TYPE VI SECRETION SYSTEM EFFECTOR TRANSPORT AND TOXICITY

MECHANISMS OF TYPE VI SECRETION SYSTEM EFFECTOR TRANSPORT AND TOXICITY

By SHEHRYAR AHMAD, B.H.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 (2021) Hamilton, Ontario (Biochemistry and Biomedical Sciences)

TITLE: Mechanisms of type VI secretion system effector transport and toxicity

AUTHOR: Shehryar Ahmad, B.H.Sc.

SUPERVISOR: John C. C. Whitney, Ph.D.

NUMBER OF PAGES: xv; 257

FOREWARD

Lay Abstract

Bacteria constantly compete with their neighbours for resources and space. The type VI secretion system is a protein complex that facilitates competition between Gram-negative bacteria by facilitating the injection of protein toxins, also known as effectors, from attacking cells into target cells. In this work, I characterize several members of a large family of membrane protein effectors. First, I showed that these effectors require a novel family of chaperone proteins for stability and recruitment to the type VI secretion system apparatus. Second, I characterized the growth-inhibitory properties of one of these effectors in-depth and showed that it possesses a toxin domain that depletes the essential nucleotides ADP and ATP in target cells by synthesizing the nucleotides adenosine penta- and tetraphosphate, (p)ppApp. Together, these studies revealed a new mechanism for the intercellular delivery of membrane protein toxins and uncovered the first known physiological role of a (p)ppApp-synthesizing enzyme in bacteria.

Abstract

The type VI secretion system (T6SS) is a protein export pathway that mediates competition between Gram-negative bacteria by facilitating the injection of toxic effector proteins from attacking cells into target cells. To function properly, many T6SSs require at least one protein that possesses a proline-alanine-alanine-arginine (PAAR) domain. These PAAR domains are often found within large, multi-domain effectors that possess additional N- and C-terminal extension domains whose function in type VI secretion is not well understood. The work described herein uncovers the function of these accessory domains across multiple PAAR-containing effectors.

First, I demonstrated that thousands of PAAR effectors possess N-terminal transmembrane domains (TMDs) and that these effectors require a family of molecular chaperones for stability in the cell prior to their export by the T6SS. Our findings are corroborated by co-crystal structures of chaperones in complex with the TMDs of their cognate effectors, capturing the first high-resolution structural snapshots of T6SS chaperone-effector interactions.

Second, I characterize a previously undescribed prePAAR effector named Tas1. My work shows that the C-terminus of Tas1 possesses a toxin domain that pyrophosphorylates ADP and ATP to synthesize the nucleotides adenosine penta- and tetraphosphate (hereafter referred to as (p)ppApp). Delivery of Tas1 into competitor cells drives the rapid accumulation of (p)ppApp, depletion of ADP and ATP, and widespread dysregulation of essential metabolic pathways, resulting in target cell death. These findings reveal a new mechanism of interbacterial antagonism, the first characterization of a (p)ppApp synthetase and the first demonstration of a role for (p)ppApp in bacterial physiology.

TMD- and toxin-containing PAAR proteins constitute a large family of over 6,000 T6SS effectors found in Gram-negative bacteria. My work on these proteins has uncovered that different regions found within effectors have distinct roles in trafficking between bacterial cells and in the growth inhibition of the target cell.

Acknowledgements

To my supervisor, John, thank you for the lessons and unwavering support over the past four years. You are an extremely patient mentor and have put up with thousands of my questions. During moments of doubt, your confidence in me and our work has encouraged me to keep trying. You have not only guided me but have helped me build my scientific creativity by letting me explore new and challenging projects. You are a strong leader who has fostered camaraderie, collaboration, and focus even during the toughest of times. Thank you for your hard work and for an incredible graduate school experience.

To my committee members, Lori and Eric, thank you for taking the time to provide helpful insights for my projects and pushing me to think critically about how to compose a more complete scientific story.

To my labmates, Tim and Nathan, thanks for fun times and the extensive scientific discussions. I could not have asked for a better pair of creative and passionate scientists to do this degree with. To my other friends in the IIDR, Wael, Kara, Liz and Caressa, thank you all for a very fun past four years. You all have allowed the lab to be a place of both hard work and silly banter. It has been a fun journey and I hope I can be lucky enough to find a group like you in the future. Wael, thank you for your wise counsel and being a positive influence on me and the lab.

To my collaborators, Boyuan and Gerd, thank you for your efforts on our projects and for significantly broadening my scope of knowledge and inquiry. I have no doubt both of you will be doing incredible work with your own labs in the future.

To Lisa Kush and the other staff members of the Biochemistry Department, thank you so much for your hard efforts. We do not get to say it enough.

To the students I have mentored in the lab especially, Tahmid, Sarah and Vanessa, thank you for your enthusiasm and dedication to the lab.

To Heidi, I am so lucky to have met you at the start of my graduate degree. I could not have asked for a better person to share the past few years with. I have never met anyone who works as hard as you do and can still be so optimistic and cheerful every day. Thank you for your friendship, love and support.

To my family and Savar, thank you all for your ongoing support over the past few years. You have all dealt with my complaints, my frustrations, kept me grounded and helped me get out of the lab. You all give me such a strong support network that I am so grateful for. I feel confident that I can take on any challenge with you guys by my side. Savar, thanks for being such a great friend since we met in our first year of undergrad. Your practical insights have helped me to see the bigger picture to so many problems and your positive impact on my undergraduate and graduate performance can't be understated. Thanks for always being there to listen and give me your input. Ibrahim, thanks for letting be your mentor and letting

me re-learn all of grade 11, grade 12 and first-year chemistry. I have no doubt you will be doing some incredible things in the future. Maha, you know better than anyone how terrible I can be at maneuvering my personal life. Thanks for always taking the time to give me advice on how to manage personal issues and lightening my frustrations with our hilarious inside jokes. Dad, thank you for always keeping me on track. Your love for the sciences certainly wore off on me. You guided me in the right directions from a young age and I am lucky to have had you by my side during this degree. Thank you for your unwavering commitment to my success. Mom, thank you for everything. I am blessed to have such a strong, intelligent and dedicated mother. I have no idea how you managed to raise three kids, take care of our home, work a full-time job and still let me complain to you about my life at the end of the day, for years on end. You have sacrificed too much to keep me, Maha and Ibrahim healthy, out of trouble and focused on our work. Not that I will ever be able to repay you for your love and care, but the completion of this thesis is to demonstrate that your efforts will never go to waste and will always be cherished.

Table of Contents

List of Figures

Chapter I Figures

Chapter II Figures

Chapter III Figures

Appendix Figures

List of Tables and Datasets

Chapter II

Chapter III

Appendix Dataset

List of Abbreviations

T6SS ..... Type VI secretion system
T6S ..... Type VI secretion
DUF ..... Domain of unknown function
Eag ..... Effector-associated gene
PAAR ..... Proline-alanine-alanine-arginine
Hcp ..... Hemolysin co-regulated protein
VgrG ..... Valine-glycine repeat protein G
Tse ..... Type VI exported
TMD ..... Transmembrane domain
Rhs ..... Rearrangement hotspot
Tas ..... Type VI secretion effector (p)ppApp synthetase
Apk ..... Adenosine 3'-pyrophosphokinase
SAS ..... Small alarmone synthetase
SAH ..... Small alarmone hydrolase
RSH ..... RelA-SpoT homolog
ATP ..... Adenosine 5'-triphosphate
ADP ..... Adenosine 5'-diphosphate
AMP ..... Adenosine 5'-monophosphate
pppApp ..... Adenosine 3'-diphosphate 5'-triphosphate
ppApp ..... Adenosine 3',5'-diphosphate
pApp ..... Adenosine 3'-diphosphate 5'-monophosphate
GTP ..... Guanosine 5'-triphosphate
GDP ..... Guanosine 5'-diphosphate
pppGpp ..... Guanosine 3'-diphosphate 5'-triphosphate
ppGpp ..... Guanosine 3',5'-diphosphate
NAD^+ ..... Nicotinamide adenine dinucleotide
$NAD(P)^+$ ..... Nicotinamide adenine dinucleotide phosphate
LB ..... Lysogeny broth
CFU ..... Colony-forming units
Cryo-EM ..... Cryogenic electron microscopy
His_6 ..... Hexahistidine
His_{10} ..... Decahistidine
VSV-G ..... Vesicular stomatitis virus G

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

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

Klein, T.A.*, **Ahmad, S.***, and Whitney, J.C. (2020). Contact-Dependent Interbacterial Antagonism Mediated by Protein Secretion Machines. Trends Microbiol *28*, 387-400.

*indicates equal contribution

Author contributions: T.A.K, S.A. and J.C.W. wrote and edited manuscript and designed all figures.

Mechanisms of interbacterial competition

Overview of interbacterial competition

Bacteria commonly reside in dense, surface-associated polymicrobial communities (Flemming and Wuertz, 2019; Nadell et al., 2016). Different members of these communities often antagonize one another in order to establish and maintain a niche while facing pervasive competition for space and resources (Cornforth and Foster, 2013). These conditions, which have existed for many millions of years, have provided the necessary selective pressure for bacteria to develop sophisticated mechanisms of nutrient acquisition, sensing and responding to abiotic environmental stress, and inhibiting the growth of or kill nearby microbes through the use of toxic molecules (Garcia-Bayona and Comstock, 2018; Peterson et al., 2020). While many mechanisms involved in the acquisition of nutrients and response to stress have been studied for decades, our knowledge of pathways that mediate interbacterial antagonism is comparatively limited. However, our understanding of how these pathways is improving as advances in next-generation sequencing and structural biology techniques have led to the rapid accumulation of genomic data for thousands of bacteria and enabled determination of macromolecular protein structures at atomic resolution, respectively (Bai et al., 2015; Koboldt et al., 2013). Together, these advances provide a platform for molecular microbiologists to study the architecture, regulation and function of these pathways in shaping complex microbial communities (Costa et al., 2015; Galán and Waksman, 2018).

One way pathways involved interbacterial antagonism can be classified is by whether they function in a contact-independent or contact-dependent manner (Garcia-Bayona and Comstock, 2018). The former classification involves the biosynthesis and secretion of small, diffusible antimicrobials, including antibiotics and bacteriocins that act on target cells at a distance from the toxin-producing cell. The earliest examples of contact-independent antagonism were found in *Streptomyces* spp., which secrete antibiotics to inhibit the growth of a wide range of bacteria (van der Heul et al., 2018). By contrast, contact-dependent mechanisms typically involve the secretion of protein toxins from attacking cells directly into recipient cells through the use of a protein secretion system (Klein et al., 2020). Almost all well-characterized bacteria contain genes involved in either or both contact-independent and contact-dependent antagonism (Peterson et al., 2020; Zhang et al., 2012). In some organisms, the collection of genes encoding these pathways can constitute a significant portion of the total coding capacity of a cell, with notable examples being as high 2-3% (Peterson et al., 2020; Schell et al., 2007). While diffusible antimicrobials may lose their efficacy at range due to dilution effects or may be resisted or detoxified by some competitors, the presence of contact-dependent mechanisms provides an alternative means to rapidly suppress the growth of competing cells and, in some cases, lyse these cells to acquire nutrients and/or acquire their DNA (Borgeaud et al., 2015; Hibbing et al., 2010; Peterson et al., 2020; Smith et al., 2020). Thus, harboring a collection of non-redundant pathways to inhibit the growth of competitors endows bacteria with a

range of strategies to target a diverse range of competing cell types thus allowing them to adapt to constantly changing environmental conditions.

Bacterial secretion systems that mediate interbacterial antagonism

The involvement of secretion systems and their exported protein toxins in interbacterial antagonism is an active and major area of ongoing research in microbiology. To better understand these pathways, a general discussion on protein secretion systems in bacteria provides a good starting point. Protein secretion systems, which are defined by their ability to actively transport protein substrates from within a cell to outside the cell (Desvaux et al., 2009), are typically comprised of dozens of structural proteins for proper function and once assembled in a bacterial cell can serve a number of biological roles that includes but is not limited to interbacterial competition (Costa et al., 2015).

Bacteria possess two general protein secretion pathways known as Sec and Tat, which are responsible for the export of the majority of secreted proteins across the cytoplasmic membrane in both Gram-negative and Gram-positive bacteria (Lee et al., 2006b; Tsirigotaki et al., 2017). These proteins are often involved in processes such as nutrient acquisition and maintenance of the bacterial cell envelope. By contrast, so-called ‘specialized’ secretion systems often mediate the interaction of bacteria with other organisms. Currently, nine specialized secretion systems known as the type I-IX secretion systems (often denoted as T1SS through to T9SS) have been characterized in bacteria (Costa et al., 2015). These systems

are exclusively found in Gram-negative bacteria with the exception of the Gram-positive T7SS (Bunduc et al., 2020; Tran et al., 2021). The barrier presented by the outer membrane of Gram-negative bacteria renders the Sec and Tat systems insufficient for protein secretion to the extracellular milieu. To overcome this, the T2SS and T5SS facilitate the outer membrane translocation of Sec- and Tat-delivered periplasmic proteins (Costa et al., 2015). The other Gram-negative specialized secretion systems form large, macromolecular structures that span both the inner and outer membrane and facilitate a one-step translocation process in which secreted proteins bypass the periplasmic space. The minimum requirements for each of these systems is an ATPase that generates the energy required for the export of substrate proteins, a receptor or chaperone for substrate recognition and a channel that guides substrate transit (Costa et al., 2015). Some one-step specialized secretion systems, such as the T1SS and T9SS, transport proteins across the Gram-negative cell envelope and into the extracellular milieu, whereas others, such as the T3SS, T4SS and T6SS, directly 'inject' proteins into other cell types (Galán and Waksman, 2018). The latter group of secreted proteins are more commonly referred to as effectors and exert their functions upon being injected into their target cell type (Galán and Waksman, 2018).

Of the nine known specialized secretion systems, five have been shown to facilitate the transport of toxic effectors between competing bacteria in a contact-dependent manner, including the T1SS (Garcia-Bayona et al., 2017), T4SS (Bayer-Santos et al., 2019; Souza et al., 2015), T5SS (Aoki et al., 2010; Aoki et al., 2005),

T6SS (Hood et al., 2010) and T7SS (Cao et al., 2016; Whitney et al., 2017). While all these systems secrete toxins that inhibit the growth of competitors, the substrates, regulation and architecture of each complex vary significantly (Costa et al., 2015). A brief overview of each pathway is provided below.

Most characterized T1SSs are involved in nutrient acquisition and host pathogenesis (Kanonenberg et al., 2013; Noegel et al., 1979). Many bacteria possess more than one T1SS, facilitating the secretion of a diverse range of substrate proteins. The role of these systems in contact-independent interbacterial competition has also been previously established, as many T1SSs secrete diffusible bacteriocins (De Kwaadsteniet et al., 2006). The role of T1SS in bacterial competition was recently expanded with the discovery of a contact-dependent inhibition by glycine zipper (Cdz) T1SS in *Caulobacter crescentus* (Garcia-Bayona et al., 2019). Glycine-zipper motifs found in the CdzCD effectors exported by this T1SS enable their aggregation on the donor cell surface, which forms fibrillar structures that are toxic to cells in close contact with a CdzCD-producing cell. Homologs of the Cdz T1SSs are distributed across many species of Proteobacteria, suggesting that this pathway may be a broadly conserved mechanism of interbacterial competition (Garcia-Bayona et al., 2019).

In contrast to T1SSs, the primary ascribed functions of T4SSs are to facilitate conjugation by transporting DNA between bacterial cells or in the uptake and release of DNA from the environment (Dillard and Seifert, 2001; Hofreuter et al., 2001). Some T4SSs are involved in pathogenesis and function to transport

protein effectors or protein-DNA complexes into target cells, which leads to disease (Seubert et al., 2003; Vergunst et al., 2000). By contrast, the recently characterized T4SSs in *Xanthomonas citri* and *Stenotrophomonas maltophilia* have been shown to secrete toxic effectors that kill susceptible bacterial competitors (Bayer-Santos et al., 2019; Souza et al., 2015). These systems appear to be distributed broadly in Xanthomonadales and the β-Proteobacteria, however, it is currently unknown whether these systems differ significantly in architecture from DNA- or other protein-transporting T4SSs mentioned above.

Similar to T1SSs, T5SSs are broadly distributed and have a large substrate diversity, with many T5SS-exported proteins shown to be important virulence factors and/or are involved in nutrient acquisition (Leo et al., 2012). As mentioned previously, T5SSs are two-step secretion pathways that require their substrates to be delivered into the periplasm by the general secretory pathway prior to their secretion from the cell via a T5SS. In general, T5SSs consist of an outer membrane transporter that recognizes a periplasmic substrate and facilitates its secretion out of the cell (Costa et al., 2015). The first example of interbacterial toxin exchange by a T5SS was shown for the CdiB-CdiA two-partner type V secretion system (TPS), which facilitates competition between closely related strains of *E. coli* (Aoki et al., 2005). CdiB is a β-barrel protein that localizes to the outer membrane. The periplasmic domain of CdiB recognizes and facilitates the transport of the large (>300 kDa), filamentous effector CdiA to the cell surface (Aoki et al., 2010). Most of the CdiA effector is secreted from the cell and remains on the cell surface until

a receptor-presenting recipient cell is recognized by a CdiA effector receptor binding domain. Upon receptor recognition, the remaining part of CdiA, which is typically a periplasmically-localized toxin domain, is secreted and transported into the recipient cell (Ruhe et al., 2018).

The bacterial T6SS is thought to function as a molecular syringe that 'injects' a payload of effector proteins that inhibit the growth of or lyse target cells. Unlike the other systems discussed above, the vast majority of T6SSs evolved to target bacteria with notable exceptions including the host cell-targeting T6SSs from *Francisella tularensis*, *Edwardsiella tarda* and several *Burkholderia* species (Eshraghi et al., 2016; Nano et al., 2004; Rao et al., 2004; Rosales-Reyes et al., 2012; Schell et al., 2007; Schwarz et al., 2010b). The T6SS is primarily found in Proteobacteria but has recently also been shown to be functionally relevant in multiple species of Bacteroidetes (Ross et al., 2019; Russell et al., 2014b; Wexler et al., 2016). The effectors secreted by the pathway typically have diverse functions and target a broad host range that in some cases, allow for competition between bacteria from different phyla (Klein et al., 2020).

Unlike the aforementioned systems, T7SSs are found in Gram-positive bacteria. There are two divergent T7SSs distributed across bacteria: T7SSa and T7SSb. The T7SSa is distributed primarily in the phylum Actinobacteria and has been extensively studied in *Mycobacterium tuberculosis* for its role as an essential virulence determinant in this pathogen (Lewis et al., 2003; Mahairas et al., 1996; Pym et al., 2002). The genetically distinct T7SSb, however, is broadly distributed

among Firmicutes bacteria. The T7SSb system has been shown to mediate pathogenesis by some bacterial species but also interbacterial competition by others (Cao et al., 2016; Taylor et al., 2021; Whitney et al., 2017). The two systems, T7SSa and T7SSb, share two core structural components, while the effectors that transit each system, the structure of each system and their regulation differ significantly (Tran et al., 2021).

While all the above secretion systems differ in their structure, distribution and target spectrum, they are united by two key attributes: 1) each system is associated with toxic effector proteins that exert their activities on essential molecules of target cells such as nucleic acid polymers, nucleotides, the cell wall and cell membranes and 2) antibacterial effectors are encoded adjacent to immunity genes, which encode for proteins that neutralize the effector, protecting toxin-producing cells from self-intoxication and permitting kin discrimination (Klein et al., 2020; Russell et al., 2014a).

Of these systems, the T6SS is most commonly ascribed a role in contact-dependent bacterial competition. The T6SS is the primary focus of this dissertation and a detailed overview of its components, effectors and accessory proteins has been provided below.

Components and mechanisms of the T6SS

Discovery and biological roles of the T6SS

Early work on virulence-associated T4SSs identified intracellular multiplication (*icm*) genes in this pathway that are required for pathogen replication within macrophages (Purcell and Shuman, 1998). A key component of this pathway is the IcmF protein, a conserved T4SS structural component, which is distributed across many Gram-negative bacteria (Das et al., 2000; Folkesson et al., 2002). Advancements in genomic sequencing at the time enabled several large-scale in silico studies of this pathway. One study found many IcmF-containing clusters that did not contain known type IV or other secretion system components, called IcmF-associated homologous protein clusters (IAHPs) (Das and Chaudhuri, 2003). Soon after this study, two IAHPs, one in *Vibrio cholera* and one in *Pseudomonas aeruginosa*, were characterized and both shown to secrete the Hemolysin co-regulated protein (Hcp) (Mougous et al., 2006; Pukatzki et al., 2006). These and many follow-up studies indicated that IAHPs constitute a novel bacterial secretion system, the T6SS, which possesses distinct protein machinery from other characterized secretion systems and exports a novel range of substrates that lack a signal sequence (Dudley et al., 2006; Pukatzki et al., 2007; Zheng and Leung, 2007).

