.org/10.1038/s41586-020-2477-4

Moretti, S., Armougom, F., Wallace, I. M., Higgins, D. G., Jongeneel, C. V., & Notredame, C. (2007). The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. *Nucleic Acids Res*, *35*(Web Server issue), W645-648. https://doi.org/10.1093/nar/gkm333

Mougous, J. D., Cuff, M. E., Raunser, S., Shen, A., Zhou, M., Gifford, C. A., . . . Mekalanos, J. J. (2006). A virulence locus of Pseudomonas aeruginosa encodes a protein secretion apparatus. *Science*, *312*(5779), 1526-1530. https://doi.org/10.1126/science.1128393

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., . . . Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr*, *67*(Pt 4), 355-367. https://doi.org/10.1107/S0907444911001314

Mészáros, B., Erdos, G., & Dosztányi, Z. (2018). IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res*, *46*(W1), W329-W337. https://doi.org/10.1093/nar/gky384

Navarre, W. W., & Schneewind, O. (1999). Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev*, *63*(1), 174-229. https://doi.org/10.1128/MMBR.63.1.174-229.1999

Nguyen, V. S., Douzi, B., Durand, E., Roussel, A., Cascales, E., & Cambillau, C. (2018). Towards a complete structural deciphering of Type VI secretion system. *Curr Opin Struct Biol*, *49*, 77-84. https://doi.org/10.1016/j.sbi.2018.01.007

Ohol, Y. M., Goetz, D. H., Chan, K., Shiloh, M. U., Craik, C. S., & Cox, J. S. (2010). Mycobacterium tuberculosis MycP1 protease plays a dual role in regulation of ESX-1 secretion and virulence. *Cell Host Microbe*, *7*(3), 210-220. https://doi.org/10.1016/j.chom.2010.02.006

Ohr, R. J., Anderson, M., Shi, M., Schneewind, O., & Missiakas, D. (2017). EssD, a Nuclease Effector of the Staphylococcus aureus ESS Pathway. *J Bacteriol*, *199*(1). https://doi.org/10.1128/JB.00528-16

Olschläger, T., Turba, A., & Braun, V. (1991). Binding of the immunity protein inactivates colicin M. *Mol Microbiol*, *5*(5), 1105-1111. https://doi.org/10.1111/j.1365-2958.1991.tb01883.x

Olson, A. B., Kent, H., Sibley, C. D., Grinwis, M. E., Mabon, P., Ouellette, C., . . . Corbett, C. R. (2013). Phylogenetic relationship and virulence inference of Streptococcus Anginosus Group: curated annotation and whole-genome comparative analysis support distinct species designation. *BMC Genomics*, *14*, 895. https://doi.org/10.1186/1471-2164-14-895

Pajuelo, D., Tak, U., Zhang, L., Danilchanka, O., Tischler, A. D., & Niederweis, M. (2021). Toxin secretion and trafficking by Mycobacterium tuberculosis. *Nat Commun*, *12*(1), 6592. https://doi.org/10.1038/s41467-021-26925-1

Pallen, M. J. (2002). The ESAT-6/WXG100 superfamily -- and a new Gram-positive secretion system? *Trends Microbiol*, *10*(5), 209-212.

Palmer, T., Finney, A. J., Saha, C. K., Atkinson, G. C., & Sargent, F. (2021). A holin/peptidoglycan hydrolase-dependent protein secretion system. *Mol Microbiol*, *115*(3), 345-355. https://doi.org/10.1111/mmi.14599

Parsot, C., Hamiaux, C., & Page, A. L. (2003). The various and varying roles of specific chaperones in type III secretion systems. *Curr Opin Microbiol*, *6*(1), 7-14. https://doi.org/10.1016/s1369-5274(02)00002-4

Paton, J. C., & Trappetti, C. (2019). Capsular Polysaccharide. *Microbiol Spectr*, *7*(2). https://doi.org/10.1128/microbiolspec.GPP3-0019-2018