In line with the suspected similarities between the T4SSs and T6SSs, the first ascribed function of the T6SS was to mediate pathogenesis, with notable examples in *Burkholderia*, *Edwardsiella* and *Francisella* species (Barker et al.,

2001; Pukatzki et al., 2006; Schell et al., 2007). However, beyond these genera there was limited evidence to suggest that other T6SSs were taking part in pathogenesis directly (Schwarz et al., 2010a). Structural studies of the T6SS components Hcp and the Valine-glycine repeat protein G (VgrG) revealed their significant structural similarity to bacteriophage tube and tailspike components, respectively (Leiman et al., 2009; Pell et al., 2009). The structural similarities to phage, genomic distribution of T6SSs in environmental strains and inconsistencies in data suggesting a direct role in pathogenesis led to the hypothesis many T6SSs may facilitate interactions with other bacteria. This hypothesis was confirmed with the discovery of the first antibacterial effector, Tse2, which was shown to be secreted by *P. aeruginosa* into neighbouring bacteria (Hood et al., 2010). Many subsequent studies corroborated these findings in other bacteria, underscoring a major function of the T6SS in mediating interbacterial interactions (Schwarz et al., 2010a). Today, three broad T6SS families have been identified primarily across Proteobacteria, Bacteroidetes and Acidobacteria species (Russell et al., 2014b). All these pathways secrete diverse effector proteins that enable targeting of a broad range of Gram-negative bacteria (Peterson et al., 2020). Collectively, these systems contribute to the lifestyles of many bacteria and have been shown to have a significant impact on the health and disease of animals and plants (Garcia-Bayona and Comstock, 2018; Peterson et al., 2020). Below, a detailed discussion on the components of the T6SS and additional details on its effector proteins is provided.

Components of the T6SS

The T6SS requires 14 structural components, 13 of which are designated type six secretion (Tss) A-M and the fourteenth is referred to as "PAAR" to reflect a proline-alanine-alanine-arginine repeat within its sequence. These structural proteins constitute two subassemblies known as the membrane complex and the bacteriophage tail-like complex, the latter of which resembles an inverted phage tail when completely assembled (Figure 1.1) (Cherrak et al., 2018; Rapisarda et al., 2019; Shneider et al., 2013). The membrane complex is critical for protein export by T6SS because it facilitates the assembly of the phage tail-like complex and acts as a conduit for effector delivery across the cell envelope. The bacteriophage tail-like subassembly consists of six components: the AAA+ ATPase ClpV (TssH), the phage tail sheath-like proteins TssB and TssC, Hemolysin co-regulated protein (Hcp/TssD), Valine-glycine repeat protein G (VgrG/TssI) and a PAAR repeat protein. Hcp protomers form hexameric rings that stack on top of one another to form a long nanotube (Ballister et al., 2008). This nanotube interacts with a VgrG trimer which possesses a flat surface that facilitates its interaction with a single PAAR repeat protein (Renault et al., 2018) (Figure 1.1).

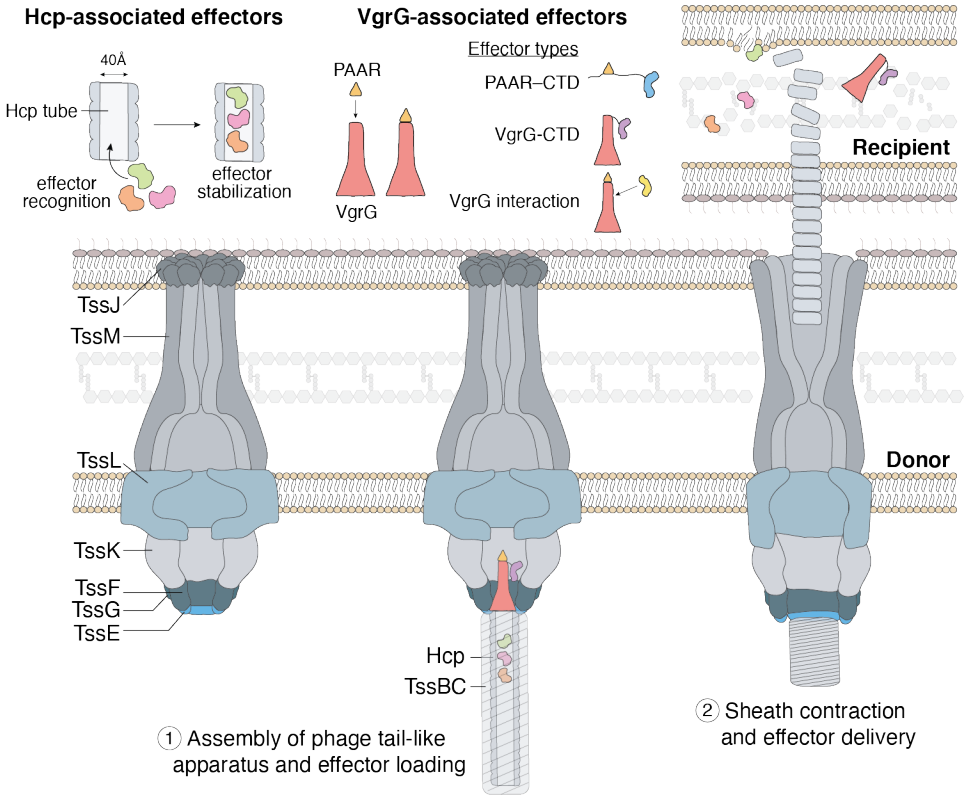


Figure 1.1. Schematic of the type VI secretion system and its proposed mechanism of effector export. A generic model for bacteria-targeting T6SSs. Upon detection of a recipient cell, a needle-like apparatus (Hcp tube, VgrG+PAAR spike, TssBC sheath) is assembled and loaded with effectors. Effectors either interact with Hcp hexamers or VgrG-PAAR complexes. Hcp-associated effectors can be found to (i) associate with the inside of the Hcp tube or (ii) form a C-terminal extension of an Hcp hexamer. VgrG- associated effectors are found as (i) a C-terminal extension of PAAR, (ii) a C-terminal extension of VgrG, or (iii) noncovalent interaction with VgrG. Sheath contraction ejects an Hcp tube directly into the target cell, where effectors are released and exert their toxic effects.

In order to deliver effectors to target cells, T6SSs undergo a phage tail-like contraction. Prior to this contraction event, the Hcp tube is capped with a single spike-shaped VgrG-PAAR complex (Cianfanelli et al., 2016a; Shneider et al., 2013). This tube-spike complex is then surrounded by a phage tail sheath-like structure that upon contraction, propels the spike-capped tube out of the bacterial

cell (Basler et al., 2012). The ClpV ATPase then disassembles the sheath following its contraction so that the individual sheath subunits can be reassembled for a subsequent firing event (Bonemann et al., 2009). The phage tail-like assembly is the site of T6SS effector recruitment and there are currently five known ways in which effectors are delivered by this structure: 1) within the lumen of Hcp hexamers, 2) as C-terminal extensions of Hcp, 3) through interaction with VgrG proteins, 4) as C-terminal extensions of VgrG proteins, and 5) as C-terminal extensions of PAAR repeat containing proteins (Flaugnatti et al., 2016; Ma et al., 2017; Pukatzki et al., 2007; Shneider et al., 2013; Silverman et al., 2013). The inner pore of the Hcp tube is approximately 40 Å in diameter and has a chaperone-like function, interacting with and promoting the stability of effectors below a specific size threshold (approx. 20 kDa) (Silverman et al., 2013). In contrast, effectors that interact non-covalently with or are C-terminal extensions of VgrG or PAAR proteins do not appear to have similar size restrictions with some effectors having a molecular weight in excess of 100 kDa (Alcoforado Diniz and Coulthurst, 2015).

Effectors of the T6SS

Overview of effectors of the T6SS

Characterized T6SSs are known to deliver “cocktail” of effectors into target cells. Intriguingly, some bacteria, such as *Pseudomonas aeruginosa* and *Burkholderia pseudomallei* possess multiple evolutionarily distinct T6SSs that are differentially regulated, with each exporting their own repertoire of effectors

(Allsopp et al., 2017; Goodman et al., 2004; Hood et al., 2010; Marden et al., 2013; Russell et al., 2013; Schell et al., 2007; Wood et al., 2019a). Similar to bacteria-targeting effectors from other systems, T6SS effectors are typically encoded by genes adjacent to their cognate immunity genes, which allows for kin discrimination during interbacterial competition (Hood et al., 2010; Russell et al., 2011). Two general categories of effectors have been identified: 1) small, single-domain effectors that only possess a toxin domain and 2) large, multi-domain effectors that contain domains that mediate various functions, typically involving effector recruitment to the T6SS and toxicity in recipient cells (Figure 1.2A). In general, members of T6SS effector families act on molecules that are essential for survival and act in either the periplasm or the cytoplasm of target cells (Figure 1.2B). Effectors that have been characterized biochemically are described in detail below.

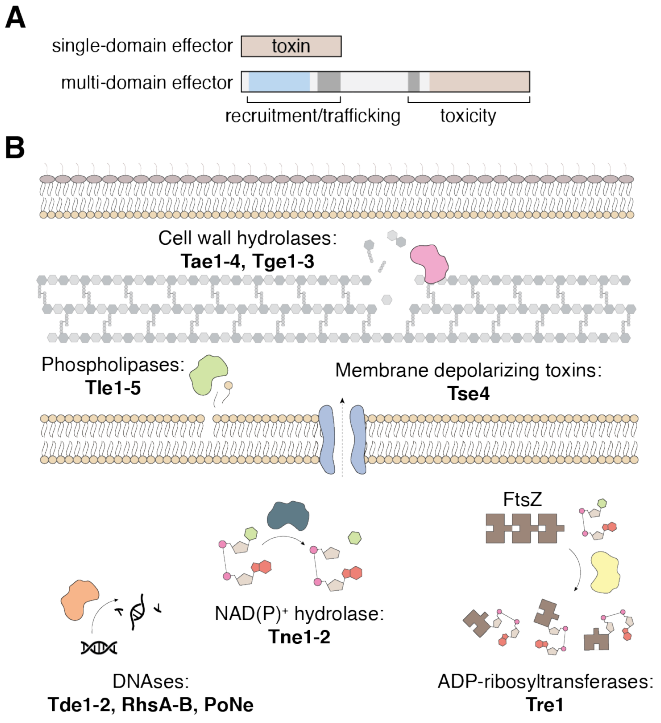


Figure 1.2. Effectors exported by the T6SS inhibit bacterial growth through diverse mechanisms. A) T6SS effectors range in sizes. Small, single-domain effectors typically require the Hcp tube for export (top). Whereas large multi-domain effectors are modular and have domains that permit both recruitment to the T6SS and toxicity in target cells (bottom). **B)** Antibacterial effectors targeting Gram-negative bacteria act in either the periplasm or cytoplasm to kill or hinder the growth of competitors. Cytoplasmic effector functions include DNase, $NAD(P)^+$ hydrolase and ADP-ribosyltransferase activities. Periplasmic effectors target one of two essential structures: (i) the cell wall or (ii) the cell membrane. Representative effectors for each category are indicated.

Cell wall-targeting T6SS effectors

Peptidoglycan (PG) maintains bacterial cell shape and protects cells from osmotic lysis, making it essential for survival (Typas et al., 2011). The polysaccharide component of PG is composed of repeating β-1,4-linked *N*-acetylmuramic acid (MurNAc) and *N*-acetylglucosamine (GlcNAc) subunits. The MurNAc groups in adjacent glycan strands are connected by peptide crosslinks resulting in a sacculus that encompasses the entirety of the cell. To date, two superfamilies of PG-targeting T6SS effectors have been identified: 1) the Type VI amidase effectors (Tae) and 2) the Type VI glycoside hydrolase effectors (Tge) (Russell et al., 2012; Whitney et al., 2013). Tae effectors exert their toxicity by hydrolyzing the peptide crosslinks of PG. Four sequence-divergent Tae families exist, and the PG peptidase activity of these families differs because they target different chemical bonds within the peptide stem. Tae1 and Tae4 effectors function as DL-endopeptidases whereas Tae2 and Tae3 are DD-endopeptidases (Russell et al., 2012). The best characterized Tae effector is the Type six exported effector 1 (Tse1) from *P. aeruginosa*, which belongs to the Tae1 family (Russell et al.,

2011). X-ray crystal structures of Tse1 show that this enzyme adopts a NlpC/P60 papain-like cysteine protease fold (Chou et al., 2012). Tse1 possesses some structural similarity to housekeeping PG amidases involved in cell wall homeostasis; however, in contrast to these enzymes, Tse1 lacks structural motifs that regulate its PG hydrolase activity – a feature that likely contributes to its potent antibacterial activity.

In contrast to Tae effectors, Tge effectors act on the glycan backbone of PG. The Tge superfamily includes enzymes targeting either the GlcNAc-MurNAc bond (glucosaminadases) or the MurNAc-GlcNAc bond (muramidases). The best characterized Tge muramidase is Tse3, the founding member of the Tge1 family (Russell et al., 2011). Tse3 adopts a lysozyme-like fold and requires a calcium cofactor for its glycan hydrolase activity (Lu et al., 2014a). A Tge2 member from *Pseudomonas protegens* has also been structurally characterized, revealing its resemblance to PG glucosaminadases (Whitney et al., 2013). There are several reports indicating that both Tae and Tge effectors are secreted by the same T6SS, which would enable the complete degradation of the PG matrix of the target cell (Russell et al., 2011; Russell et al., 2012; Whitney et al., 2013).

Cell membrane-targeting T6SS effectors

There are several T6SS effectors known to target cellular membranes and these can be classified as either 1) phospholipases or 2) pore-forming toxins. Similar to cell wall-targeting effectors, membrane-targeting effectors act in the

periplasm and are neutralized by immunity proteins that localize to this compartment. Phospholipase effectors are widely distributed among many T6SS-containing bacteria and are comprised of five sequence-divergent families designated Tle1-5 (Type VI lipase effector 1-5). The Tle1-4 families all contain a conserved GxSxG motif most commonly found in esterases, whereas Tle5 possesses dual HxKxxxD motifs characteristic of phospholipase D enzymes (Russell et al., 2013). Structural analyses of Tle1, Tle4 and Tle5 enzymes showed that these effectors adopt folds that are similar to the catalytic domains of well characterized phospholipases, but also revealed novel α-helical membrane anchoring domains that have been postulated to improve toxin efficacy through interactions with the inner membrane once inside a recipient cell (Aloulou et al., 2012; Hu et al., 2014; Lu et al., 2014b; Yang et al., 2017).

In addition to enzymatic hydrolysis of membrane phospholipids, T6SS membrane-targeting effectors can act by forming pores that puncture membranes of target cells. Some of these effectors bear homology to pore-forming colicins while others are similar to CdzCD in that they possess glycine zipper motifs reminiscent of those found in multimeric membrane protein channels and amyloid-β-peptide (Kim et al., 2005; LaCourse et al., 2018; Miyata et al., 2013). The best understood example is the Tse4 effector from *P. aeruginosa* which possesses glycine zipper motifs and exerts its toxicity in the periplasm of target bacteria by forming ion-selective membrane pores that facilitate potassium efflux, resulting in membrane depolarization and dissipation of the proton motive force (LaCourse et

al., 2018). Similarly, the Ssp6 effector from *S. marcescens* was recently shown to inhibit bacterial growth by forming cation-selective pores in membranes (Mariano et al., 2019).

Nuclease and nucleotide-targeting T6SS effectors

Though the mechanisms by which T6SS effectors that act in the cytoplasm reach this cellular compartment are incompletely understood, it is well established that molecules such as nucleic acids and nucleotides are targeted by these toxins (Ho et al., 2017; Quentin et al., 2018; Vettiger and Basler, 2016). For example, several families of nuclease toxins have been identified and shown to inhibit the growth of the target bacterial cells by degrading chromosomal DNA (Jana et al., 2019; Koskiniemi et al., 2013; Ma et al., 2014). The molecular basis for this enzymatic activity is not well understood but current evidence suggests that many of these effectors act as non-specific DNases (Alcoforado Diniz and Coulthurst, 2015; Jana et al., 2019; Ma et al., 2014). In contrast, nucleotide-targeting effectors deplete cells of high energy molecules required for cell viability. Thus far, the primary class of effectors of this class are NAD^+/$NADP^+$ ($NAD(P)^+$) hydrolases.

$NAD(P)^+$ hydrolases studied to date belong to one of two Type VI NADase effector (Tne) families (Tang et al., 2018). The best characterized $NAD(P)^+$ hydrolase and founding member of the Tne1 group is the Tse6 toxin from *P. aeruginosa* (Whitney et al., 2015). The structure of the Tse6 toxin domain resembles protein-targeting ADP-ribosyltransferases; however, in contrast to these

transferases, Tse6 lacks an open active site that would permit transfer of ADP-ribose from NAD^+ to a protein target. Instead, Tse6 degrades $NAD(P)^+$ at a rate that is approximately 1000-fold higher than structurally similar enzymes with ADP-ribosyltransferase activity. A second family of $NAD(P)^+$ hydrolase effectors was more recently described and a representative member from *Pseudomonas protegens* was shown to possess enzymatic properties similar to Tse6 (Tang et al., 2018).

Protein-targeting T6SS effectors

The majority of protein toxins that target bacteria function by degradation or depletion of molecules essential for cell viability. Less common are antibacterial toxins that act by modulating the function of a target protein, presumably because it would be easier for intoxicated cells to evolve resistance to this type of activity. The first example of a T6SS effector to act in this way is the recently described ADP-ribosylating toxin Tre1 from *Serratia proteamaculans* (Ting et al., 2018). This unique effector causes growth cessation by inhibiting cell division, which it accomplishes via ADP-ribosylation of the prokaryotic tubulin homologue FtsZ. FtsZ protomers polymerize to form a contractile “Z-ring” that plays an essential role in cytokinesis during cell division (Bi and Lutkenhaus, 1991). ADP-ribosylation of FtsZ prevents Z-ring formation causing a defect in cell division, resulting in filamentation and the eventual death of targeted cells. Remarkably, the Tre1-specific immunity protein, Tri1, can protect cells from Tre1 intoxication via its ADP-ribosylhydrolase

activity. In contrast to all other characterized effector-immunity pairs, this mechanism of immunity is not reliant on a highly specific protein-protein interaction and can therefore confer broad immunity to diverse interbacterial ADP-ribosyltransferases (Ting et al., 2018).

Chaperones of T6SS effectors

The Hcp component of T6SSs possesses chaperone-like properties that promote the stability of small, single-domain effectors prior to their export (Silverman et al., 2013). By contrast, large multi-domain effectors, which interact with VgrG proteins to facilitate their export, often require an additional protein for secretion. Three Domain of Unknown Function (DUF) protein superfamilies, DUF4123, DUF1795 and DUF2169, have been demonstrated to interact with and promote the secretion of VgrG-interacting effectors (Alcoforado Diniz and Coulthurst, 2015; Bondage et al., 2016; Liang et al., 2015; Whitney et al., 2015). For example, the DUF1795 chaperone EagT6 from *P. aeruginosa* binds its cognate PAAR domain-containing effector Tse6 and is speculated to facilitate complex formation with its cognate VgrG (Whitney et al., 2015). The precise roles of DUF2169 and DUF4123 chaperones have not been established; however, a DUF4123 chaperone from *P. aeruginosa* was shown to require interaction with a co-chaperone to bind its cognate effector and together, these proteins facilitate loading of a PAAR domain-containing effector onto a VgrG spike (Burkinshaw et

al., 2018). Further structural and biochemical analyses of these chaperone families are needed to better understand their function in effector delivery.

Most of the aforementioned studies exploring T6SS-associated effectors have studied the distribution (Russell et al., 2012), mechanisms of recognition (Silverman et al., 2013) and toxic functions of Hcp-associated effectors (Russell et al., 2011). However, comparatively less is known about VgrG-associated effectors. Beyond the presence of a PAAR domain (Cianfanelli et al., 2016a) and the isolated characterization of several toxin domains within VgrG-associated effectors (Ting et al., 2018; Whitney et al., 2015), no strides have been made to characterize general attributes of this family of effectors and explore the diversity of toxins associated with it.

Objectives of this dissertation

The work described in this thesis aims to characterize novel components of a widely distributed family of PAAR effectors that possess N-terminal transmembrane domains (TMDs) and C-terminal toxin domains.

Chapter 2 describes my work on the characterization of the N-terminus of TMD-containing PAAR effectors and their protein partners. This work shows that DUF1795-containing proteins function as molecular chaperones that bind and stabilize the TMDs within these effectors prior to their secretion. In addition to chaperone interactions, the work in this chapter also characterizes an important N-

terminal motif within this family of effectors, called prePAAR, that is necessary for effector recruitment to the T6SS.

In chapter 3, I characterize a novel C-terminal toxin in the TMD-containing PAAR effector Tas1. I show that upon its delivery to target cells, Tas1 causes cell death by depleting ADP and ATP via the synthesis of adenosine penta- and tetraphosphate ((p)ppApp). This work reveals a novel mechanism of interbacterial antagonism and describes the first characterization of a (p)ppApp-synthesizing toxin. I extend this work by identifying hundreds of Tas1 homologs that are not associated with a T6SS but instead likely transit other protein secretion machineries. I show that a number of these Tas1 homologous proteins also synthesize (p)ppApp suggesting that (p)ppApp may play additional roles in bacterial physiology beyond T6SS-dependent interbacterial antagonism.

Chapter 4 provides a conclusion to the thesis by highlighting important insights this work provides for T6SS effector trafficking and toxicity and discusses the potential roles of (p)ppApp in diverse bacterial species.