Pazos, M., Otten, C., & Vollmer, W. (2018). Bacterial Cell Wall Precursor Phosphatase Assays Using Thin-layer Chromatography (TLC) and High Pressure Liquid Chromatography (HPLC). *Bio Protoc*, *8*(6), e2761. https://doi.org/10.21769/BioProtoc.2761

Pei, J., & Grishin, N. V. (2001). AL2CO: calculation of positional conservation in a protein sequence alignment. *Bioinformatics*, *17*(8), 700-712. https://doi.org/10.1093/bioinformatics/17.8.700

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, *25*(13), 1605-1612. https://doi.org/10.1002/jcc.20084

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., . . . Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci*, *30*(1), 70-82. https://doi.org/10.1002/pro.3943

Phan, T. H., Ummels, R., Bitter, W., & Houben, E. N. (2017). Identification of a substrate domain that determines system specificity in mycobacterial type VII secretion systems. *Sci Rep*, *7*, 42704. https://doi.org/10.1038/srep42704

Piton, J., Pojer, F., Wakatsuki, S., Gati, C., & Cole, S. T. (2020). High resolution CryoEM structure of the ring-shaped virulence factor EspB from. *J Struct Biol X*, *4*, 100029. https://doi.org/10.1016/j.yjsbx.2020.100029

Poulsen, C., Panjikar, S., Holton, S. J., Wilmanns, M., & Song, Y. H. (2014). WXG100 protein superfamily consists of three subfamilies and exhibits an α-helical C-terminal conserved residue pattern. *PLoS One*, *9*(2), e89313. https://doi.org/10.1371/journal.pone.0089313

Poweleit, N., Czudnochowski, N., Nakagawa, R., Trinidad, D. D., Murphy, K. C., Sassetti, C. M., & Rosenberg, O. S. (2019). The structure of the endogenous ESX-3 secretion system. *Elife*, *8*. https://doi.org/10.7554/eLife.52983

Pym, A. S., Brodin, P., Brosch, R., Huerre, M., & Cole, S. T. (2002). Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti. *Mol Microbiol*, *46*(3), 709-717. https://doi.org/10.1046/j.1365-2958.2002.03237.x

Quentin, D., Ahmad, S., Shanthamoorthy, P., Mougous, J. D., Whitney, J. C., & Raunser, S. (2018). Mechanism of loading and translocation of type VI secretion system effector Tse6. *Nat Microbiol*, *3*(10), 1142-1152. https://doi.org/10.1038/s41564-018-0238-z

Renshaw, P. S., Lightbody, K. L., Veverka, V., Muskett, F. W., Kelly, G., Frenkiel, T. A., . . . Carr, M. D. (2005). Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. *EMBO J*, *24*(14), 2491-2498. https://doi.org/10.1038/sj.emboj.7600732

Riley, M. A., & Wertz, J. E. (2002). Bacteriocins: evolution, ecology, and application. *Annu Rev Microbiol*, *56*, 117-137. https://doi.org/10.1146/annurev.micro.56.012302.161024

Rivera-Calzada, A., Famelis, N., Llorca, O., & Geibel, S. (2021). Type VII secretion systems: structure, functions and transport models. *Nat Rev Microbiol*, *19*(9), 567-584. https://doi.org/10.1038/s41579-021-00560-5

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*, *42*(Web Server issue), W320-324. https://doi.org/10.1093/nar/gku316

Rosenberg, O. S., Dovala, D., Li, X., Connolly, L., Bendebury, A., Finer-Moore, J., . . . Cox, J. S. (2015). Substrates Control Multimerization and Activation of the Multi-Domain ATPase Motor of Type VII Secretion. *Cell*, *161*(3), 501-512. https://doi.org/10.1016/j.cell.2015.03.040

Ross, B. D., Verster, A. J., Radey, M. C., Schmidtke, D. T., Pope, C. E., Hoffman, L. R., . . . Mougous, J. D. (2019). Human gut bacteria contain acquired interbacterial defence systems. *Nature*, *575*(7781), 224-228. https://doi.org/10.1038/s41586-019-1708-z

Ruhe, Z. C., Low, D. A., & Hayes, C. S. (2013). Bacterial contact-dependent growth inhibition. *Trends Microbiol*, *21*(5), 230-237. https://doi.org/10.1016/j.tim.2013.02.003