CHAPTER II – Structural basis for transmembrane domain recognition by type VI secretion system chaperones

Preface

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

Quentin, D., **Ahmad, S.**, Shanthamoorthy, P., Mougous, J.D., Whitney, J.C., and Raunser, S. (2018). Mechanism of loading and translocation of type VI secretion system effector Tse6. Nat Microbiol *3*, 1142-1152.

Author contributions: S.R., J.C.W. and J.D.M. designed the project. J.C.W. and J.D.M. provided protein complexes. D.Q. prepared specimens, recorded, analysed and processed the EM data, performed the liposome-based in vitro assay and prepared figures. S.R. managed the project. S.A. and J.C.W. performed the biochemical and cellular in vivo experiments. P.S. introduced point mutations into EagT6. D.Q., J.C.W. and S.R. wrote the manuscript with input from all authors.

Ahmad, S., Tsang, K.K.*, Sachar, K.*, Quentin, D., Tashin, T.M., Bullen, N.P., Raunser, S., McArthur, A.G., Prehna, G., and Whitney, J.C. (2020). Structural basis for effector transmembrane domain recognition by type VI secretion system chaperones. Elife *9*.

*indicates equal contribution

Author contributions: S.A., G.P. and J.C.W. conceived experiments and designed the project. S.A. completed cloning, strain generation, bacterial competition assays and biochemical experiments. K.K.T. performed bioinformatics analyses. K.S. and G.P conducted protein crystallization, X-ray data collection and analysis. D.Q. performed negative-stain EM experiments. T.M.T. and N.P.B provided assistance with cloning and biochemical experiments. S.A., K.K.T., G.P., J.C.W. completed figure design, manuscript writing and editing. G.P. and J.C.W. supervised project. S.R., A.G.M., G.P. and J.C.W. provided project funding.

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

Abstract

Type VI secretion systems (T6SSs) deliver antibacterial effector proteins between neighbouring bacteria. Many effectors harbor N-terminal transmembrane domains (TMDs) implicated in effector translocation across target cell membranes. However, the distribution of these TMD-containing effectors remains unknown. Here we discover prePAAR, a conserved motif found in over 6,000 putative TMD-containing effectors encoded predominantly by 15 genera of Proteobacteria. Based on differing numbers of TMDs, effectors group into two distinct classes that both require a member of the Eag family of T6SS chaperones for export. Co-crystal structures of class I and class II effector TMD-chaperone complexes from *Salmonella* Typhimurium and *Pseudomonas aeruginosa*, respectively, reveal that Eag chaperones mimic transmembrane helical packing to stabilize effector TMDs. In addition to participating in the chaperone-TMD interface, we find that prePAAR residues mediate effector-VgrG spike interactions. Taken together, our findings reveal mechanisms of chaperone-mediated stabilization and secretion of two distinct families of T6SS membrane protein effectors.

Main

Bacteria secrete proteins to facilitate interactions with their surrounding environment. In Gram-negative bacteria, the transport of proteins across cellular membranes often requires the use of specialized secretion apparatuses found within the cell envelope. One such pathway is the type VI secretion system (T6SS),

which in many bacterial species functions to deliver antibacterial effector proteins from the cytoplasm directly into an adjacent bacterial cell via a one-step secretion event (Russell et al., 2011). A critical step that precedes type VI secretion is the selective recruitment of effectors to the T6SS apparatus. Recent work has shown that for many effectors this process requires chaperone proteins, which are thought to maintain effectors in a 'secretion-competent' state (Unterweger et al., 2017). However, to-date, no molecular-level evidence exists to support this idea.

The T6SS is comprised of two main components: a cell envelope-spanning membrane complex and a cytoplasmic bacteriophage tail-like complex. The latter contains a tube structure formed by many stacked copies of hexameric ring-shaped hemolysin co-regulated protein (Hcp) capped by a single homotrimer of valine-glycine repeat protein G (VgrG) (Mougous et al., 2006; Spínola-Amilibia et al., 2016). Together, these proteins form an assembly that resembles the tail-tube and spike components of contractile bacteriophage (Renault et al., 2018). Additionally, VgrG proteins interact with a single copy of a cone-shaped proline-alanine-alanine-arginine (PAAR) domain-containing protein that forms the tip of the VgrG spike (Shneider et al., 2013). Altogether, PAAR, Hcp and VgrG are necessary for T6SS function, and during a secretion event these components are themselves delivered into target cells (Cianfanelli et al., 2016a). Prior to its export from the cell, the bacteriophage tail-like complex is loaded with toxic effector proteins. In contrast to proteins that are exported by the general secretory pathway, T6SS effectors do not contain linear signal sequences that facilitate their recognition by the T6SS

apparatus. Instead, effectors transit the T6SS via physical association with Hcp, VgrG or PAAR proteins (Cianfanelli et al., 2016b).

In addition to its role in effector export, Hcp also possesses chaperone-like properties that facilitate cytoplasmic accumulation of Hcp-interacting effectors prior to their secretion (Silverman et al., 2013). This chaperone activity has been attributed to the interior of the ~4 nm pore formed by hexameric Hcp rings, which are wide enough to accommodate small, single-domain effectors. Individual Hcp rings appear to possess affinity towards multiple unrelated effectors. However, the molecular basis for this promiscuous substrate recognition is unknown.

In contrast to their Hcp-associated counterparts, VgrG-linked effectors are typically comprised of multiple domains and often require effector-specific chaperones for stability and/or to facilitate their interaction with the VgrG spike. Thus far, three effector-specific chaperone families belonging to the DUF1795, DUF2169 and DUF4123 protein families have been described. Studies on representative DUF2169 and DUF4123 proteins indicate that these chaperones minimally form ternary complexes with their cognate effector and a PAAR protein to facilitate the 'loading' of the PAAR domain and effector onto their cognate VgrG (Bondage et al., 2016; Burkinshaw et al., 2018). In contrast, DUF1795 proteins, also known as effector associated gene (Eag) chaperones, interact with so-called 'evolved' PAAR proteins in which the PAAR and toxin domains are found as a single polypeptide chain (Alcoforado Diniz and Coulthurst, 2015; Whitney et al., 2015).

In this work, we report the discovery of prePAAR, a highly conserved motif that enabled the identification of over 6,000 putative T6SS effectors, all of which possess N-terminal TMDs and co-occur in genomes with Eag chaperones. Further informatics analyses found that these candidate effectors can be categorized into one of two broadly defined classes. Class I effectors belong to the Rhs family of proteins, are comprised of ~1200 amino acids and possess a single region of N-terminal TMDs. Class II effectors are ~450 amino acids in length and possess two regions of N-terminal TMDs. We validate our informatics approach by showing that a representative member of each effector class requires a cognate Eag chaperone for T6SS-dependent delivery into susceptible bacteria. Crystal structures of Eag chaperones in complex with the TMDs of cognate class I and class II effectors reveal the conformation of effector TMDs prior to their secretion and insertion into target cell membranes. In addition to participating in chaperone-effector interactions, structure-guided mutagenesis of hydrophilic residues within prePAAR show that this motif also enables effector interaction with its cognate VgrG. We also find that in the presence of liposomes a representative TMD-containing effector spontaneously inserts into membranes in vitro, causing release from its cognate chaperone and self-translocation of the effector toxin domain across the membrane. Collectively, our data provide the first high-resolution structural snapshots of T6SS effector-chaperone interactions, define the molecular determinants for effector TMD stabilization and recruitment to the T6SS apparatus and demonstrate a role for effector TMDs in type VI secretion.

prePAAR is a motif found in TMD-containing effectors that interact with Eag chaperones

Characterization of Eag chaperones and their associated PAAR effectors has thus far been limited to the EagT6-Tse6 and EagR1-RhsA chaperone-effector pairs from *P. aeruginosa* and *Serratia marcescens*, respectively (Cianfanelli et al., 2016a; Whitney et al., 2015). In both cases, the chaperone gene is found upstream of genes encoding its cognate PAAR effector and an immunity protein that protects the toxin-producing bacterium from self-intoxication (Figure 2.1A). Both EagT6 and EagR1 directly interact with and are necessary function of their cognate effectors, however, the molecular basis for this interaction remains unknown. To address this, we used an informatics approach to identify other Eag-associated effectors and search for potential conserved regions within these proteins that permit Eag binding. Homology-based searches for additional Eag chaperones can yield difficult to interpret results due to a scarcity of conserved residues and homology of this protein family to the phage protein DcrB (Samsonov et al., 2002), which is widely distributed in both T6SS-positive and T6SS-negative organisms. We instead screened different regions of the known Eag-interacting effector Tse6 using *jackhmmer* to generate sequence alignment hidden Markov models (HMM) using an iterative search procedure that queried the UniProtKB database (Johnson et al., 2010). Multiple regions in Tse6 are not well conserved. For example, Tse6 and RhsA are PAAR effectors that contain N-terminal TMDs, however, using the Tse6 PAAR domain as a seed sequence resulted in the identification of several PAAR proteins lacking N- and C-terminal extensions (Shneider et al., 2013) and aside

from being comprised of hydrophobic residues, the TMDs themselves are poorly conserved. Remarkably, however, the HMM obtained from the N-terminal 60 residues of Tse6 revealed a nearly invariant AARxxDxxxH motif, which in Tse6 is found in the first 15 residues of the protein and is immediately N-terminal to its first TMD (Figure 2.1B). In total, our query identified over 2,054 proteins containing this motif (Table 2.1). Among these candidate effectors, our search identified the recently characterized toxins Tre1 and DddA as well as many toxins of unknown function indicating that our approach may have identified T6SS effectors with novel biochemical activities (Mok et al., 2020; Ting et al., 2018). Interestingly, prior to any knowledge of PAAR domains or Eag chaperones being involved in T6SS function, Zhang and colleagues noted the existence of this N-terminal motif in PAAR-containing proteins through an informatics analysis of bacterial nucleic acid degrading toxins (Zhang et al., 2011). Here, they refer to the motif and its adjacent TMD region as “prePAARTM” because these sequence elements co-occur with one another and because they are both found N-terminal to PAAR domains. We have chosen to refer to the motif as “prePAAR” because, as described below, our data indicate it has a function that is distinct from the TMD regions.

Examination of our putative effector sequences revealed that prePAAR is substantially enriched in bacterial genera with characterized T6SSs including *Pseudomonas*, *Burkholderia*, *Salmonella*, *Shigella*, *Escherichia*, *Enterobacter*, *Yersinia*, and *Serratia*. Interestingly, no prePAAR motifs were identified in *Vibrio* despite an abundance of species within this genus possessing highly active

bacteria-targeting T6SSs. We next obtained all 56,324 available genomes from NCBI for the abovementioned genera and found that 26,327 genomes encode at least one prePAAR motif. After removing all redundant sequences, 6,101 unique prePAAR-containing proteins present across 5,584 genomes were used for further analyses (Table 2.2, List C). In these genomes, we determined that approximately 90% encode a single prePAAR motif, although instances where prePAAR is present up to six times within a single genome were also identified (Figure 2.1C). To determine if these unique proteins are probable TMD-containing T6SS effectors that require Eag chaperones for secretion, we next examined each prePAAR-containing protein and its associated genome for the following three criteria: 1) the existence of an Eag chaperone encoded in the same genome, 2) the presence of a downstream PAAR domain and 3) predicted TMDs in the first 300 amino acids of the protein (Käll et al., 2007; Krogh et al., 2001). The location restriction in our TMD search was used in order to exclude C-terminal toxin domains that possess TMDs, which differ from N-terminal translocation TMDs in that they may not require chaperones for secretion (Mariano et al., 2019). We searched each genome for Eag proteins using an HMM for DUF1795 and found that 99.5% (5,554/5,584) of prePAAR-containing genomes also possessed at least one *eag* gene (Jones et al., 2014). In approximately 14% of the 5,554 genomes analyzed, the number of prePAAR motifs matched the number of Eag homologues. In the remainder of cases, the number of Eag homologous proteins exceeded the number of prePAAR motifs, with a weighted average of 2.5 paralogues per genome. As is the case with

eagT6-tse6 and *eagR1-rhsA*, ~90% of the identified prePAAR-containing effector genes appear directly beside an *eag* gene whereas the remaining ~10% are found in isolation suggesting that their putative chaperone is encoded elsewhere in the genome. We removed pre-PAAR-containing protein fragments (proteins less than 100 amino acids in length) and further reduced redundancy by clustering sequences with 95% identity. Remarkably, in all but two of the remaining 1,166 prePAAR-containing proteins, we identified a PAAR domain, indicating a probable functional relationship between prePAAR and PAAR. The two prePAAR-containing proteins lacking a PAAR domain were either adjacent to a gene encoding a PAAR domain-containing protein or directly beside T6SS structural genes. Finally, we searched 1,166 prePAAR-containing proteins for TMDs and found that all protein sequences contained predicted TMDs with 86% having one region of TMDs and 14% having two regions of TMDs. In sum, our prePAAR-based search procedure identified thousands of candidate effector proteins that possess a striking genetic association with Eag proteins.

To further analyze our collection of prePAAR-containing effectors, we built a phylogenetic tree from 1,166 non-redundant effector sequences that represent the diversity present in our collection of sequences (Figure 2.2A). Interestingly, two distinct sizes of proteins emerged from this analysis: large prePAAR effectors that are on average 1,196 amino acids in length and small prePAAR effectors comprised of an average of 443 amino acids (Figure 2.2B). As noted previously, all effectors contained predicted TMDs; however, large effectors almost exclusively

contained a single region of TMDs N-terminal to their PAAR domain whereas most small effectors contained TMD regions N- and C-terminal to their PAAR domain. To distinguish between these two domain architectures, we hereafter refer to large, single TMD region-containing prePAAR effectors as class I and small, two TMD region-containing prePAAR effectors as class II. Notably, class I effectors also contain numerous YD repeat sequences, which are a hallmark of rearrangement hotspot (Rhs) proteins that function to encapsulate secreted toxins (Busby et al., 2013). We also found a small subset of these effectors are encoded by two separate ORFs, the first encoding prePAAR-TMD-PAAR and the second encoding a protein containing Rhs repeats and a C-terminal toxin domain (Figure 2.2C). Conversely, class II effectors are distinguished by a GxxxxGxxLxGxxxD motif in addition to their second TMD region.

As a first step towards validating our informatics approach for identifying Eag chaperone-effector pairs, we assessed the ability of several newly identified Eag chaperones to interact with the prePAAR-containing effector encoded in the same genome (Figure 2.2C). We included the previously characterized class II effector Tse6 known to interact with EagT6 and we similarly found that when expressed in *E. coli*, Eag chaperones from *Enterobacter cloacae*, *Salmonella* Typhimurium, *Shigella flexneri* and *Serratia proteamaculans* co-purified with their predicted cognate effector (Figure 2.2D). Collectively, these findings indicate that prePAAR proteins constitute two classes of TMD-containing T6SS effectors and that representative members from both classes interact with Eag chaperones.

Eag chaperones are specific for cognate prePAAR effectors

We next sought to examine the specificity of Eag chaperones towards prePAAR effectors in a biologically relevant context. To accomplish this, we inspected our list of prePAAR effectors and found that the soil bacterium *Pseudomonas protegens* Pf-5 possesses both a class I and class II effector, encoded by the previously described effector genes *rhsA* and *tne2*, respectively (Tang et al., 2018). Furthermore, the genome of this bacterium encodes two putative Eag chaperones, PFL_6095 and PFL_6099, which have 25% sequence identity between them (Figure 2.3A). PFL_6095 is found upstream of *rhsA* and is likely co-transcribed with this effector whereas PFL_6099 is not found next to either effector gene. To examine the relationship between these genes, we generated strains bearing single deletions in each effector and chaperone gene and conducted intraspecific growth competition assays against *P. protegens* recipient strains lacking the *rhsA-rhsI* or *tne2-tni2* effector-immunity pairs. We noted that protein secretion by the T6SS of *P. protegens* is substantially inhibited by the threonine phosphorylation pathway, so we additionally inactivated the threonine phosphatase encoding gene *pppA* in recipients to induce a 'tit-for-tat' counterattack by wild-type donor cells (Figure 2.3B-C) (Basler et al., 2013; Mougous et al., 2007). Consistent with the effector-immunity paradigm for bacteria-targeting T6SSs, wild-type *P. protegens* readily outcompeted Δ*rhsA* Δ*rhsI* Δ*pppA* and Δ*tne2* Δ*tni2* Δ*pppA* strains in a *rhsA*- and *tne2*-dependent manner, respectively (Figure 2.3D). Additionally, we found that a strain lacking PFL_6095 no longer exhibited a co-

culture fitness advantage versus a Δ*rhsA* Δ*rhsI* Δ*pppA* recipient but could still outcompete *tne2* sensitive recipients to the same extent as the wild-type strain. Conversely, a ΔPFL_6099 strain outcompeted Δ*rhsA* Δ*rhsI* Δ*pppA* but not Δ*tne2* Δ*tni2* Δ*pppA* recipients. Together, these data indicate that the delivery of RhsA and Tne2 into susceptible target cells requires effector-specific *eag* genes.

To test the ability of PFL_6095 and PFL_6099 to act as RhsA- and Tne2-specific chaperones, respectively, we co-expressed each chaperone-effector pair in *E. coli* and examined intracellular effector levels by western blot. Consistent with functioning to promote cognate effector stability, accumulation of RhsA only occurred in the presence of PFL_6095 whereas Tne2 accumulated in cells containing PFL_6099 (Figure 2.3E). We next examined the stability-enhancing properties of PFL_6095 and PFL_6099 when expressed at native levels in *P. protegens*. Due to challenges associated with detecting RhsA and Tne2 in unconcentrated cell lysates, we constructed chromosomally encoded N-terminal decahistidine-tagged (his_{10}) fusions of RhsA and Tne2 to facilitate the enrichment of these proteins from *P. protegens* and confirmed that these fusions did not compromise the ability of these effectors to intoxicate recipients (Figure 2.3F). Following affinity purification, RhsA and Tne2 levels were assessed using RhsA and vesicular stomatitis virus glycoprotein epitope (VSV-G) antibodies, respectively. In line with our data in *E. coli*, we were unable to detect RhsA in the absence of PFL_6095 whereas Tne2 was absent in a strain lacking PFL_6099 (Figure 2.3G). Collectively, these data suggest that Eag chaperones exhibit a high

degree of specificity for their cognate effectors. Based on our characterization of these genes, we propose to rename PFL_6095 and PFL_6099 to *eagR1* and *eagT2*, respectively, to reflect their newfound role as chaperones for the prePAAR-containing effectors RhsA and Tne2.

Previous biochemical studies on the class II prePAAR effector Tse6 suggested that its cognate chaperone, EagT6, interacts with the N-terminus of this effector (Whitney et al., 2015). Based on our informatics work, the N-terminal region of Eag-associated effectors always contains both a prePAAR motif and TMD(s), which could constitute the binding site for Eag chaperones. To test this idea, we constructed N-terminal truncation mutants of the representative class I and class II effector proteins, RhsA and Tse6, respectively (Figure 2.4A, E). We first co-expressed a variant of RhsA lacking its N-terminal region ($RhsA_{\Delta NT}$) with EagR1 in *E. coli*. Consistent with our hypothesis, affinity purification of $RhsA_{\Delta NT}$ showed that this truncated variant does not co-purify with EagR1 (Figure 2.4B). Additionally, expression of the deleted 74 residue N-terminal fragment of RhsA in isolation was sufficient for EagR1 binding (Figure 2.4C). Our data also demonstrated that in contrast to wild-type RhsA, $RhsA_{\Delta NT}$ is stable in the absence of EagR1 when expressed in *E. coli* indicating that RhsA's N-terminus imparts instability on the protein in the absence of its cognate chaperone. In *P. protegens*, we could readily detect $rhsA_{\Delta NT}$ in a strain lacking *eagR1*, corroborating our findings in *E. coli* (Figure 2.4D). We conducted a similar analysis for Tse6, which in contrast to RhsA has two TMDs within its N-terminus. Thus, we generated three mutants of

Tse6, the first lacking its N-terminal prePAAR and TMD1 ($Tse6_{\Delta NT}$), the second lacking only TMD2 ($Tse6_{\Delta TMD2}$) and lastly, a mutant lacking prePAAR, TMD1 and TMD2 ($Tse6_{\Delta NT/TMD2}$) (Figure 2.4E). To assess binding to EagT6, we co-expressed wild-type Tse6 and each of the abovementioned variants with EagT6 in *E. coli* and conducted pulldown assays. Individual deletions of prePAAR+TMD1 or TMD2 reduced but did not abolish EagT6 binding, however, deletion of both regions completely abrogated effector binding to EagT6 (Figure 2.4F). Importantly, none of the truncation mutants were affected in their ability to bind EF-Tu, which is necessary for Tse6-dependent growth inhibition of target cells (Whitney et al., 2015). Expression of either the N-terminal region containing prePAAR+TMD1 or TMD2 in isolation was sufficient for EagT6 binding (Figure 2.4G). Similar to our findings for RhsA, the $Tse6_{\Delta NT/TMD2}$ variant was stable in the absence of EagT6 and accumulated to wild-type levels in *P. aeruginosa* lacking *eagT6* (Figure 2.4H). The previously solved apo structure of the EagT6 homodimer reveals that it has a concave cavity enriched with hydrophobic residues (Figure 2.4I, PDB: 1TU1), which likely facilitates binding to the TMDs. To test this, we introduced mutations at the *eagT6* locus of *P. aeruginosa* that encode site-specific variants of the EagT6 protein and examined whether mutated EagT6 is still able to stabilize Tse6 in vivo (Figure 2.4J). Indeed, substitutions of hydrophobic residues in the concave cavity reduced Tse6 stability, whereas mutations in other regions of the EagT6 dimer had no effect (Figure 2.4J). Collectively, our data suggest that Eag chaperones binding

to their cognate class I and class II effectors is mediated by regions harbouring TMDs.