Ruhe, Z. C., Nguyen, J. Y., Xiong, J., Koskiniemi, S., Beck, C. M., Perkins, B. R., . . . Hayes, C. S. (2017). CdiA Effectors Use Modular Receptor-Binding Domains To Recognize Target Bacteria. *mBio*, *8*(2). https://doi.org/10.1128/mBio.00290-17

Ruhe, Z. C., Subramanian, P., Song, K., Nguyen, J. Y., Stevens, T. A., Low, D. A., . . . Hayes, C. S. (2018). Programmed Secretion Arrest and Receptor-Triggered Toxin Export during Antibacterial Contact-Dependent Growth Inhibition. *Cell*, *175*(4), 921-933.e914. https://doi.org/10.1016/j.cell.2018.10.033

Russell, A. B., Hood, R. D., Bui, N. K., LeRoux, M., Vollmer, W., & Mougous, J. D. (2011). Type VI secretion delivers bacteriolytic effectors to target cells. *Nature*, *475*(7356), 343-347. https://doi.org/10.1038/nature10244

Russell, A. B., LeRoux, M., Hathazi, K., Agnello, D. M., Ishikawa, T., Wiggins, P. A., . . . Mougous, J. D. (2013). Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature*, *496*(7446), 508-512. https://doi.org/10.1038/nature12074

Russell, A. B., Peterson, S. B., & Mougous, J. D. (2014). Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol*, *12*(2), 137-148. https://doi.org/10.1038/nrmicro3185

Russell, A. B., Singh, P., Brittnacher, M., Bui, N. K., Hood, R. D., Carl, M. A., . . . Mougous, J. D. (2012). A widespread bacterial type VI secretion effector superfamily identified using a heuristic approach. *Cell Host Microbe*, *11*(5), 538-549. https://doi.org/10.1016/j.chom.2012.04.007

Sala, A., Bordes, P., & Genevaux, P. (2014). Multitasking SecB chaperones in bacteria. *Front Microbiol*, *5*, 666. https://doi.org/10.3389/fmicb.2014.00666

Santucci, P., Diomandé, S., Poncin, I., Alibaud, L., Viljoen, A., Kremer, L., . . . Canaan, S. (2018). Delineating the Physiological Roles of the PE and Catalytic Domains of LipY in Lipid Consumption in Mycobacterium-Infected Foamy Macrophages. *Infect Immun*, *86*(9). https://doi.org/10.1128/IAI.00394-18

Sayes, F., Sun, L., Di Luca, M., Simeone, R., Degaiffier, N., Fiette, L., . . . Majlessi, L. (2012). Strong immunogenicity and cross-reactivity of Mycobacterium tuberculosis ESX-5 type VII secretion: encoded PE-PPE proteins predicts vaccine potential. *Cell Host Microbe*, *11*(4), 352-363. https://doi.org/10.1016/j.chom.2012.03.003

Schlegel, L., Grimont, F., Ageron, E., Grimont, P. A. D., & Bouvet, A. (2003). Reappraisal of the taxonomy of the Streptococcus bovis/Streptococcus equinus complex and related species: description of Streptococcus gallolyticus subsp. gallolyticus subsp. nov., S. gallolyticus subsp. macedonicus subsp. nov. and S. gallolyticus subsp. pasteurianus subsp. nov. *Int J Syst Evol Microbiol*, *53*(Pt 3), 631-645. https://doi.org/10.1099/ijs.0.02361-0

Schägger, H. (2006). Tricine-SDS-PAGE. *Nat Protoc*, *1*(1), 16-22. https://doi.org/10.1038/nprot.2006.4

Serafini, A., Boldrin, F., Palù, G., & Manganelli, R. (2009). Characterization of a Mycobacterium tuberculosis ESX-3 conditional mutant: essentiality and rescue by iron and zinc. *J Bacteriol*, *191*(20), 6340-6344. https://doi.org/10.1128/JB.00756-09