Our data also suggest that in contrast to the wild-type proteins, chaperone 'blind' $RhsA_{\Delta NT}$ and $Tse6_{\Delta NT/TMD2}$ are stable in the absence of their cognate Eag chaperones. To confirm that these truncation mutants are not misfolded, we expressed them in *E. coli* and found that both $RhsA_{\Delta NT}$ and $Tse6_{\Delta NT/TMD2}$ are equally toxic to *E. coli* as their wild-type counterparts (Figure 2.5A, C). We also confirmed through depletion of the Tse6-specific immunity protein, Tsi6, that a $Tse6_{\Delta NT/TMD2}$ is toxic to *P. aeruginosa* (Figure 2.5D). Despite the stability and toxicity of these variants, neither a *P. protegens* strain expressing $RhsA_{\Delta NT}$ or a *P. aeruginosa* strain expressing $Tse6_{\Delta NT/TMD2}$ was able to outcompete RhsA-sensitive or Tse6-sensitive recipient cells, respectively, demonstrating an essential role for the chaperone-bound N-terminus during interbacterial competition (Figure 2.5B, E).

PAAR effectors are necessary for the function of their associated T6SS (Cianfanelli et al., 2016a). We hypothesized that perturbation of N-terminal domains in RhsA or Tse6 may affect the folding of their PAAR domain, which would inhibit secretion of other substrates in a strain lacking other essential PAAR effectors. To test this, we generated a strain of *P. aeruginosa* expressing $Tse6_{\Delta NT/TMD2}$, lacking the only other known PAAR effector secreted by this pathway, Tse5, and subsequently tracked the secretion of the known well characterized effector Tse1. A strain lacking Tse5 and Tse6 was unable to secrete Tse1, suggesting that the T6SS of *P. aeruginosa* also requires at least one PAAR

domain-containing effector for function. Interestingly, we achieve the same result in Δ*tse5* strains expressing $Tse6_{\Delta NT/TMD2}$ (Figure 2.5F). This result, in line with our growth competition assay, indirectly suggests that the interaction of the TMDs of Tse6 with EagT6 somehow facilitates the formation of the phage tail spike-like capping structure that is required from T6SS function.

To further understand the mechanism preventing secretion of these mutants, we next examined the ability of $RhsA_{\Delta NT}$ or $Tse6_{\Delta NT/TMD2}$ to interact with a cognate secreted structural component of the T6SS apparatus. T6SS effectors encoded downstream of *vgrG* genes typically rely on the encoded VgrG protein for delivery into target cells (Whitney et al., 2014). Consistent with this pattern, PFL_6094 encodes a predicted VgrG protein, herein named VgrG1, which we confirmed is required for RhsA-mediated growth inhibition of susceptible target cells (Figure 2.6A), whereas VgrG1a from *P. aeruginosa* has previously been shown to be necessary for Tse6 secretion (Whitney et al., 2015). Using a *P. protegens* strain expressing His_{10}-tagged RhsA and FLAG-tagged VgrG1 from their native loci, we found that these proteins physically interact to form a complex (Figure 2.6B). To test if the absence of the chaperone-bound N-terminus affects the formation of this complex, we used our *E. coli* co-expression system to purify RhsA-EagR1-VgrG1 complexes. These experiments show that $RhsA_{\Delta NT}$ is not able to interact with VgrG1, even though this truncated protein possesses its PAAR domain, which in T6SS effectors lacking prePAAR and TMDs in their N-terminus, is sufficient for VgrG interaction (Figure 2.6C) (Bondage et al., 2016). To gain

insight into how EagR1 binding facilitates RhsA interaction with VgrG1, we next performed negative-stain electron microscopy (EM) to examine the configuration of each subunit within this complex. To facilitate the accurate identification of each component, we obtained class averages of purified VgrG1, $RhsA_{\Delta NT}$, RhsA-EagR1 complex and RhsA-EagR1-VgrG1 complex (Appendix Figure A2.1). As expected, isolated VgrG1 and $RhsA_{\Delta NT}$ proteins appeared as characteristic spike- and barrel-shaped proteins, respectively (Busby et al., 2013; Spínola-Amilibia et al., 2016); Figure 2.6F-G). Intriguingly, images of RhsA-EagR1 complexes contained a sphere-shaped object that likely represents a subcomplex between EagR1 and the N-terminus of RhsA (Figure 2.6H). Lastly, the class-averages of RhsA-EagR1-VgrG1 complexes revealed a close association of EagR1 and RhsA with the tip of the VgrG spike, which is likely mediated by the PAAR domain of RhsA (Figure 2.6I). Interestingly, though both complexes exhibit significant rotational flexibility, the average distance between the subcomplex formed by EagR1 and the N-terminus of RhsA is substantially greater in the absence of VgrG1 (average distance: 2.68 nm, *n* = 27 classes versus 1.20 nm, *n* = 26 classes) (Appendix Figure A2.1). When taken together with our biochemical experiments, these structural data indicate that EagR1 stabilizes the N-terminus of RhsA, which may also orient the effector such that it can interact with its cognate VgrG.

We next turned to validating VgrG1a interaction with the Tse6 truncation mutants described previously. We previously showed that two forms of Tse6 can be detected following electrophoresis: 1) As part of a high-molecular weight Tse6-

VgrG1a complex that sits above the resolving gel or 2) as 'free' Tse6 that migrates at its expected molecular weight of 45 kDa. Deletion of one or both TMD regions in Tse6 prevents the formation of a Tse6-VgrG1a complex, which is readily detectable in *P. aeruginosa* strains expressing Tse6 and does not affect the monomeric form of the protein (Figure 2.6D). Expression of these variants in *E. coli* revealed that only wild-type Tse6 is able to bind and form SDS-resistant complexes with VgrG1a (Figure 2.6E). These findings were corroborated by a structure of the 'pre-firing' VgrG1a-Tse6-EagT6$_2$-EF-Tu-Tsi6 complex, which was determined by cryo-EM and single-particle analysis by our collaborators (Moriya et al., 2017; Quentin et al., 2018). The overall resolution of the reconstruction was 4.2 Å and comprises VgrG1a, the N-terminal PAAR and TMDs of Tse6 and two homodimers of EagT6 (Figure 2.6J). Model refinement for the VgrG1a was completed using the previously solved unpublished VgrG1a crystal structure (PDB: 4MTK). The density directly adjacent to the VgrG was assigned to the PAAR domain of Tse6, which is known to bind the flat C-terminal β-helical surface of VgrG (Shneider et al., 2013). In close proximity to the PAAR domain are two horse-shoe shaped structures that correspond to EagT6 dimers. This density was readily fit with two copies of the crystal structure of this protein (PDB: 1TU1) and supports our biochemical data implicating the concave surfaces of EagT6 proteins in binding the TMDs of Tse6. The C-terminal toxin domain of Tse6 (Tse6$_{tox}$), EF-Tu and Tsi6 could not be resolved, indicating that this subcomplex is highly flexible. In line with our

biochemical analysis, both EagT6 chaperones were observed to contain additional density in their concave cavity that likely corresponds to the TMDs (Figure 2.6K).

Eag chaperones bind effector TMDs by mimicking transmembrane helical packing

In addition to a TMD-containing region, the N-terminus of prePAAR effectors also harbours the prePAAR motif itself. However, the negative stain EM images of RhsA-EagR1-VgrG1 particles presented herein and our single-particle cryo-EM structure of a complex containing Tse6-EagT6-VgrG1 are of insufficient resolution to resolve the structures of chaperone-bound effector TMDs or the prePAAR motif. Therefore, to better understand the molecular basis for chaperone-TMD interactions and to gain insight into prePAAR function we initiated X-ray crystallographic studies on both class I and class II effector-chaperone complexes. Efforts to co-crystallize *P. protegens* EagR1 with the prePAAR and TMD-containing N-terminus of RhsA were unsuccessful. However, the EagR1 homologue SciW from *Salmonella* Typhimurium crystallized in isolation and in the presence of the N-terminus of the class I prePAAR effector Rhs1 ($Rhs1_{NT}$), allowing us to determine apo and effector bound structures to resolutions of 1.7Å and 1.9Å, respectively (Figure 2.7A, C and Table 2.4). Similar to RhsA, we confirmed that an $Rhs1_{\Delta NT}$ variant was unable to bind its cognate chaperone, SciW (Figure 2.7B). The structure of the EagT6 chaperone was previously solved as part of a structural genomics effort and we were additionally able to obtain a 2.6Å co-crystal structure

of this chaperone in complex with the N-terminal prePAAR and first TMD region of the class II effector Tse6 ($Tse6_{NT}$) (Figure 2.7E, F and Table 2.4).

The overall structure of SciW reveals a domain-swapped dimeric architecture that is similar to the previously described apo structure of *P. aeruginosa* EagT6 though each chaperone differs in its electrostatic surface properties (Figure 2.8A-D) (Whitney et al., 2015). A comparison of the chaperone structures in their apo and effector bound states shows that upon effector binding, both chaperones 'grip' the prePAAR-TMD region of their cognate effector in a claw-like manner (Figure 2.7D, G). Although our biochemical data indicate that Eag chaperones exhibit a high degree of specificity for their associated effector, the internal surface of the claw-shaped dimer contains a number of conserved residues that make critical interactions with the TM helices in both complexes (Figure 2.9A-F). For example, I22 and I24 of EagT6 create a hydrophobic surface in the 'palm' of the claw, which is flanked on either side by symmetrical hydrophobic surfaces comprised of A62, L66, L98, F104 and I113 (Figure 2.9B-D). Furthermore, the conserved hydrophilic residues S37, S41, Q58, and Q102 also interact with the bound effectors by making bifurcated hydrogen bonds to amide or carbonyl groups in the peptide backbone of the TM helices (Figure 2.9E-F). These polar interactions between chaperone and effector TM helices are striking because they are reminiscent of polar interactions seen within the helical packing of alpha helical transmembrane proteins, which often use serine and glutamine residues to mediate inter-helical interactions via bifurcated hydrogen bonds between side

chain and main chain atoms (Adamian and Liang, 2002; Dawson et al., 2003; Dawson et al., 2002). Additionally, EagT6 and SciW provide 'knob-hole-like' interactions, which also feature prominently in membrane protein packing (Curran and Engelman, 2003). Knob-hole interactions involve a large hydrophobic residue on one TM helix acting as a 'knob' to fill the hole provide by a small residue such as glycine or alanine on another TM-helix. TM holes are typically created by GxxxG/A motifs such as those found in G19-A24 (Rhs1) and G25-A30 (Tse6). In this case, the conserved Eag chaperone residue L66 provides a knob for the A24/30 hole (Figure 2.9E-F). Given that the Eag chaperone dimer creates a hydrophobic environment with complementary knob-hole interactions for its cognate effector TM helices and interacts with TM helices via side chain to main chain hydrogen bonds, we conclude that Eag chaperones provide an environment that mimics transmembrane helical packing to stabilize prePAAR effector TMDs in the cytoplasm prior to effector export from the cell.

prePAAR facilitates PAAR domain folding and interaction with the VgrG spike

We next compared the conformation of the bound prePAAR-TMD fragments between our effector-chaperone co-crystal structures. Interestingly, despite the abovementioned similarities between the SciW and EagT6 structures, the conformation of the N-terminal fragment of their bound prePAAR effector differs significantly. In the SciW complex, $Rhs1_{NT}$ adopts an asymmetric binding mode whereby the effector fragment does not make equivalent molecular contacts with

both chains of the two-fold symmetrical chaperone dimer (Figures 2.7C and 2.9E). The first TM helix (residues 19-33) binds to the hydrophobic cavity of one SciW protomer whereas the remaining hydrophobic region of Rhs1, which consists of two anti-parallel alpha-helices connected by a short 3_{10} helix, occupies the remainder of the binding surface. Phenylalanine residues F20 and F43 likely play an important role in the asymmetric binding of Rhs1 to SciW because their hydrophobic side chains insert into equivalent hydrophobic pockets found in each SciW protomer (Figure 2.9E). By contrast, $Tse6_{NT}$ exhibits a pseudosymmetric binding mode with EagT6 (Figures 2.7F and 2.9F). In this structure, two alpha-helices of Tse6 each occupy equivalent Eag binding pockets and run in the opposite direction to match the antiparallel arrangement of the EagT6 dimer. For example, A7 and A30 of Tse6 interact with equivalent sites in their respective chaperone protomers (Figures 2.7F and 2.9F). These two helices, which consist of prePAAR and a TM helix, flank a central TM helix whose C-terminus extends into the solvent, likely indicating the location of the downstream PAAR domain in the full-length effector.

A lack of interpretable electron density prevented modelling of Rhs1's entire AARxxDxxxH prePAAR motif in our $Rhs1_{NT}$-SciW co-crystal structure. However, the DxxxH portion of this motif is part of a short 3_{10} helix that orients the aspartate and histidine side chains such that they face outward into solvent (Figure 2.8E-G). By contrast, we were able to model the entire prePAAR motif of $Tse6_{NT}$ and in this case, the motif forms an alpha helix that binds the hydrophobic pocket of an EagT6

protomer. In this structure, the two conserved alanine residues of prePAAR make contact with the EagT6 chaperone whereas the arginine, aspartate and histidine residues are solvent exposed (Figures 2.7F and 2.9F). Remarkably, despite existing in different secondary structure elements, the D11 and H15 prePAAR residues of Tse6 are located in a similar 3D location as their D9 and H13 counterparts in Rhs1 (Figure 2.9G). It should be noted that the modelled conformation of $Tse6_{NT}$ appears to be locked into place by crystal packing suggesting that in solution, Tse6's prePAAR motif may exhibit significant conformational flexibility and can dissociate from EagT6 as is observed for the prePAAR motif of Rhs1 (Figure 2.8H-I). In support of this, we found that the detergent β-D-decylmaltopyranoside disrupts the interaction between EagT6 and Tse6 suggesting that Eag chaperone-effector interactions are labile, likely because chaperone dissociation is required prior to effector delivery into target cells (Figure 2.13D) (Whitney et al., 2015). Intriguingly, docking our high resolution EagT6-$Tse6_{NT}$ crystal structure into our previously determined lower resolution Tse6-EagT6-VgrG1 cryo-EM map orients the D11 and H15 prePAAR residues of Tse6 in a position that suggests they interact with its PAAR domain (Figure 2.9H). In sum, our structural analyses of prePAAR shows that this region is likely dynamic, and its mode of interaction varies for class I and class II prePAAR effectors. However, both Eag chaperones bind the N-terminus of their cognate effector such that the conserved aspartate and histidine residues of prePAAR are positioned to

potentially be involved in interactions with the downstream PAAR domain, and thus may play a role in effector-VgrG interactions.

To test if prePAAR influences PAAR function, we next conducted mutagenesis analysis on Tse6 because its PAAR-dependent interaction with its cognate VgrG protein, VgrG1a, can be monitored *in vivo* by western blot. During denaturing electrophoresis, Tse6 appears in two forms: 1) a high-molecular weight species corresponding to Tse6-VgrG1a complex and 2) a low-molecular weight species indicative of free Tse6 (Whitney et al., 2015). Deletion of *vgrG1a* only affects complex formation whereas deletion of the *eagT6* gene results in a substantial reduction in the levels of both species providing a means to differentiate residues involved in effector-chaperone versus effector-VgrG interactions. Using this readout, we engineered *P. aeruginosa* strains expressing Tse6 D11A and H15A single amino acid substitutions and a D11A/H15A double substitution and examined the consequences of these prePAAR mutations on Tse6 interactions. In support of a role in promoting proper folding of PAAR, Tse6-VgrG1a complex formation was substantially reduced in a strain expressing the Tse6^{D11A} variant and abolished in a strain expressing Tse6$^{D11A, H15A}$ (Figure 2.10A-B). We next examined the effect of these mutations on T6SS-dependent delivery of Tse6 into target cells by subjecting these *P. aeruginosa* strains to growth competition assays against Tse6-sensitive recipients. In agreement with our biochemical data, strains expressing Tse6 harboring a D11A mutation exhibited a substantial reduction in

co-culture fitness consistent with an inability of these mutant proteins to form a complex with VgrG1a (Figure 2.10C).

To better understand why Tse6's PAAR domain requires prePAAR for function, we compared its sequence and predicted structure to the X-ray crystal structure of the 'orphan' PAAR domain c1882 from *E. coli*, which does not contain additional components such as TMDs or a toxin domain (Shneider et al., 2013). Interestingly, this analysis suggested that the PAAR domain of Tse6 lacks an N-terminal segment, which, based on the structure of c1882, is potentially important for the proper folding of this domain (Figure 2.11A). We next extended this structural analysis to include all PAAR domains of the prePAAR effectors that we experimentally confirmed bind Eag chaperones. In all cases, the N-terminal segment of each PAAR domain was missing (Figure 2.11B). We also noted that the prePAAR motif possesses significant sequence homology to the N-terminal segment of c1882, suggesting that even though this stretch of amino acids exists on the opposite side of the first TMD region of Tse6, it may comprise the missing segment of Tse6's PAAR domain (Figure 2.11C). Lending further support to this hypothesis, when we artificially fused Tse6's prePAAR motif (residues 1-16) with its PAAR domain (residues 77-163) and generated a structural model, we found that the first 16 residues of Tse6 fill the missing structural elements of Tse6's PAAR domain (Figure 2.11D). Based on this information, we hypothesized that prePAAR is necessary for PAAR domain function and thus, facilitates VgrG binding. To explore this experimentally, we assessed the effect of prePAAR on VgrG binding

in vivo and *in vitro*. We started by generating two Tse6 mutants expressed from their native locus in *P. aeruginosa*. The first mutant lacks prePAAR, TMD1 and TMD2 (Tse6$^{\Delta prePAAR,\ \Delta TMDs}$) while the second contains prePAAR fused to PAAR and thus lacks only its TMDs (Tse6$^{\Delta TMDs}$). We next assessed the ability of these truncated forms of Tse6 to form a complex with VgrG1a *in vivo*, as described above. In these experiments, we found that prePAAR and PAAR are together both necessary and sufficient for the formation of Tse6-VgrG1a complexes *in vivo* (Figure 2.11E). Of note, the amount of the complex formed by the Tse6$^{\Delta TMDs}$ mutant is less than the parent strain, which may be due to unstable structural elements that arose from suboptimal boundaries selected for truncating the effector. We next assessed the formation of this complex *in vitro* and found that co-incubation of Tse6, EagT6 and VgrG1a after overexpression in *E. coli* leads to the formation of SDS-resistant Tse6-VgrG1a complexes whereas doing so with a strain expressing the Tse6$^{D11A,\ H15A}$ prePAAR double mutant does not (Figures 2.11F-G). Importantly, these mutations do not affect overall levels of Tse6 in cells or affect its ability to bind to EagT6, indicating that these mutations do not have a global destabilizing effect on Tse6 (Figure 2.11F). Together, these data suggest that the PAAR domains of prePAAR effectors exist as 'split PAAR' due to the presence of N-terminal TMDs.

In orphan PAAR proteins, such as c1882, DxxxH motifs are necessary for Zn^{2+}-coordination and are therefore necessary for proper folding of this domain (Shneider et al., 2013). In agreement with this precedent, the conserved histidine

residue in the DxxxH portion of Tse6's prePAAR motif is predicted to be in the same 3D position as the first zinc-coordinating histidine residue of c1882 (Figure 2.12A). To extend this comparison further, we conducted in silico analyses to examine the evolutionary relationship and potential Zn^{2+}-binding residues in 564 orphan PAARs and 1,765 prePAAR effectors. We found that orphan PAAR sequences are ancestral to split PAAR domains and that while orphan PAAR proteins typically contain a total of four histidine and/or cysteine Zn^{2+}-coordinating residues, prePAAR effectors only contain three in their PAAR domain with the fourth likely being provided by the prePAAR motif (Figures 2.12B-C). In support of this prediction, we found that Tse6-VgrG1a complexes formed by the D11A or H15A variants were susceptible to heat treatment under denaturing conditions whereas the wild-type complex remained intact (Figures 2.10B and 2.11G).

The toxin domain of Tse6 is translocated across membranes

Our biochemical and structural data informed us on the recruitment of TMD-containing effectors to the T6SS apparatus and provide important insights into the state of these effectors in a donor cell prior to a T6SS firing event. Despite the requirement of Eag proteins for effector stability, we found that the EagT6 chaperone is not secreted by *P. aeruginosa* and is likely retained within the cell following T6SS firing, corroborating mass spectrometric analysis conducted on the secretome of *Serratia marcescens* showing that the Eag chaperones from this organisms are likely not secreted (Cianfanelli et al., 2016a) (Figure 2.13A). Current

evidence suggests that T6SS effectors are delivered to the periplasm of target cells and do not require a cell surface receptor for entry (Russell et al., 2011). We hypothesized that upon injection and release from their cognate Eag chaperone, TMD-containing effectors insert into target cell membranes to maintain effector stability and provide a conduit for cytotoxic effectors to enter the cytoplasm. To test this hypothesis, we used the candidate Tse6 effector, which possesses a glycohydrolase toxin domain ($Tse6_{tox}$) that hydrolyzes the essential dinucleotides NAD^+ and $NADP^+$ (Whitney et al., 2015). We employed a liposome-based assay in which we incubated NAD^+-loaded liposomes with the VgrG1a-Tse6-EF-Tu-Tsi6 complex and measured the levels of NAD^+ inside the liposomes. Negative stain EM of the liposomes clearly showed that the Tse6-loaded VgrG1 complexes decorated the liposomes (Figure 2.13B). In comparison to control liposomes, a significant decrease in NAD^+ concentration was observed in the presence of the Tse6-loaded VgrG1 complex, indicating successful translocation of the $Tse6_{tox}$ domain into the interior of the liposomes (Figure 2.13C). To ensure that the needle-shaped particles themselves do not puncture liposomes and cause leakage of NAD^+, we also tested complexes containing a catalytically inactive variant of Tse6 ($Tse6^{Q333D,\ D396A}$) or a variant of Tse6 lacking its C-terminal toxin domain ($Tse6_{\Delta tox}$). We observed no significant decrease in NAD^+ levels when incubating the mutant complexes with liposomes, indicating that the integrity of liposomes is not compromised by this complex (Figure 2.13C). To extend this analysis, we reconstituted lipid nanodiscs with the VgrG1a-Tse6-EF-Tu-Tsi6 complex and found

that following reconstitution, Tsi6 and EagT6 were no longer associated with the Tse6-loaded VgrG1 complex, indicating that both proteins spontaneously dissociate from the complex upon its insertion into the membrane (Figure 2.13D). These results demonstrate that Tse6 enters the membrane without the need of an additional receptor and self-translocates its toxin domain across the bilayer using its TMDs.

Collectively, our experimental data and informatics analyses indicate that unlike orphan PAAR proteins, which contain all the necessary molecular determinants for proper folding, prePAAR effectors may contain inherently unstable PAAR domains that require a prePAAR motif to ensure their proper folding and enable their interaction with their cognate VgrG protein. These effectors also contain TMDs that require Eag chaperones for stability in the donor cell, but upon injection into target cells, are released from these chaperones and insert into the inner membrane of target cells, facilitating transfer of their toxin domain into the target cell cytoplasm (Figure 2.14).

Discussion

The proper functioning of protein secretion systems requires the precise recruitment of effector proteins among hundreds of cytoplasmic proteins. Here we investigate the mechanism of recruitment for a widespread family of membrane protein effectors exported by the T6SS. Our work demonstrates that the N-terminal

region of these effectors possesses two structural elements that are critical for their delivery between bacterial cells by the T6SS apparatus.

First, this region contains TMDs, which interact with the Eag family of chaperones and are proposed to play a role in effector translocation across the inner membrane of recipient cells. Additionally, this region possesses prePAAR, which we show is required for the proper folding of PAAR, thereby facilitating the interaction of this domain with its cognate VgrG protein and enabling effector export by the T6SS.

As demonstrated in Figure 2.3, EagR1 and EagT2 from *P. protegens* interact with and stabilize RhsA and Tne2, respectively, but not vice versa. The inability of these chaperones to interact with non-cognate effectors can be explained by the two starkly contrasting binding modes observed in our co-crystal structures. The class I prePAAR effector and RhsA homolog Rhs1 interacts with its cognate chaperone SciW in an asymmetric manner whereas the class II effector and Tne2 homolog Tse6 adopts a pseudosymmetric binding mode whereby two separate alpha helices interact with each EagT6 chaperone protomer in a similar location. Our structural analyses suggest that Rhs1 residues F20 and F43 play a critical role in its asymmetric binding mode because the aromatic side chains of these amino acids insert into hydrophobic pockets present in SciW that are lacking in EagT6. The back of these hydrophobic binding pockets in SciW is formed by residue G108 whereas in EagT6 the equivalent space is occupied by the side chain of F104. As such, the F20 and F43 side chains of Rhs1's TMD are able to insert

into the hydrophobic pockets of SciW but are unable to do so with EagT6 because of a steric clash that would occur with F104. Conversely, the equivalent residue to F20 in the TMD of Tse6 is A26. The small side chain of alanine lacks the volume needed to fill the hydrophobic binding pocket of SciW and thus this interaction would likely contribute less free energy of binding with SciW. We expect that these same structural features contribute to the effector specificity observed for EagR1 and EagT2 because these chaperones also possess an alanine and phenylalanine at the corresponding positions in their hydrophobic pockets. Furthermore, the equivalent positions in the TMD of RhsA has similarly large hydrophobic residues (I23/L46) whereas the TMD of Tne2 has an alanine residue (A26) that aligns to A26 of Tse6. Although our structural models and bioinformatic analyses of chaperone-effector pairs suggest this pattern of Eag chaperone specificity is widespread, additional biochemical experimentation will be needed to demonstrate the generalizability of these findings.

Another important difference observed in the class I and class II chaperone-effector interactions is the position of the prePAAR region in each chaperone-effector pair. The asymmetric chaperone-TMD interaction allows SciW to ‘shield’ the hydrophobic regions of Rhs1’s N-terminus from the aqueous milieu while also positioning its prePAAR motif in such a way that would allow it to interact with PAAR. By contrast, the pseudosymmetric binding mode of Tse6 to EagT6 appears to be much more dynamic as interpretable electron density for bound Tse6 was only observed when the effector fragment was held in place by interactions with an

adjacent complex in the crystallographic asymmetric unit. Consequently, we speculate that even though the Tse6's prePAAR motif appears less accessible than that of Rhs1, it is likely highly dynamic in solution and thus may adopt a markedly different conformation when in complex with PAAR.

Despite containing a primarily beta-sheet secondary structure, Eag chaperones interact with effector TMDs by mimicking the interactions that occur between the helices of alpha-helical membrane proteins, which, to our knowledge, is a unique mechanism for a chaperone-effector interaction. Upon binding their cognate effector, we hypothesize that Eag chaperones not only shield effector TMDs from solvent but also distort their structure to prevent potential hairpin formation and erroneous insertion into the inner membrane of the effector-producing cell. Because Eag-interacting TMDs have likely evolved to insert into bacterial membranes, a mechanism to prevent self-insertion is probably necessary prior to export. Recent work studying the secretion of TMD-containing effectors of the bacterial type III and type IV secretion systems found that shielding TMDs to prevent inner membrane insertion is a critical step for proper targeting to the secretion apparatus (Krampen et al., 2018). However, membrane protein effectors of these secretion systems have evolved to target eukaryotic, not bacterial, membranes and thus may not require stringent control of TMD conformation prior to export. Indeed, unlike the Eag chaperones presented here, a previously studied T3SS chaperone was shown not to distort the conformation of effector TMDs,

whose conformation remained similar before and after membrane insertion (Nguyen et al., 2015).

Current evidence also suggests that Eag chaperones are not secreted by the T6SS (Cianfanelli et al., 2016a; Quentin et al., 2018). This leads to two important questions: 1) when do Eag chaperones dissociate from their cognate effector? 2) How do effector TMDs remain stable after their dissociation from the chaperone? Although no definitive answers exist for either of these questions, given that effector-chaperone interactions are maintained after effector-VgrG complex formation, chaperone dissociation presumably occurs immediately before or during a T6SS firing event. One way this could be accomplished is through chaperone interactions with components of the T6SS membrane and/or baseplate subcomplexes, which might induce chaperone-effector dissociation. The lumen of the T6SS apparatus may also serve to mitigate the susceptibility to degradation observed for prePAAR effectors in the absence of Eag chaperones because the inner chamber of the T6SS apparatus may shield effectors from the protein homeostasis machinery of the cell.

Crystal structures of single domain PAAR proteins suggest that this domain folds independently and is highly modular (Shneider et al., 2013). Indeed in many instances, PAAR domains appear in isolation (orphan PAAR) and do not require additional binding partners to interact with VgrG (Wood et al., 2019b). The initial characterization of PAAR domains established seven groups of PAAR proteins, with the most abundant being orphan PAARs (55% of 1353 PAAR proteins) while

the remaining groups represent PAAR proteins with N- and/or C-terminal extensions (Shneider et al., 2013). Our data demonstrate that PAAR domains with N-terminal extensions possess prePAAR, which we predict may be required for the proper folding of the downstream PAAR domain. Based on our structural modelling and sequence alignments, the ability of prePAAR to assist with PAAR domain folding may in part be due to its participation in coordinating the zinc ion found near the tip of this cone-shaped protein. Our sequence analysis also suggests that while orphan PAARs contain four zinc-coordinating histidine and/or cysteine residues, the PAAR domain of prePAAR effectors contains only three, suggesting that the fourth ligand required for tetrahedrally coordinated Zn^{2+} is provided by prePAAR. In this way, the PAAR domain of prePAAR effectors is split into two components, which likely come together to form a structure that can interact with VgrG and undergo T6SS-mediated export. One consequence of this 'split PAAR' domain arrangement is that the TMDs are tethered to PAAR via their N- and C-terminus, which would restrict the mobility of the TMDs and ensure their positioning on the surface of PAAR. We speculate that the proper arrangement of prePAAR effector TMDs on the surface of PAAR is likely critical for the ability of the T6SS spike complex to effectively penetrate target cell membranes during a T6SS firing event. Future studies focused on capturing high-resolution structural snapshots of assembled prePAAR-TMD-PAAR complexes will be needed to further support this proposed mechanism.

In summary, our mechanistic dissection of prePAAR effectors and their cognate chaperones has revealed fundamental new insights into bacterial toxin export and membrane protein trafficking. The unique ability of T6SSs to potently target a wide range of bacteria in a contact-dependent manner may permit their use in different biomedical applications, such as the selective depletion of specific bacterial species in complex microbial communities (Ting et al., 2020). An in-depth understanding of the mechanisms that that underlie T6SS effector recruitment and delivery will be of critical importance for such future bioengineering efforts.

Figures

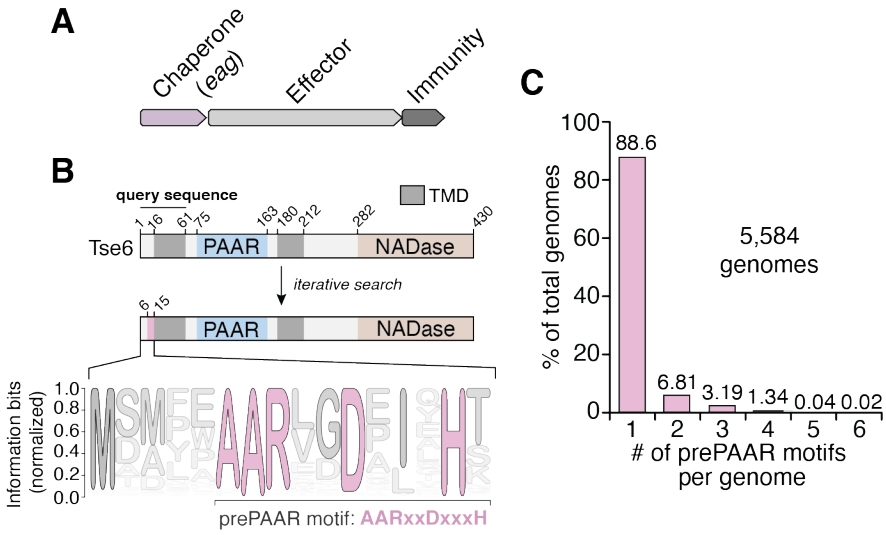


Figure 2.1. The prePAAR motif is conserved in PAAR- and TMD-containing effectors. A) Genomic arrangement of T6SS chaperone-effector-immunity genes for characterized effector-associated gene family members (*eag*; shown in purple), which encode DUF1795 domain-containing chaperones. **B)** Protein architecture and sequence logo for the prePAAR motif found in the N-terminus of Tse6. An alignment of 2,054 sequences was generated using the 61 N-terminal residues of Tse6 as the search query. The relative frequency of each residue and information content in bits was calculated at every position of the sequence and then normalized by the sum of each position's information bits. Transparency is used to indicate probability of a residue appearing at a specific position. Residues colored in pink correspond to the prePAAR motif: AARxxDxxxH. **C)** Genomes from genera of Proteobacteria known to contain functional T6SSs including: *Burkholderia*, *Escherichia*, *Enterobacter*, *Pseudomonas*, *Salmonella*, *Serratia*, *Shigella* and *Yersinia* were screened for unique prePAAR effectors. Percentage of total genomes that contained 1 to 6 prePAAR motifs is indicated.

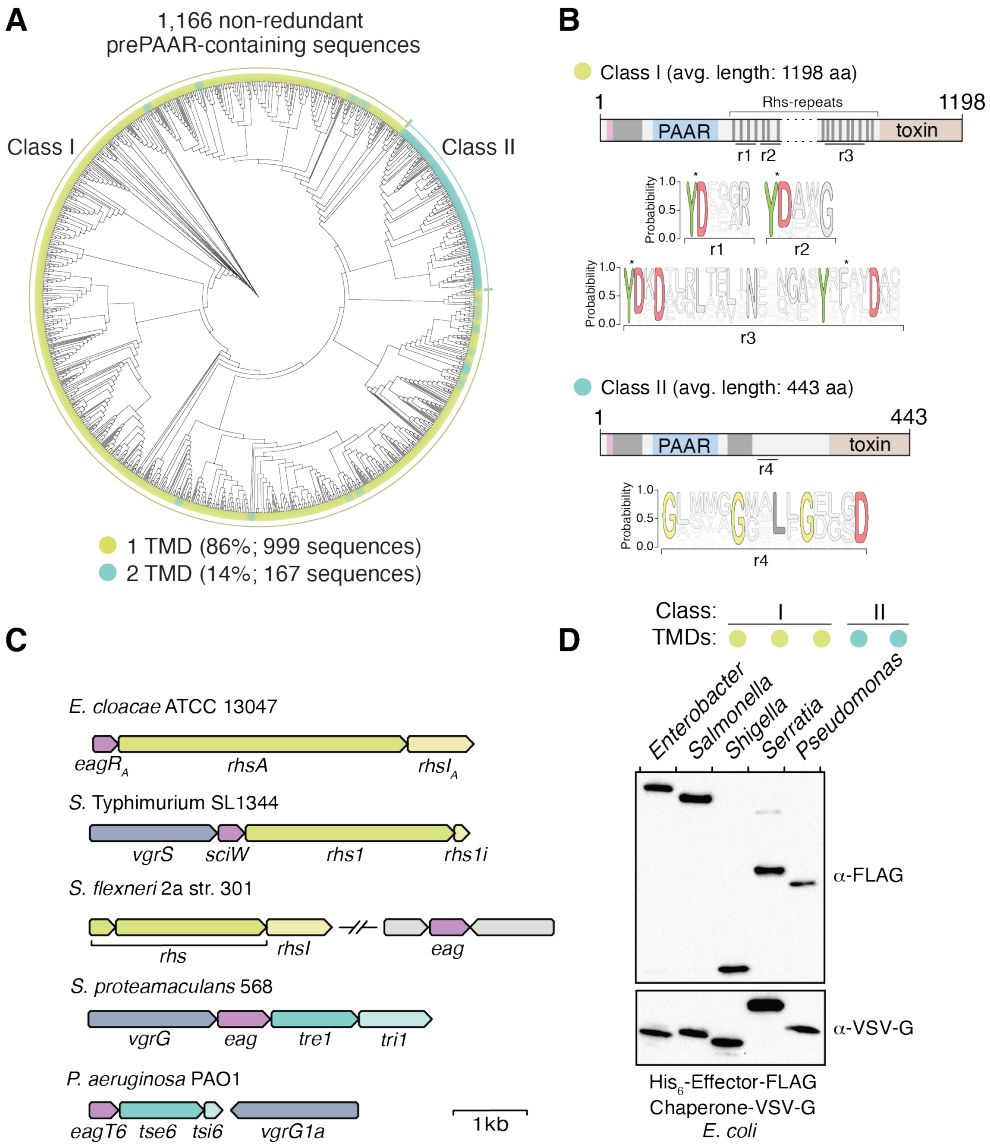


Figure 2.2. Two classes of prePAAR motif-containing effectors interact with Eag proteins. A) Phylogenetic distribution of 1,166 non-redundant prePAAR-containing effectors identified in Figure 2.1C. TM prediction algorithms were used to quantify the number of TM regions in each effector. The two classes that emerged are labeled in green (class I; 1 TM region-containing effectors) and blue (class II; 2 TM region-containing effectors). Branch lengths indicates evolutionary distances. **B)** Effector sequences within class I or class II were aligned and a sequence logo was generated based on the relative frequency of each residue at each position to identify characteristic motifs of both classes. Four different regions (r1–r4) after the PAAR and TM regions were found to harbor conserved residues. Class I effectors contain YD repeat regions (r1-3) characteristic of Rhs proteins,

whereas a GxxxxGxxLxGxxxD motif (r4) was identified in class II effectors. **C)** Genomic arrangement of the five chaperone-effector pairs used for co-purification experiments. Shading was used to differentiate effector (dark) from potential immunity (light) genes. Class I effectors and associated immunity genes are shown in green, class II effectors and associated immunity genes are shown in blue, *vgrG* genes are shown in dark blue, *eag* chaperone-encoding genes are shown in purple. Previously established names for each open reading frame are indicated below the gene diagram. Locus tags for each pair (e, effector; c, chaperone) are as follows: *Enterobacter* (e: ECL_01567, c: ECL_01566), *Shigella* (e: SF0266, c: SF3490), *Salmonella* (e: SL1344_0286, c: SL1344_0285), *Serratia* (e: Spro_3017, c: Spro_3016), *Pseudomonas* (e: PA0093, c: PA0094). Note that the Rhs component of the class I prePAAR effector SF0266 is encoded by the downstream open-reading frame SF0267. Scale bar indicates 1 kilobase. **D)** Western blot analysis of five effector-chaperone pairs described in **C**. Each pair was co-expressed in *E. coli* and co-purified using nickel affinity chromatography. The class and number of TM regions from each pair are indicated.

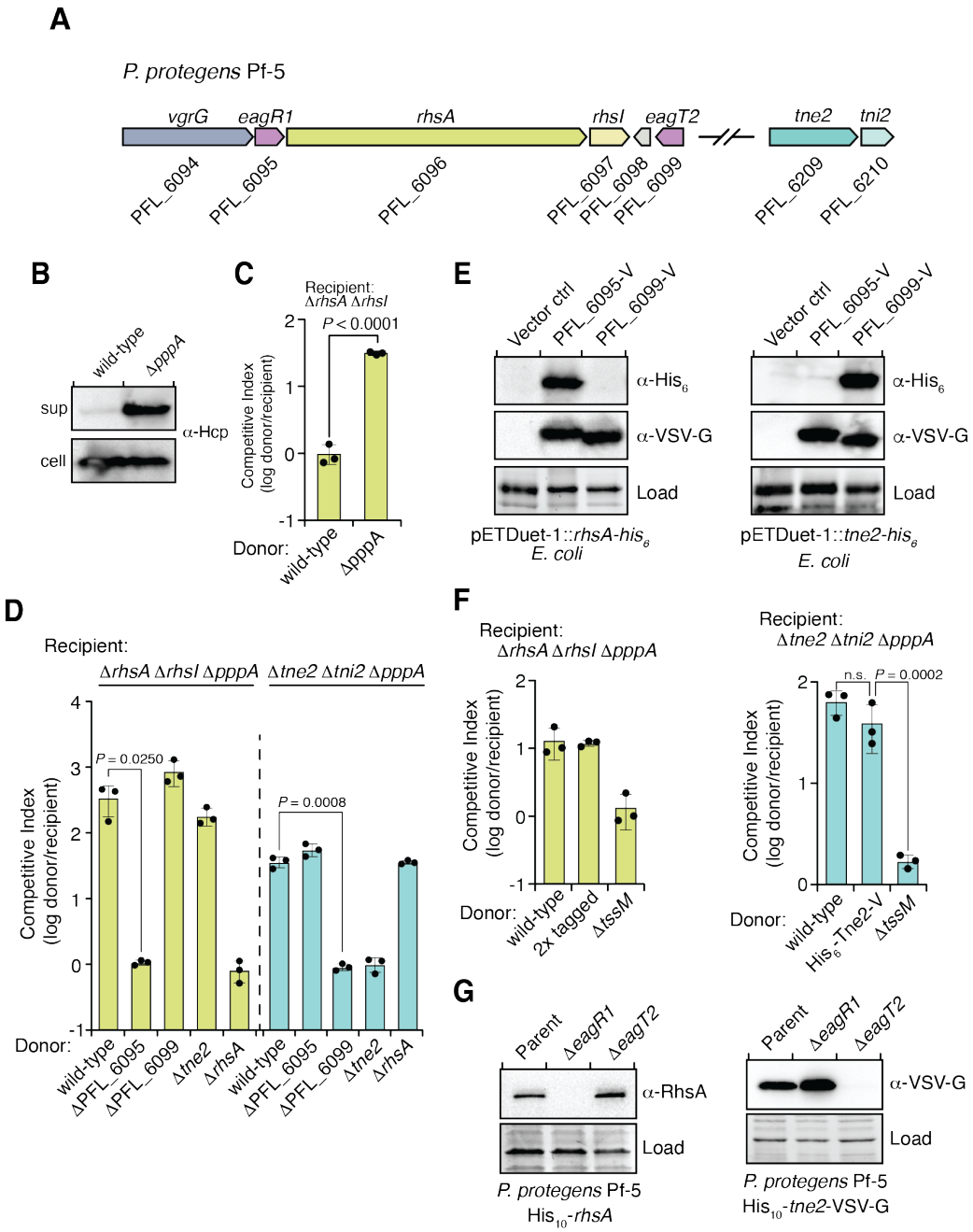


Figure 2.3. Eag chaperones are specific for their cognate prePAAR effector and are necessary for effector stability in vivo. A) Genomic context of two prePAAR-containing effector-immunity pairs from *P. protegens* Pf-5. RhsA is a class I effector (shown in green) and Tne2 is a class II effector (shown in blue). Shading is used to differentiate effector (dark) and immunity genes (light). Predicted *eag* genes are shown in purple. **B)** Western blot of supernatant (sup) and cell fractions of the indicated *P. protegens* Pf-5 strains grown to OD 0.8. An

Hcp (PFL_6089)-specific antibody was used to assess T6SS activity. **C)** Intraspecific growth competition assay of the indicated donor *P. protegens* strains against a recipient susceptible to intoxication by the class I prePAAR effector RhsA. Data are mean ± s.d. for *n* = 3 biological replicates; p value shown is from a two-tailed, unpaired t-test. **D)** Outcome of intraspecific growth competition assays between the indicated *P. protegens* donor and recipient strains. Donor strains were competed with recipient strains lacking *rhsA-rhsI* (green) or *tne2-tni2* (blue). Both recipients are lacking *pppA* to stimulate type VI secretion. Data are mean ± s.d. for *n* = 3 biological replicates and are representative of two independent experiments; p values shown are from two-tailed, unpaired t-tests. **E)** Western blot analysis of *E. coli* cell lysates from cells co-expressing the indicated effectors (RhsA, left or Tne2, right) and either empty vector, PFL_6095 V or PFL_6099 V. **F)** Growth competition assays between the indicated *P. protegens* donor strains and either RhsA (green, left) or Tne2 (blue, right) susceptible recipients. **G)** Affinity-tagged RhsA or Tne2 were purified from cell fractions of the indicated *P. protegens* strains and visualized using western blot analysis. Deletion constructs for each *eag* gene were introduced into each of the indicated parent backgrounds. A non-specific band present in the SDS-PAGE gel was used as a loading control. **E, G)** Data are representative of two independent experiments.