Serafini, A., Pisu, D., Palù, G., Rodriguez, G. M., & Manganelli, R. (2013). The ESX-3 secretion system is necessary for iron and zinc homeostasis in Mycobacterium tuberculosis. *PLoS One*, *8*(10), e78351. https://doi.org/10.1371/journal.pone.0078351

Sgro, G. G., Oka, G. U., Souza, D. P., Cenens, W., Bayer-Santos, E., Matsuyama, B. Y., . . . Farah, C. S. (2019). Bacteria-Killing Type IV Secretion Systems. *Front Microbiol*, *10*, 1078. https://doi.org/10.3389/fmicb.2019.01078

Sheldrick, G. M. (2010). Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr*, *66*(Pt 4), 479-485. https://doi.org/10.1107/S0907444909038360

Shimanovskaya, E., Viscardi, V., Lesigang, J., Lettman, M. M., Qiao, R., Svergun, D. I., . . . Dong, G. (2014). Structure of the C. elegans ZYG-1 cryptic polo box suggests a conserved mechanism for centriolar docking of Plk4 kinases. *Structure*, *22*(8), 1090-1104. https://doi.org/10.1016/j.str.2014.05.009

Siegrist, M. S., Unnikrishnan, M., McConnell, M. J., Borowsky, M., Cheng, T. Y., Siddiqi, N., . . . Rubin, E. J. (2009). Mycobacterial Esx-3 is required for mycobactin-mediated iron acquisition. *Proc Natl Acad Sci U S A*, *106*(44), 18792-18797. https://doi.org/10.1073/pnas.0900589106

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., . . . Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*, *7*, 539. https://doi.org/10.1038/msb.2011.75

Silverman, J.M., Agnello, D.M., Zheng, H., Andrews, B.T., Li, M., Catalano, C.E., Gonen, T., amd Mougous, J.D. (2013). Haemolysin coregulated protein is an exported receptor and chaperone of type VI secretion substrates. *Mol Cell, 51,* 584-593.

Skjøt, R. L., Oettinger, T., Rosenkrands, I., Ravn, P., Brock, I., Jacobsen, S., & Andersen, P. (2000). Comparative evaluation of low-molecular-mass proteins from Mycobacterium tuberculosis identifies members of the ESAT-6 family as immunodominant T-cell antigens. *Infect Immun*, *68*(1), 214-220. https://doi.org/10.1128/IAI.68.1.214-220.2000

Solomonson, M., Setiaputra, D., Makepeace, K. A. T., Lameignere, E., Petrotchenko, E. V., Conrady, D. G., . . . Strynadka, N. C. J. (2015). Structure of EspB from the

ESX-1 type VII secretion system and insights into its export mechanism. *Structure*, *23*(3), 571-583. https://doi.org/10.1016/j.str.2015.01.002

Souza, D. P., Oka, G. U., Alvarez-Martinez, C. E., Bisson-Filho, A. W., Dunger, G., Hobeika, L., . . . Farah, C. S. (2015). Bacterial killing via a type IV secretion system. *Nat Commun*, *6*, 6453. https://doi.org/10.1038/ncomms7453

Spencer, B. L., Tak, U., Mendonça, J. C., Nagao, P. E., Niederweis, M., & Doran, K. S. (2021). A type VII secretion system in Group B Streptococcus mediates cytotoxicity and virulence. *PLoS Pathog*, *17*(12), e1010121. https://doi.org/10.1371/journal.ppat.1010121

Stanger, F. V., de Beer, T. A. P., Dranow, D. M., Schirmer, T., Phan, I., & Dehio, C. (2017). The BID Domain of Type IV Secretion Substrates Forms a Conserved Four-Helix Bundle Topped with a Hook. *Structure*, *25*(1), 203-211. https://doi.org/10.1016/j.str.2016.10.010

Stanley, S. A., Raghavan, S., Hwang, W. W., & Cox, J. S. (2003). Acute infection and macrophage subversion by Mycobacterium tuberculosis require a specialized secretion system. *Proc Natl Acad Sci U S A*, *100*(22), 13001-13006. https://doi.org/10.1073/pnas.2235593100