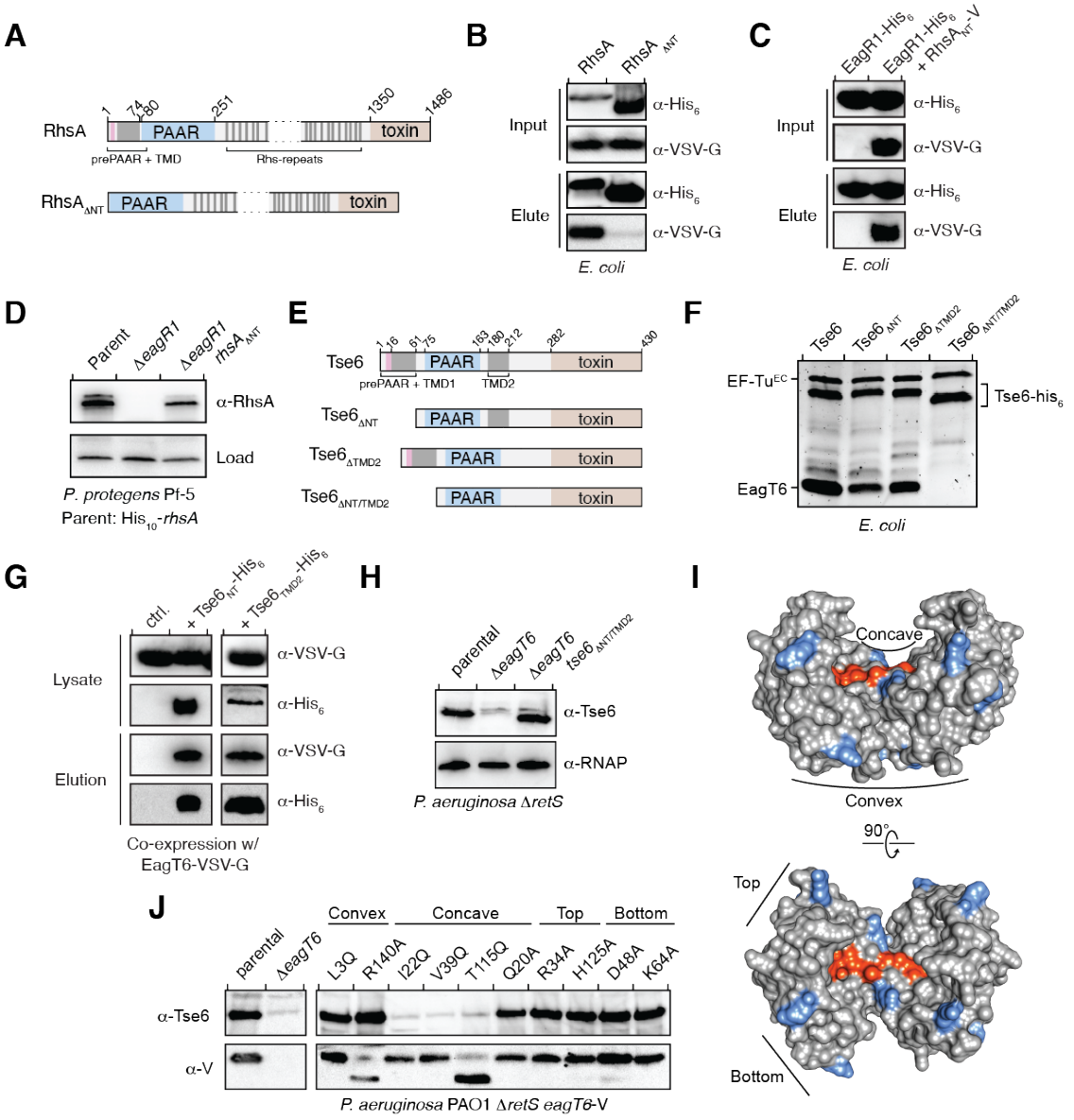


Figure 2.4. Eag chaperones promoter the stability of cognate effectors by interacting with their prePAAR- and TMD-containing regions. A) Domain architecture of *P. protegens* RhsA and a truncated variant lacking its prePAAR and TMD-containing N-terminus (RhsA$_{\Delta NT}$). **B)** EagR1 interacts with the N-terminus of RhsA. His$_6$-tagged RhsA or RhsA$_{\Delta NT}$ were co-expressed with EagR1 in *E. coli*, purified using affinity chromatography and detected by western blot. **C)** Western blot of lysate (input) and pull-down elution fractions of His$_6$-tagged EagR1 co-expressed with an empty vector or VSV-G tagged RhsA$_{NT}$ (residues 1–74) in *E. coli*. **D)** Affinity purification of chromosomally His$_{10}$-tagged RhsA or RhsA$_{\Delta NT}$ from cell fractions of the indicated *P. protegens* strains. The parent strain expresses chromosomally encoded His$_{10}$-tagged RhsA. The loading control is a non-specific

band on the blot. **E)** Domain architecture of *P. aeruginosa* Tse6 and truncated variants lacking its prePAAR and TMD1-containing N-terminus ($Tse6_{\Delta NT}$), TMD2 ($Tse6_{\Delta TMD2}$) or both regions ($Tse6_{\Delta NT/TMD2}$). **F)** SDS-PAGE analysis of His_6-tagged Tse6 or the indicated variants co-expressed with EagT6 in *E. coli*, purified using affinity chromatography and visualized using Coomassie Blue staining. **G)** Western blot of lysate and pull-down elution fractions of VSV-G-tagged EagT6 co-expressed with an empty vector (ctrl.) or the indicated His_6-tagged Tse6 fragment (NT (prePAAR+TMD1) or TMD2) in *E. coli*. **H)** Western blot of cell fractions harvested from parental or the indicated mutant *P. aeruginosa* strains grown to mid-log OD. Tse6 was detected using a Tse6-specific polyclonal antibody. An antibody that binds the α-subunit of RNA polymerase (RNAP) was used to detect RNAP for a loading control. The parental background is Δ*retS*. **I)** Structure of the EagT6 homodimer (PDB: 1TU1) with sites used for mutagenesis experiments highlighted on the concave, convex, top or bottom surfaces. Orange residues indicates sites that had a significant effect on the interaction of EagT6 with Tse6, while blue residues indicate sites that did not have an effect on this interaction. **J)** Western blot analysis of Tse6 and EagT6-VSV-G from *P. aeruginosa* parent or mutant strains expressing the indicated EagT6 variant. The surface that each residue maps to is indicated and shown in panel **I**. The parent *P. aeruginosa* genotype is Δ*retS eagT6*-VSV-G.

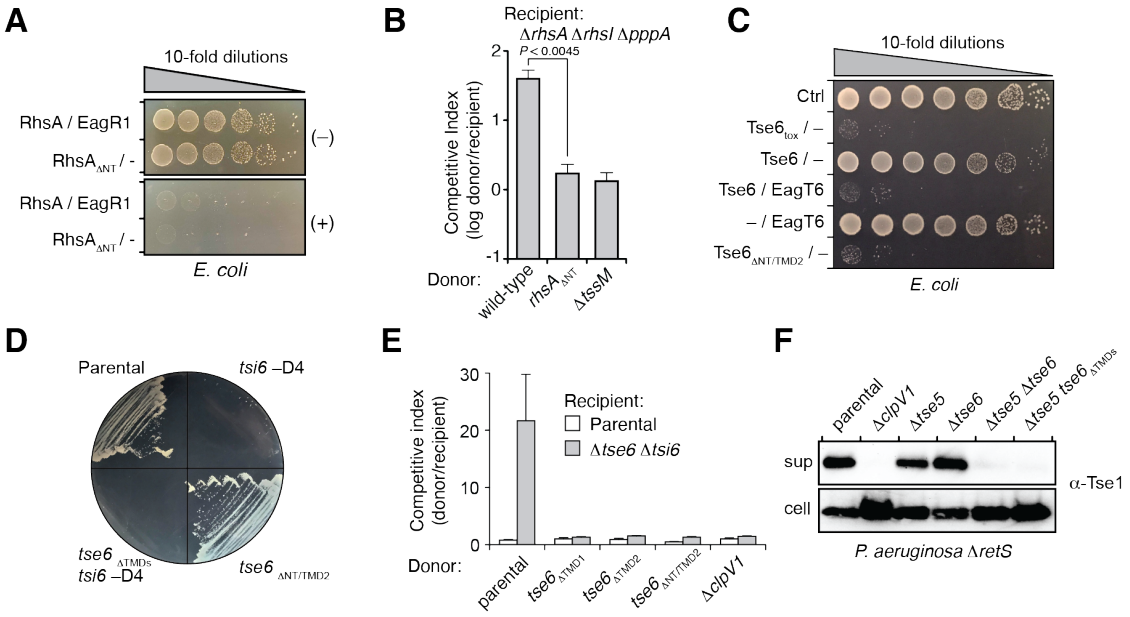


Figure 2.5. Chaperone-blind RhsA and Tse6 variants are toxic when expressed intracellularly, but not during interbacterial competition. A) Growth of *E. coli* co-expressing inducible plasmids harboring RhsA and EagR1 or RhsA$_{\Delta NT}$ with an empty vector. Overnight cultures were plated on media containing (+) or lacking (-) inducers and were imaged after 24 hours of growth. **B)** Outcome of growth competition assays between the indicated donor and recipient strains of *P. protegens*. Data are mean ± s.d. for *n* = 3 biological replicates; p value shown is from a two-tailed, unpaired t-test. **C)** Dilution plating of stationary-phase *E. coli* cultures co-expressing inducible plasmids harboring the indicated constructs on inducer-containing LB agar. **D)** Growth of *P. aeruginosa* parent or the indicated mutants on LB containing 250 μM IPTG. The parent strain is *P. aeruginosa* Δ*retS* Δ*sspB* pPSV39-CV::*sspB*. In the presence of IPTG, SspB is expressed and recognizes D4-tagged Tsi6 and facilitates its depletion by the ClpXP protease (McGinness et al., 2006). **E)** Outcome of growth competition assays between the indicated donor and recipient strains of *P. aeruginosa*. The parent background is *P. aeruginosa* Δ*retS*. Data are mean ± s.d. for *n* = 3 biological replicates. **F)** Detection of the H1-T6SS exported effector Tse1 by western blot (Hood et al., 2010). Cell or TCA-precipitated supernatant (sup) from indicated *P. aeruginosa* strains were prepared and a Tse1-specific antibody was used to detect Tse1.

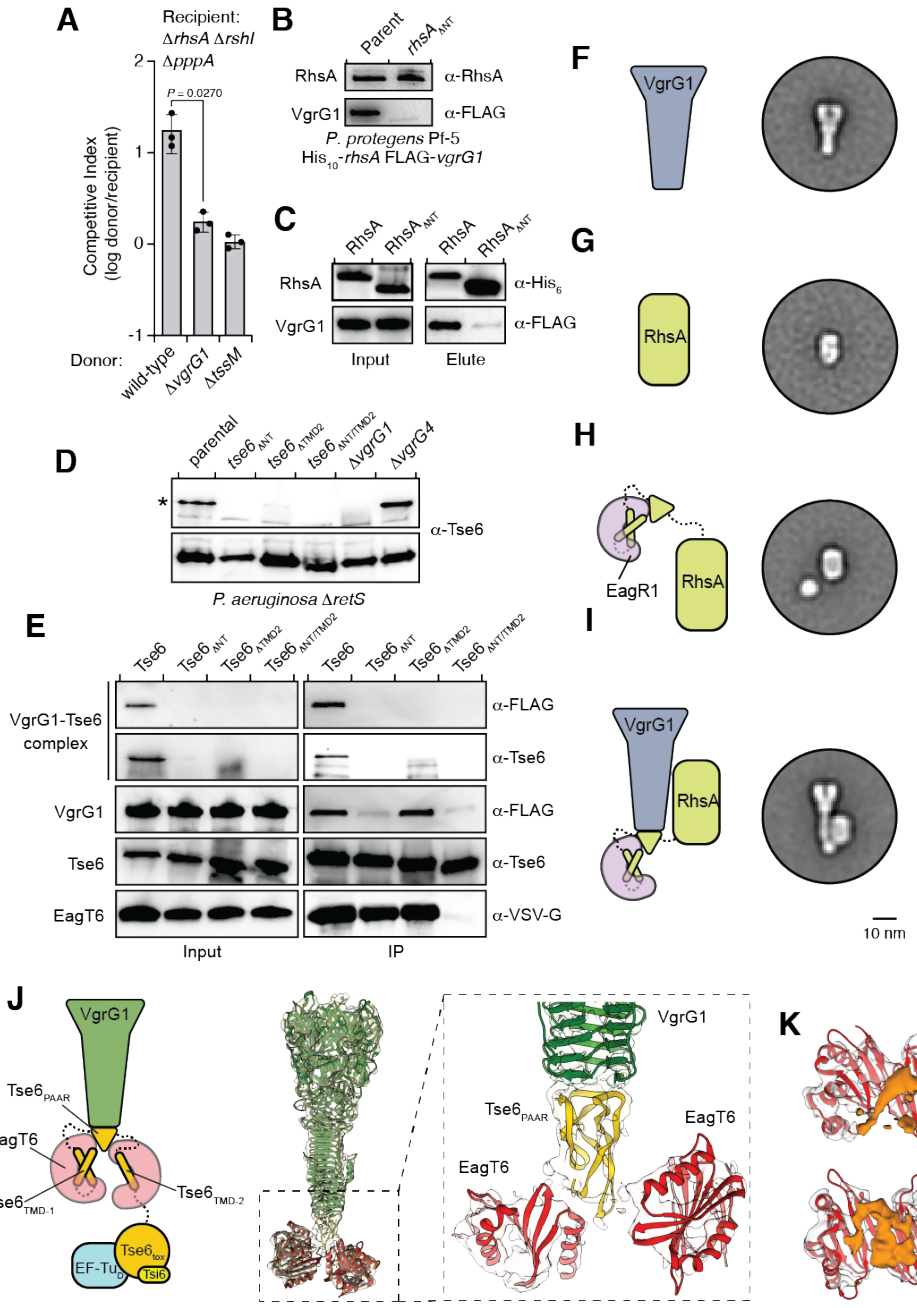


Figure 2.6. Chaperone-blind prePAAR effectors do not interact with their cognate VgrG spike. A) Outcome of growth competition assays between the indicated donor and recipient strains of *P. protegens*. Data are mean ± s.d. for *n* = 3 biological replicates; p value shown is from a two-tailed, unpaired t-test. **B)** Affinity purification of His_{10}-RhsA or His_{10}-$RhsA_{\Delta NT}$ from a *P. protegens* Pf-5 strain containing a chromosomally encoded FLAG epitope tag fused to *vgrG1*. FLAG-tagged VgrG1 and RhsA was detected by western blot. An RhsA-specific antibody

was used for RhsA detection. **C)** Affinity pull-down of His_6-tagged RhsA or $RhsA_{\Delta NT}$ co-expressed with VgrG-FLAG and EagR1-VSV-G in *E. coli*. Samples were analysed by western blot using the indicated antibodies. **D)** Western blot of Tse6 within cell fractions of the indicated *P. aeruginosa* strains. A Tse6-specific antibody can detect Tse6 in complex with VgrG1a (high molecular weight band, indicated by asterisk) or as a free monomer (low molecular weight band). The parental genotype is indicated. **E)** Detection of the indicated His_6-tagged Tse6 variants, VgrG1a-FLAG and EagT6-VSV-G by western blot in the input and co-immunoprecipitated (IP) fractions following co-expression in *E. coli*. **F-I)** Representative negative-stain EM class averages for purified VgrG1 (**F**), $RhsA_{\Delta NT}$ (**G**), EagR1-RhsA complex (**H**) and EagR1-RhsA-VgrG1 complex (**I**). Scale bar represents 10 nm for all images. All proteins were expressed and purified from *E. coli*. See Figure A2.1 in the Appendix for unprocessed micrographs and other class averages. **J)** Schematic (left) and cryo-EM density map fit with available atomic structures (middle) of the VgrG1-Tse6-$EagT6_2$-EF-Tu-Tsi6 complex and a zoom of the interface between EagT6 dimers and the Tse6 PAAR domain (right). The Tse6 toxin domain, EF-Tu and Tsi6 are not well resolved due to excessive flexibility within this part of the complex. **K)** Side view EagT6 dimers from the cryo-EM model in **J** fit with EagT6 atomic model (PDB: 1TU1) to highlight the TMD-interacting cavity. Density within the EagT6 cavity indicating is highlighted in orange.

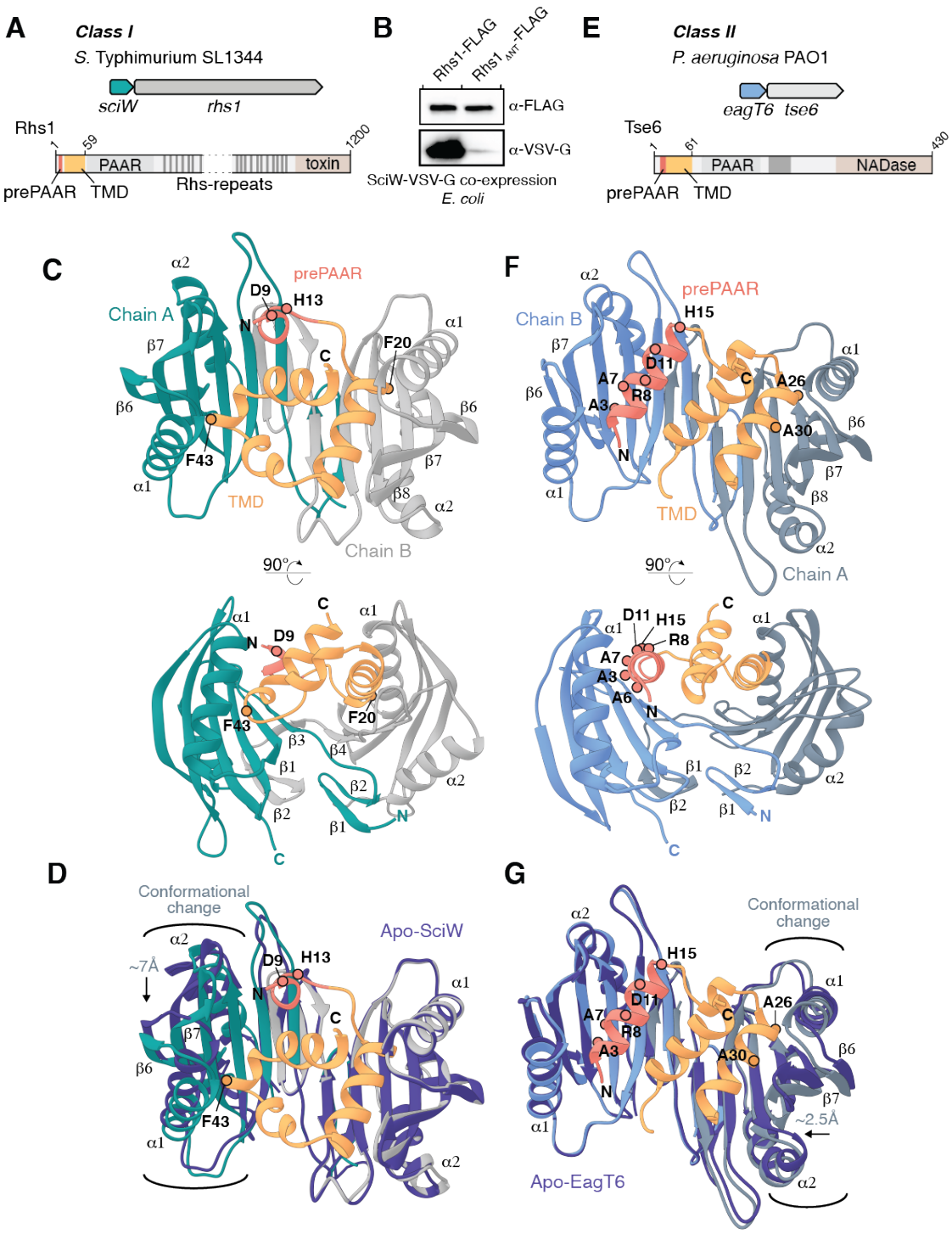


Figure 2.7. Co-crystal structures of the N-terminus of class I and class II prePAAR effectors in complex with their cognate Eag chaperones. A) Top, the Eag chaperone SciW (green) and Rhs1 (dark gray) class I chaperone-effector pair from *Salmonella* Typhimurium used for X-ray crystallography. Bottom, Rhs1

protein architecture highlighting prePAAR (red) and TMD (orange) within residues 1-59 of the effector. **B)** Western blot of FLAG-tagged Rhs1 or the $Rhs1_{\Delta NT}$ variant (lacking residues 1-59) and VSV-G tagged SciW following co-expression in *E. coli* and affinity pulldown. **C)** An X-ray crystal structure of the SciW chaperone bound to the N-terminus of Rhs1 ($Rhs1_{NT}$, residues 8–57 are modeled) shown in two views related by a 90° rotation. **D)** Structural overlay of the apo-SciW structure with SciW-$Rhs1_{NT}$ complex demonstrates that a considerable conformational change in SciW occurs upon effector binding. **E)** Top, the EagT6 (blue) and Tse6 (light gray) class II chaperone-effector pair from *P. aeruginosa* used for X-ray crystallography. Bottom, Tse6 protein architecture highlighting prePAAR (red) and TMD (orange) within residues 1-61 of the effector. **F)** An X-ray crystal structure of the EagT6 bound to the N-terminus of Tse6 ($Tse6_{NT}$, residues 1–38 and 41–58 are modeled) shown in two views related by a 90° rotation. **G)** Structural overlay of the apo-EagT6 structure (PDB: 1TU1) with the EagT6-$Tse6_{NT}$ complex shows a minor conformational change in EagT6 occurring upon effector binding. **C-D, F-G)** Eag chaperones are colored by chain, N-terminal transmembrane domains (TMDs) are colored in orange, the prePAAR motif in red, and apo chaperone structure in dark blue. Positions of residues of interest in the effector N-terminal regions are labeled. See Table 2.4 for crystal structure data collection and refinement statistics.

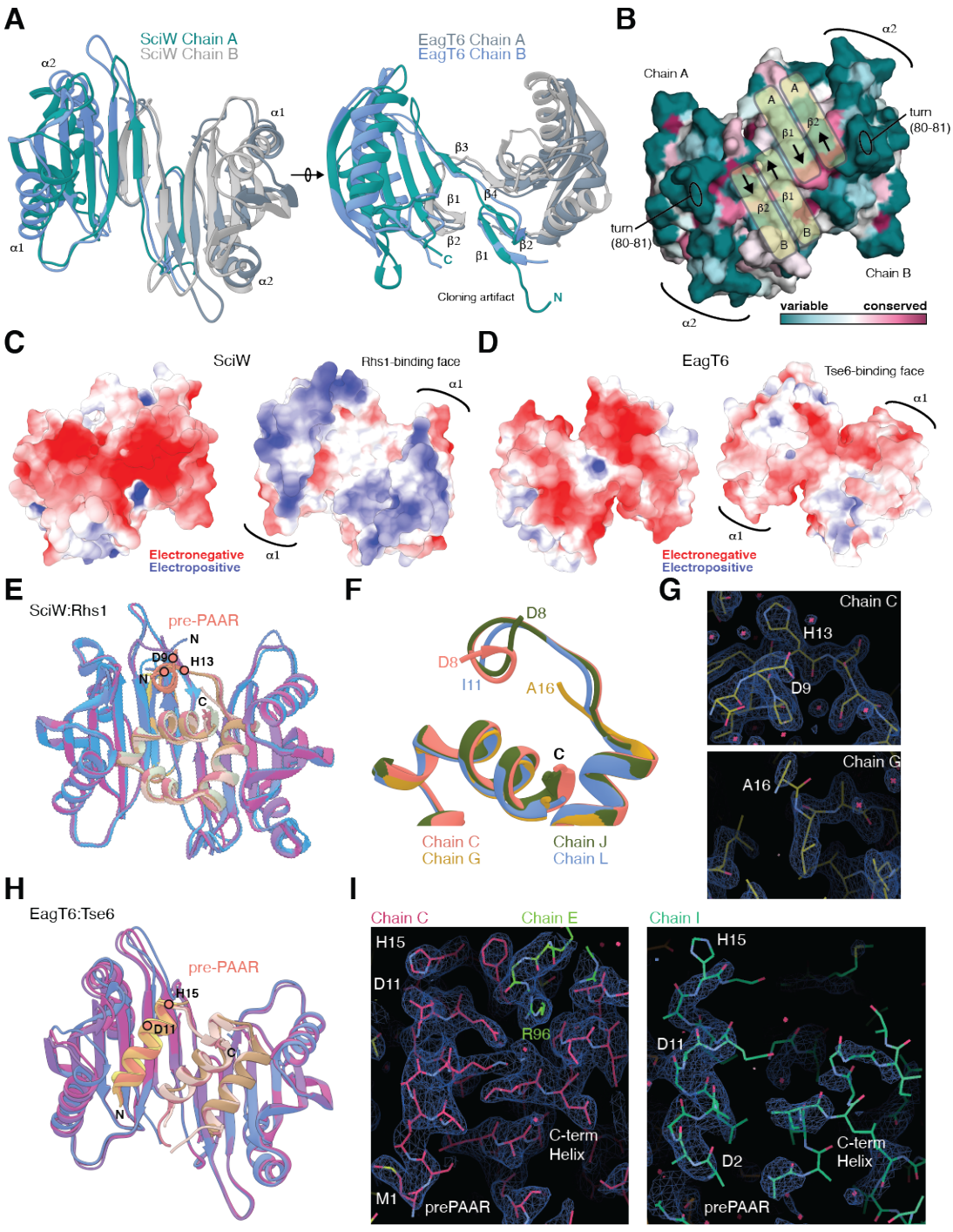


Figure 2.8. Structural comparison of Eag chaperones and effector complexes. A) Structural comparison of apo-SciW and apo-EagT6. Two views are shown related by 90° rotation. Each chaperone is colored by chain as in Figure 2.7. **B)** Conserved surface residues in SciW and EagT6 as determined by the

Consurf server (Ashkenazy et al., 2016). The domain-swap created by the beta-strands from chain A and chain B are labeled and shown with yellow bar overlays. **C)** Electrostatic surface potential of apo-SciW. The convex (left) and concave Rhs1 binding surfaces (right) are shown. **D)** Electrostatic surface potential of apo-EagT6. The convex (left) and concave (Tse6 binding) surfaces (right) are shown. **E)** Structural overlay of the four SciW-$Rhs1_{NT}$ complexes in the asymmetric unit of the crystal structure. The modeled prePAAR and C-terminus of Rhs1 are indicated and colored by chain. **F)** View of the Rhs1 prePAAR region of each complex in the crystal structure. The N-terminal residue for each chain is listed. **G)** Electron density maps of SciW-$Rhs1_{NT}$ Chain C and Chain G contoured at 1.4 rmsd (0.6816e/Å^3). **H)** Structural overlay of the three EagT6-$Tse6_{NT}$ complexes in the asymmetric unit of the crystal structure. The modeled prePAAR and C-terminus of Tse6 are indicated and colored by chain. **I)** Electron density maps of EagT6-$Tse6_{NT}$ Chain C and Chain I contoured at 1.2 rmsd (0.0344e/Å^3). The prePAAR and modelled C-terminal helix of the TMD region are labeled. A crystal packing artifact from Chain E including residue R96 that locks the prePAAR-TMD into place is shown. Electrostatic surface potentials were calculated by the adaptive-Poisson-Boltzmann server. Potentials are colored from −5 to 5 kT/e at pH 7.0. Images were created using UCSF Chimera, Coot, and Pymol.

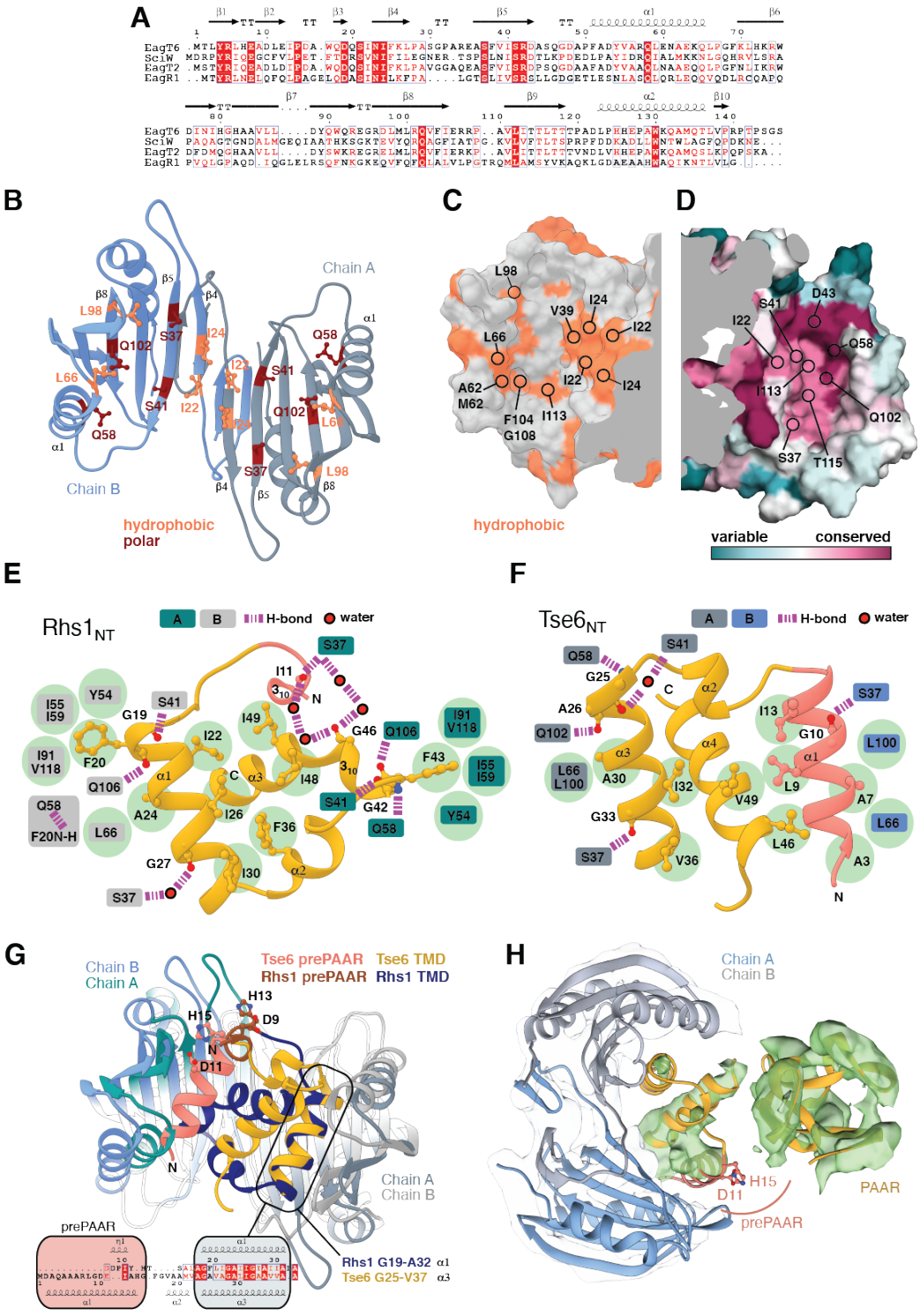


Figure 2.9. Eag chaperones interact with effector TMDs by mimicking interhelical interactions of alpha helical membrane proteins. A) Alignment of Eag chaperones that interact with class I (SciW, EagR1) or class II (EagT6 and

EagT2) prePAAR effectors. Secondary structure elements are shown above the alignment. **B)** Hydrophobic (orange) and polar residues (dark red) that are conserved among SciW, EagR1, EagT6, and EagT2 are shown. Residue numbers are based on EagT6. **C-D)** The conserved hydrophobic molecular surface of the chaperones in **B** is shown in light orange **(C)** and their molecular surface residue conservation is shown as determined by the Consurf server **(D)** (Ashkenazy et al., 2016). Conserved residues making contacts with the TMDs in both co-crystal structures are shown. **E)** Molecular contact map of $Rhs1_{NT}$ (residues 1–59) and SciW. prePAAR is shown in pink and the TMD regions in gold. Amino acids making contacts with the conserved residues of the Eag chaperones are shown by side chain/and or by main chain atoms (red for oxygen and blue for nitrogen). Residues in the Eag chaperone are highlighted by color of chain A or B. Polar (H-bond) contacts are drawn with a purple dashed line and are made with the side chain of the listed Eag residue. Outlined red circles indicate a water molecule. Light green circles indicate van der Waals interactions and hydrophobic interactions. The central group of hydrophobic residues without a listed chaperone residue all pack into the Eag hydrophobic face in Figure 2.9C (EagT6 I22/24 and V39). **F)** Molecular contact map of $Tse6_{NT}$ (residues 1–61) and EagT6. Schematic is the same as panel **E**. Q102 in EagT6 corresponds to Q106 in SciW. **G)** Structural alignment of SciW-$Rhs1_{NT}$ and EagT6-$Tse6_{NT}$ co-crystal structures using the structurally conserved TM helix as a reference. Eag chain colouring is the same as Figure 2.7. $Rhs1_{NT}$ is colored in dark blue with a brown prePAAR and $Tse6_{NT}$ in gold with a pink prePAAR. The conserved solvent accessible prePAAR residues D9/11 and H13/15 are shown in ball and stick model. Inset sequence alignment reflects the structurally aligned residues of $Rhs1_{NT}$ (top) and $Tse6_{NT}$ (bottom) as calculated by UCSF Chimera (Pettersen et al., 2004). Secondary structural elements are labeled. **H)** Docking of the EagT6-TMD crystal structure from Figure 2.7F into the previously obtained cryo-EM density map of the $EagT6_2$-Tse6-EF-Tu-Tsi6-VgrG1a complex (EMD-0135). Cryo-EM density corresponding to EagT6 is depicted in transparent gray and Tse6-TMD and Tse6-PAAR in green; prePAAR residues are shown in pink. Note that Tse6-TMD was docked independent of EagT6 into the Tse6 density. One of three possible orientations for the PAAR domain is shown.

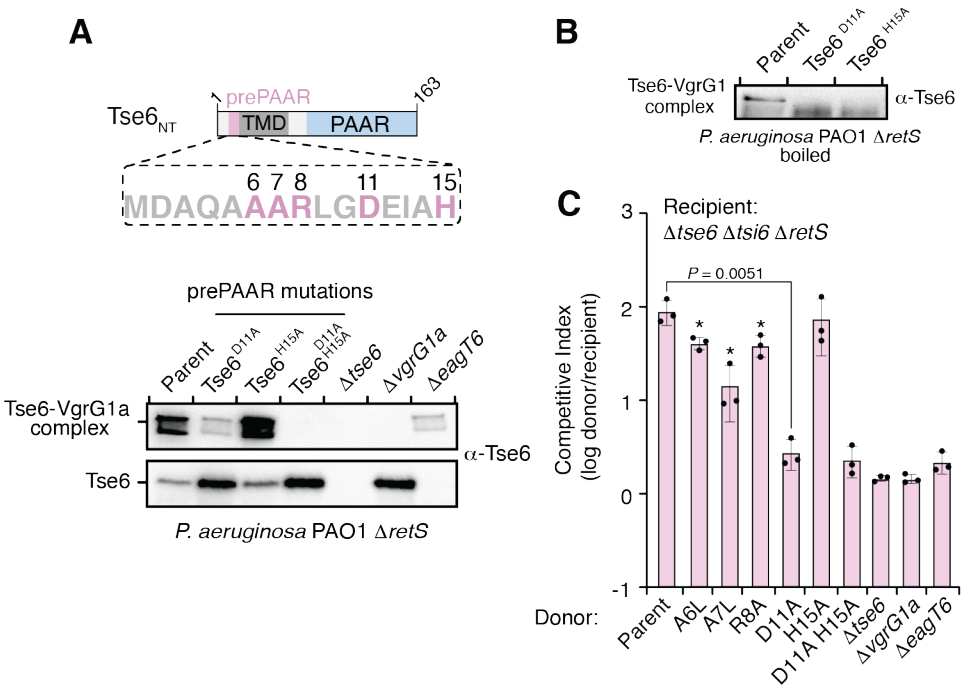


Figure 2.10. prePAAR is required for PAAR domain interaction with the VgrG spike. A) Western blot analysis of Tse6 from cell fractions of the indicated *P. aeruginosa* strains. Low-molecular-weight band indicates Tse6 alone whereas high-molecular-weight band indicates Tse6-VgrG1a complex. The parental strain contains a Δ*retS* deletion to transcriptionally activate the T6SS (Goodman et al., 2004). Schematic shows the N-terminal construct of Tse6 ($Tse6_{NT}$), prePAAR is indicated in pink. **B)** Same western blot analysis as shown in **A** except samples were boiled prior to electrophoresis. **C)** Outcome of growth competition assay between the indicated *P. aeruginosa* donor and recipient strains. The parent strain is *P. aeruginosa* Δ*retS*. Data are mean ± s.d. for *n* = 3 biological replicates; p value shown for the D11A mutant is from a two-tailed, unpaired t-test; asterisks indicates other prePAAR mutants that were also significantly different from the parent strain ($p < 0.05$).