Strong, M., Sawaya, M. R., Wang, S., Phillips, M., Cascio, D., & Eisenberg, D. (2006). Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A*, *103*(21), 8060-8065. https://doi.org/10.1073/pnas.0602606103

Sun, J., Siroy, A., Lokareddy, R. K., Speer, A., Doornbos, K. S., Cingolani, G., & Niederweis, M. (2015). The tuberculosis necrotizing toxin kills macrophages by hydrolyzing NAD. *Nat Struct Mol Biol*, *22*(9), 672-678. https://doi.org/10.1038/nsmb.3064

Sundaramoorthy, R., Fyfe, P. K., & Hunter, W. N. (2008). Structure of Staphylococcus aureus EsxA suggests a contribution to virulence by action as a transport chaperone and/or adaptor protein. *J Mol Biol*, *383*(3), 603-614. https://doi.org/10.1016/j.jmb.2008.08.047

São-José, C., Baptista, C., & Santos, M. A. (2004). Bacillus subtilis operon encoding a membrane receptor for bacteriophage SPP1. *J Bacteriol*, *186*(24), 8337-8346. https://doi.org/10.1128/JB.186.24.8337-8346.2004

São-José, C., Lhuillier, S., Lurz, R., Melki, R., Lepault, J., Santos, M. A., & Tavares, P. (2006). The ectodomain of the viral receptor YueB forms a fiber that triggers ejection of bacteriophage SPP1 DNA. *J Biol Chem*, *281*(17), 11464-11470. https://doi.org/10.1074/jbc.M513625200

Tak, U., Dokland, T., & Niederweis, M. (2021). Pore-forming Esx proteins mediate toxin secretion by Mycobacterium tuberculosis. *Nat Commun*, *12*(1), 394. https://doi.org/10.1038/s41467-020-20533-1

Tang, J. Y., Bullen, N. P., Ahmad, S., & Whitney, J. C. (2018). Diverse NADase effector families mediate interbacterial antagonism via the type VI secretion system. *J Biol Chem*, *293*(5), 1504-1514. https://doi.org/10.1074/jbc.RA117.000178

Tareen, A., & Kinney, J. B. (2020). Logomaker: beautiful sequence logos in Python. *Bioinformatics*, *36*(7), 2272-2274. https://doi.org/10.1093/bioinformatics/btz921

Tassinari, M., Doan, T., Bellinzoni, M., Chabalier, M., Ben-Assaya, M., Martinez, M., . . . Gubellini, F. (2020). Central role and structure of the membrane pseudokinase YukC in the antibacterial Bacillus subtilis Type VIIb Secretion System. *bioRxiv*, 2020.2005.2009.085852. https://doi.org/10.1101/2020.05.09.085852

Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., . . . Hung, L. W. (2009). Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr*, *65*(Pt 6), 582-601. https://doi.org/10.1107/S0907444909012098

Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., . . . Adams, P. D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta*

Crystallogr D Biol Crystallogr, *64*(Pt 1), 61-69. https://doi.org/10.1107/S090744490705024X

Testa, B., Carrupt, P. A., Gaillard, P., Billois, F., & Weber, P. (1996). Lipophilicity in molecular modeling. *Pharm Res*, *13*(3), 335-343. https://doi.org/10.1023/a:1016024005429

Ting, S. Y., Bosch, D. E., Mangiameli, S. M., Radey, M. C., Huang, S., Park, Y. J., . . . Mougous, J. D. (2018). Bifunctional Immunity Proteins Protect Bacteria against FtsZ-Targeting ADP-Ribosylating Toxins. *Cell*, *175*(5), 1380-1392.e1314. https://doi.org/10.1016/j.cell.2018.09.037

Ting, S. Y., Martínez-García, E., Huang, S., Bertolli, S. K., Kelly, K. A., Cutler, K. J., . . . Mougous, J. D. (2020). Targeted Depletion of Bacteria from Mixed Populations by Programmable Adhesion with Antagonistic Competitor Cells. *Cell Host Microbe*, *28*(2), 313-321.e316. https://doi.org/10.1016/j.chom.2020.05.006

Tiwari, S., Casey, R., Goulding, C. W., Hingley-Wilson, S., & Jacobs, W. R. (2019). Infect and Inject: How. *Microbiol Spectr*, *7*(3). https://doi.org/10.1128/microbiolspec.BAI-0024-2019

Tomoyasu, T., Tabata, A., Hiroshima, R., Imaki, H., Masuda, S., Whiley, R. A., . . . Nagamune, H. (2010). Role of catabolite control protein A in the regulation of intermedilysin production by Streptococcus intermedius. *Infect Immun*, *78*(9), 4012-4021. https://doi.org/10.1128/IAI.00113-10

Tran, H. R., Grebenc, D. W., Klein, T. A., & Whitney, J. C. (2021). Bacterial type VII secretion: An important player in host-microbe and microbe-microbe interactions. *Mol Microbiol*, *115*(3), 478-489. https://doi.org/10.1111/mmi.14680

Tsirigotaki, A., De Geyter, J., Šoštaric, N., Economou, A., & Karamanou, S. (2017). Protein export through the bacterial Sec pathway. *Nat Rev Microbiol*, *15*(1), 21-36. https://doi.org/10.1038/nrmicro.2016.161

Tufariello, J. M., Chapman, J. R., Kerantzas, C. A., Wong, K. W., Vilchèze, C., Jones, C. M., . . . Jacobs, W. R. (2016). Separable roles for Mycobacterium tuberculosis

ESX-3 effectors in iron acquisition and virulence. *Proc Natl Acad Sci U S A*, *113*(3), E348-357. https://doi.org/10.1073/pnas.1523321113

Tuukkanen, A. T., Freire, D., Chan, S., Arbing, M. A., Reed, R. W., Evans, T. J., . . . Korotkov, K. V. (2019). Structural Variability of EspG Chaperones from Mycobacterial ESX-1, ESX-3, and ESX-5 Type VII Secretion Systems. *J Mol Biol*, *431*(2), 289-307. https://doi.org/10.1016/j.jmb.2018.11.003

Ulhuq, F. R., Gomes, M. C., Duggan, G. M., Guo, M., Mendonca, C., Buchanan, G., . . . Palmer, T. (2020). A membrane-depolarizing toxin substrate of the. *Proc Natl Acad Sci U S A*, *117*(34), 20836-20847. https://doi.org/10.1073/pnas.2006110117

Unnikrishnan, M., Constantinidou, C., Palmer, T., & Pallen, M. J. (2017). The Enigmatic Esx Proteins: Looking Beyond Mycobacteria. *Trends Microbiol*, *25*(3), 192-204. https://doi.org/10.1016/j.tim.2016.11.004

Usón, I., Patzer, S. I., Rodríguez, D. D., Braun, V., & Zeth, K. (2012). The crystal structure of the dimeric colicin M immunity protein displays a 3D domain swap. *J Struct Biol*, *178*(1), 45-53. https://doi.org/10.1016/j.jsb.2012.02.004

van der Wel, N., Hava, D., Houben, D., Fluitsma, D., van Zon, M., Pierson, J., . . . Peters, P. J. (2007). M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells. *Cell*, *129*(7), 1287-1298. https://doi.org/10.1016/j.cell.2007.05.059

van Winden, V. J., Ummels, R., Piersma, S. R., Jiménez, C. R., Korotkov, K. V., Bitter, W., & Houben, E. N. (2016). Mycosins Are Required for the Stabilization of the ESX-1 and ESX-5 Type VII Secretion Membrane Complexes. *mBio*, *7*(5). https://doi.org/10.1128/mBio.01471-16

Vollmer, W., Blanot, D., & de Pedro, M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiol Rev*, *32*(2), 149-167. https://doi.org/10.1111/j.1574-6976.2007.00094.x

Wallace, I. M., O'Sullivan, O., Higgins, D. G., & Notredame, C. (2006). M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res*, *34*(6), 1692-1699. https://doi.org/10.1093/nar/gkl091

Wang, Q., Boshoff, H. I. M., Harrison, J. R., Ray, P. C., Green, S. R., Wyatt, P. G., & Barry, C. E. (2020). PE/PPE proteins mediate nutrient transport across the outer membrane of. *Science*, *367*(6482), 1147-1151. https://doi.org/10.1126/science.aav5912

Wheeler, T. J., Clements, J., & Finn, R. D. (2014). Skylign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics*, *15*, 7. https://doi.org/10.1186/1471-2105-15-7

Whitney, J. C., Peterson, S. B., Kim, J., Pazos, M., Verster, A. J., Radey, M. C., . . . Mougous, J. D. (2017). A broadly distributed toxin family mediates contact-dependent antagonism between gram-positive bacteria. *Elife*, *6*. https://doi.org/10.7554/eLife.26938

Whitney, J. C., Quentin, D., Sawai, S., LeRoux, M., Harding, B. N., Ledvina, H. E., . . . Mougous, J. D. (2015). An interbacterial NAD(P)(+) glycohydrolase toxin requires elongation factor Tu for delivery to target cells. *Cell*, *163*(3), 607-619. https://doi.org/10.1016/j.cell.2015.09.027

Williamson, Z. A., Chaton, C. T., Ciocca, W. A., Korotkova, N., & Korotkov, K. V. (2020). PE5-PPE4-EspG. *J Biol Chem*, *295*(36), 12706-12715. https://doi.org/10.1074/jbc.RA120.012698

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., . . . Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr*, *67*(Pt 4), 235-242. https://doi.org/10.1107/S0907444910045749

Winter, G., Lobley, C. M., & Prince, S. M. (2013). Decision making in xia2. *Acta Crystallogr D Biol Crystallogr*, *69*(Pt 7), 1260-1273. https://doi.org/10.1107/S0907444913015308

Xu, J., & Zhang, Y. (2010). How significant is a protein structure similarity with TM-score = 0.5? *Bioinformatics*, *26*(7), 889-895. https://doi.org/10.1093/bioinformatics/btq066

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nat Methods*, *12*(1), 7-8. https://doi.org/10.1038/nmeth.3213

Yu, W., Hallinen, K. M., & Wood, K. B. (2018). Interplay between Antibiotic Efficacy and Drug-Induced Lysis Underlies Enhanced Biofilm Formation at Subinhibitory Drug Concentrations. *Antimicrob Agents Chemother*, *62*(1). https://doi.org/10.1128/AAC.01603-17

Zhang, D., de Souza, R. F., Anantharaman, V., Iyer, L. M., & Aravind, L. (2012). Polymorphic toxin systems: Comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. *Biol Direct*, *7*, 18. https://doi.org/10.1186/1745-6150-7-18

Zhang, K., Riba, A., Nietschke, M., Torow, N., Repnik, U., Pütz, A., . . . Hornef, M. (2018). Minimal SPI1-T3SS effector requirement for Salmonella enterocyte invasion and intracellular proliferation in vivo. *PLoS Pathog*, *14*(3), e1006925. https://doi.org/10.1371/journal.ppat.1006925

Zhang, Y., & Skolnick, J. (2004). Scoring function for automated assessment of protein structure template quality. *Proteins*, *57*(4), 702-710. https://doi.org/10.1002/prot.20264

Zoltner, M., Ng, W. M., Money, J. J., Fyfe, P. K., Kneuper, H., Palmer, T., & Hunter, W. N. (2016). EssC: domain structures inform on the elusive translocation channel in the Type VII secretion system. *Biochem J*, *473*(13), 1941-1952. https://doi.org/10.1042/BCJ20160257

Zoltner, M., Norman, D. G., Fyfe, P. K., El Mkami, H., Palmer, T., & Hunter, W. N. (2013). The architecture of EssB, an integral membrane component of the type VII secretion system. *Structure*, *21*(4), 595-603. https://doi.org/10.1016/j.str.2013.02.007

Zoued, A., Brunet, Y. R., Durand, E., Aschtgen, M. S., Logger, L., Douzi, B., . . . Cascales, E. (2014). Architecture and assembly of the Type VI secretion system.

Biochim Biophys Acta, *1843*(8), 1664-1673.
https://doi.org/10.1016/j.bbamcr.2014.03.018