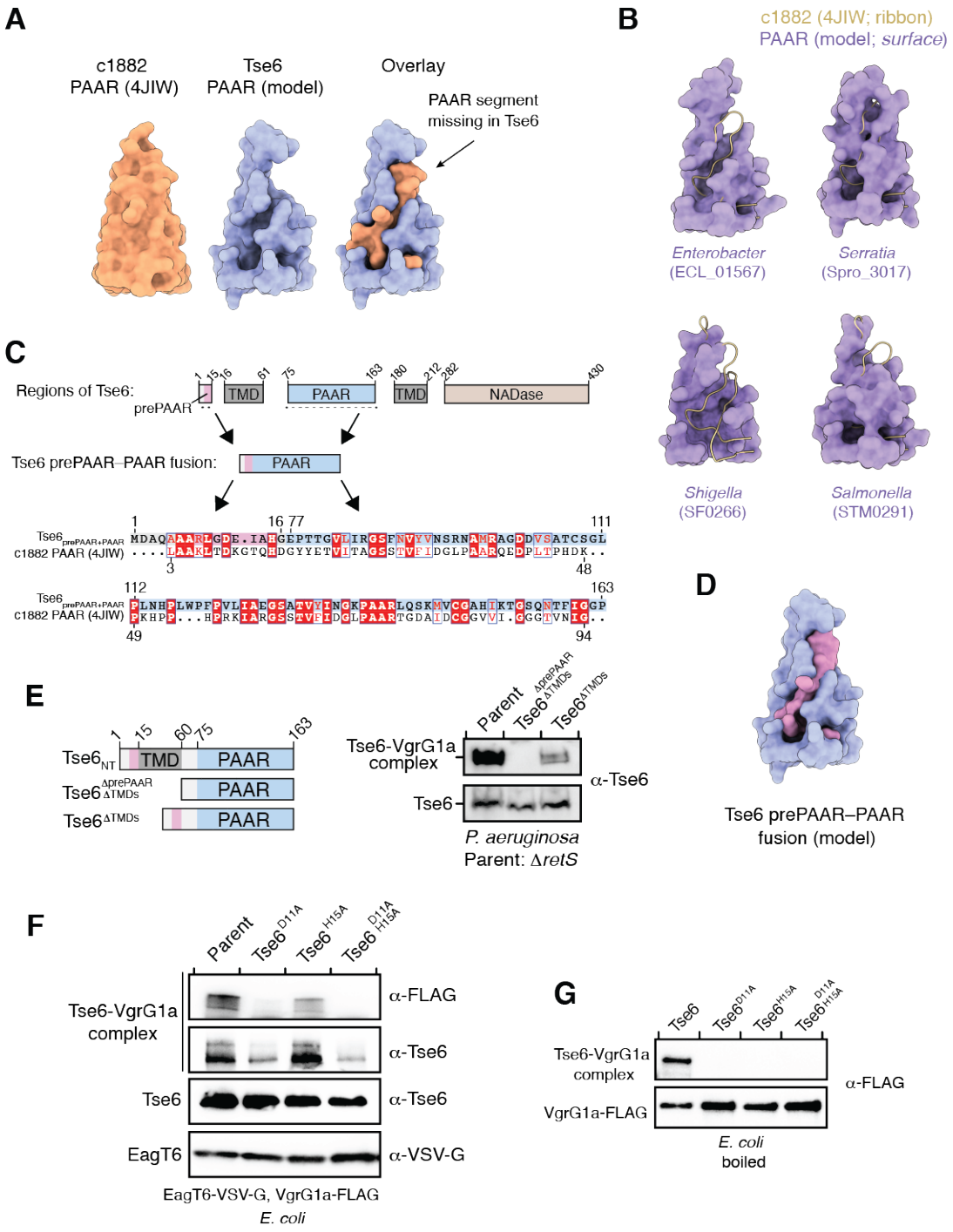


Figure 2.11. The PAAR domain of prePAAR effectors lacks an N-terminal segment. A) Structural comparison of the c1882 PAAR protein from *E. coli* (PDB: 4JIW) with a model of the PAAR domain of Tse6 generated using Phyre2 (Kelley et al., 2015). The overlay shows the additional N-terminal segment present in

c1882 that is absent in Tse6. **B)** Surface representation of structural models of the PAAR domain from each of the indicated prePAAR effector proteins (purple) overlaid with a ribbon representation of the c1882 PAAR protein crystal structure (beige). Structural alignments were performed using ChimeraX. **C-D)** Schematic showing the residue boundaries of the different regions of Tse6. The prePAAR (pink) and PAAR (blue) sequences were artificially fused to create $Tse6_{prePAAR+PAAR}$ and used to generate an alignment with c1882 **(C)** and a structural model **(D)**. Pink space-filling representation indicates the region of the model comprised of prePAAR. **E)** Detection of Tse6 in cell fractions of the indicated *P. aeruginosa* strains by western blot. The identity of the low-molecular weight and high-molecular weight bands are the same as described in Figure 2.10. The parent strain contains a *retS* deletion. Schematic shows mutants of Tse6 natively expressed in *P. aeruginosa*. Only the N-terminus of Tse6 is shown for simplicity. prePAAR is indicated in pink. Constructs lacking TMD1 also lack TMD2. **F)** His_6-tagged Tse6 or the indicated variants were co-expressed with FLAG-tagged VgrG1a and VSV-G-tagged EagT6 in *E. coli*. Following affinity pulldown, elution fractions were subject to western blot to detect the indicated epitope-tagged proteins. **G)** Same samples from **F**, except samples were boiled before being subject to electrophoresis.

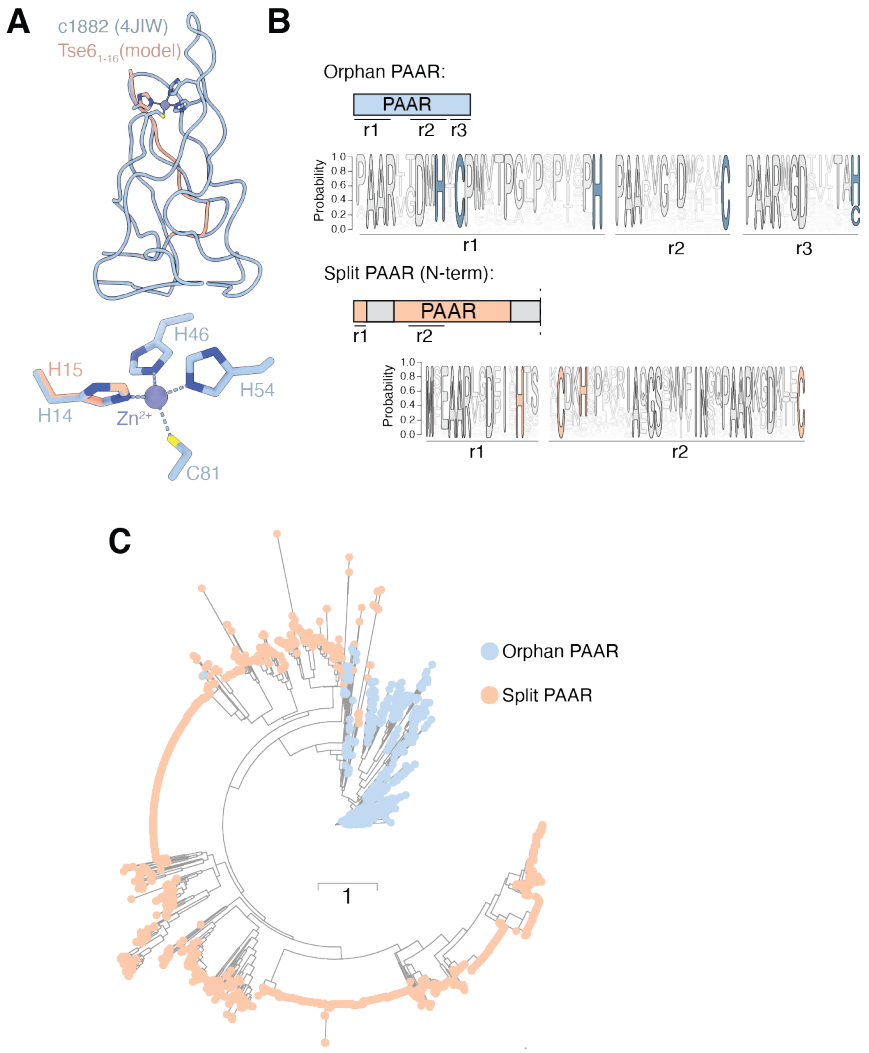


Figure 2.12. prePAAR effectors possess 'evolved' PAAR domains. A) Structural overlay of the prePAAR segment (peach) from the artificially fused Tse6$_{prePAAR+PAAR}$ sequence in Figure 2.11 with the entire c1882 PAAR protein (blue). The zoom shows the Zn^{2+}-coordinating residues of c1882 and the overlap of H15 from Tse6's prePAAR with H14 of c1882. **B)** Sequence logos developed from multiple sequence alignments of 564 orphan PAAR sequences and the N-terminus of 1,765 prePAAR-containing effectors. Sequence logos were developed for different regions (r1, r2, r3) in each construct that contained the Zn^{2+}-coordinating residues histidine and cysteine (Shneider et al., 2013). H and C residues are coloured blue in orphan PAAR sequence logos and coloured peach in split PAAR sequence logos. **C)** Phylogenetic distribution of 564 orphan PAAR sequences (blue) and 1,765 split PAAR (peach) sequences. The scale bar indicates the substitutions per base.

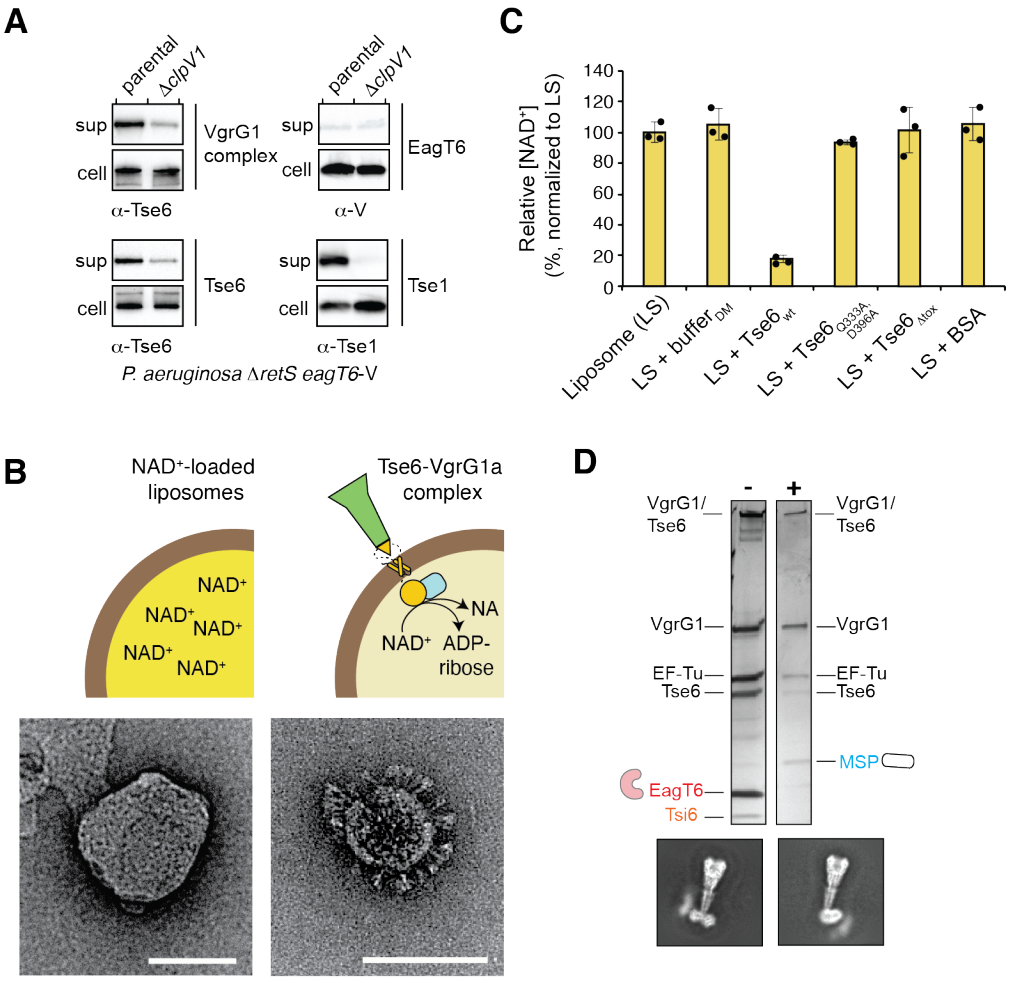


Figure 2.13. The TMDs of Tse6 spontaneously insert into membranes and enable self-translocation of its NADase toxin domain. A) Western blot of Tse6, the Tse6-VgrG1a complex, EagT6-VSV-G or Tse1 from supernatant (sup) and cell fractions of either the parent or the indicated Δ*clpV1 P. aeruginosa* strains. Tse6 and Tse1 were detected using antibodies specific for either effector. The parental background is indicated. **B)** Schematic (top) or micrographs (bottom) of NAD^+-loaded liposomes before (left) and after treatment with the Tse6-VgrG1a complex. Following toxin translocation into the liposome, NAD^+ is rapidly hydrolysed to ADP-ribose and nicotinamide (NA). Scale bars indicate 100 nm. **C)** NAD^+ levels in liposomes following treatment with the indicated conditions. DM-containing buffer ($buffer_{DM}$) condition is to control for detergent-dependent lysis of liposomes. **D)** Silver stained SDS-PAGE gel of the Tse6-loaded VgrG1a complex in its pre-firing conformation (−) and reconstituted in lipid nanodiscs (+) (see Methods for details).

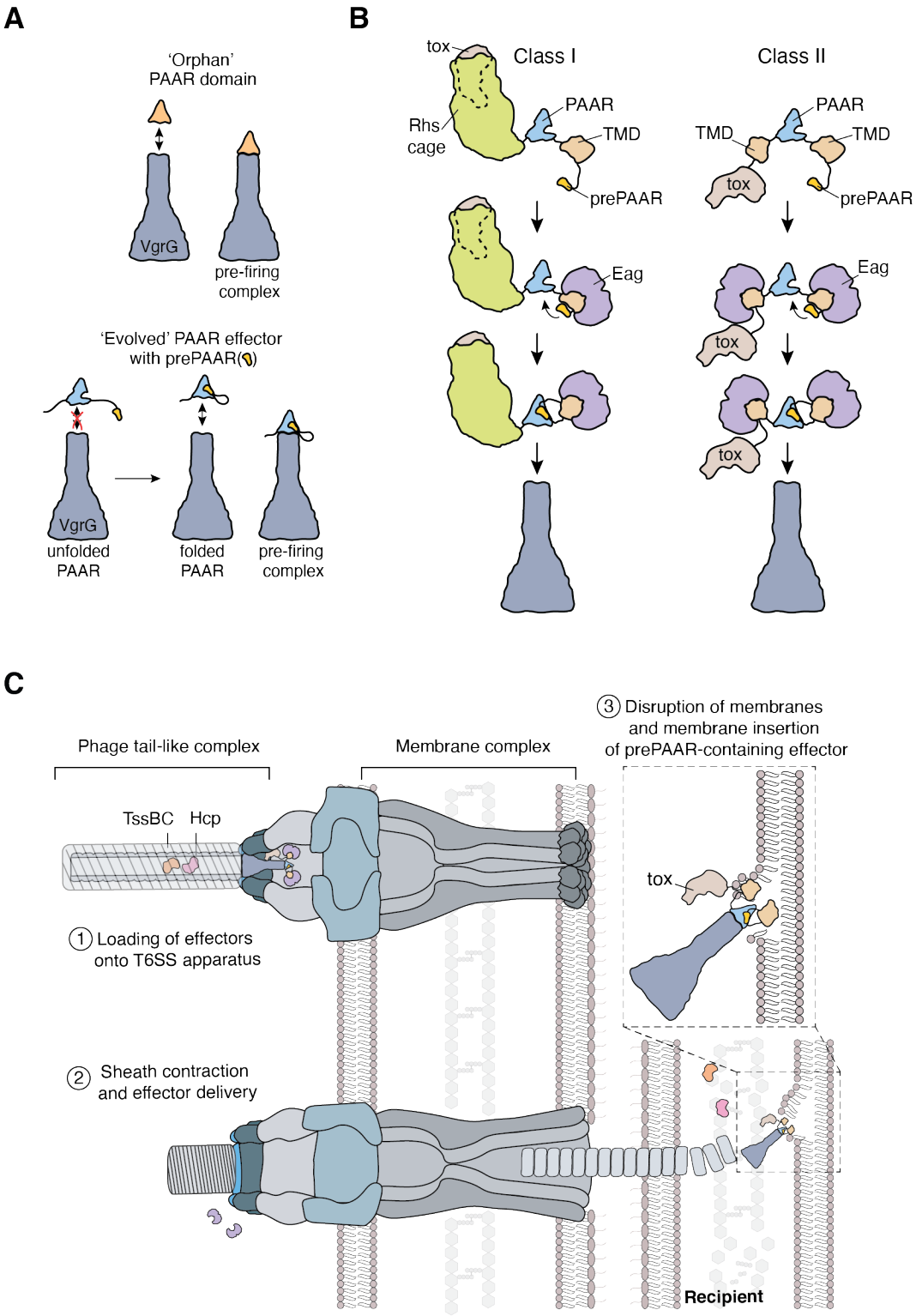

A
'Orphan' PAAR domain
VgrG
pre-firing complex
'Evolved' PAAR effector with prePAAR
VgrG
unfolded PAAR
folded PAAR
pre-firing complex
B
tox
Class I
Rhs cage
PAAR
TMD
prePAAR
Eag
Class II
TMD
PAAR
TMD
tox
prePAAR
Eag
tox
tox
C
Phage tail-like complex
Membrane complex
3 Disruption of membranes and membrane insertion of prePAAR-containing effector
TssBC
Hcp
tox
1 Loading of effectors onto T6SS apparatus
2 Sheath contraction and effector delivery
Recipient

Figure 2.14. Model depicting the role of Eag chaperones and prePAAR in type VI secretion. A) PAAR proteins exist with or without prePAAR domains. Those that lack prePAAR (orphan), can interact with VgrG and form a functional T6SS spike complex without any additional factors. By contrast, prePAAR-containing effectors contain multiple domains (evolved PAAR effectors) and likely require the prePAAR motif for proper folding of the PAAR domain and thus, loading onto the T6SS apparatus. **B)** prePAAR-containing effectors can be divided into two classes: class I effectors have a single TMD and contain a C-terminal toxin domain that is likely housed within a Rhs cage whereas class II effectors contain two TMDs and do not possess a Rhs cage. TMD-chaperone and prePAAR-PAAR interactions are required for effector stability and VgrG interaction, respectively, for both classes of prePAAR effectors. **C)** Depiction of a prePAAR-containing effector being exported by the T6SS into recipient cells. Inset shows the hydrophobic TMDs of a class II prePAAR effector disrupting the inner membrane of the target bacterium to allow entry of the effector toxin domain into the cytoplasm.

Tables

Table 2.1. List of prePAAR motif-containing proteins identified in the UniProtKB Database (provided as Supp_File_1_UniprotKB_prePAAR _D01.xlsx file). The document contains two separate sheets. List A corresponds to 2,054 prePAAR-containing sequences that were identified through an iterative search of the UniprotKB using $Tse6_{NT}$. List B corresponds to 975 sequences collected following filtering of list A (see Methods for details). The table can be downloaded here: https://cdn.elifesciences.org/articles/62816/elife-62816-supp1-v3.xlsx

Table 2.2. List of prePAAR motif-containing proteins from assembled genomes of all species belonging to the genera *Burkholderia*, *Escherichia*, *Enterobacter*, *Pseudomonas, Salmonella*, *Serratia, Shigella* and *Yersinia* (provided as Supp_File_2_8_genera_prePAAR_D01.xlsx file). The document contains two separate sheets. List C corresponds to 6,101 prePAAR-containing sequences that were identified through an iterative search of the UniprotKB using $Tse6_{NT}$. List D corresponds to 1,166 sequences collected following filtering of list C (see Methods for details) The table can be downloaded here: https://cdn.elifesciences.org/articles/62816/elife-62816-supp2-v3.xlsx

Table 2.3. Summary of the number of genomes and effector sequences used in our informatics analyses (provided as Supp_File_3_methods_D01.xlsx file). This document contains three separate sheets. The "UniprotKB-effectors" sheet shows the quantity of initial prePAAR-containing sequences that were identified in our search and the number of sequences that were used following filtering and removal of redundancy. The numbers in bold indicate the number of sequences in Table 2.1. The "8 genera - genomes" sheet corresponds to the number of genomes from the 8 genera (*Burkholderia*, *Escherichia*, *Enterobacter*, *Pseudomonas, Salmonella*, *Serratia, Shigella* and *Yersinia*) that contained one prePAAR-containing sequence and the number that remained following filtering and removal of redundancy. The "8-genera – effectors" sheet corresponds to initial and final numbers of prePAAR-containing sequences that were identified in the 8 genera listed above. The final number of sequences in this sheet were used to construct the cladogram in Figure 2.2A. The numbers in bold indicate the numbers of sequences in the lists in Table 2.2. The table can be downloaded here: https://cdn.elifesciences.org/articles/62816/elife-62816-supp3-v3.xlsx

Table 2.4. X-ray data collection, phase and structure refinement statistics for the crystal structures presented in chapter II.

	SciW (native)	SciW (Iodide)	SciW-Rhs1$_{1-59}$	EagT6-Tse6$_{1-61}$
Data Collection				
Wavelength (Å)	1.5418	1.5418	0.97895	1.5418
Space group	$P2_12_12_1$	$P2_12_12_1$	$P3_12_1$	$P3_2$
Cell dimensions				
a, b, c (Å)	55.27 75.1 76.6	55.6 75.3 76.4	105.3 105.3 248.4	68.9 68.9 173.1
α, β, γ (°)	90 90 90	90 90 90	90 90 120	90 90 120
Resolution (Å)	29.03-1.75 (1.82-1.75)	19.63-2.21 (2.33-2.21)	91.20-1.90 (1.98-1.90)	28.22-2.55 (2.65-2.55)
Unique reflections	32309 (3162)[a]	29933 (4888)	126298 (12473)	29267 (2832)
CC(1/2)	99.8 (89.1)	99.6 (81.4)	99.9 (53.9)	99.6 (52.8)
R_{merge} (%)[b]	6.2 (91.3)	6.1 (44.7)	5.7 (34.6)	15.5 (179.8)
$I/\sigma I$	14.2 (1.9)	8.0 (1.8)	11.6 (1.26)	7.27 (0.92)
Completeness (%)	99.5 (98.8)	96.0 (97.9)	99.9 (99.9)	99.3 (96.9)
Redundancy	7.0 (6.8)	2.0 (1.9)	9.9 (9.7)	4.9 (4.8)
Refinement				
R_{work} / R_{free} (%)[c]	19.8/22.6		18.7/21.4	22.9/26.6
Average B-factors (Å^2)	46.1		42.9	71.7
Protein	45.1		42.5	72.1
Ligands	60.8		123.4	
Water	53.9		42.2	59.3
No. atoms				
Protein	2331		10492	7827
Ligands	10		60	
Water	256		1119	248
Rms deviations				
Bond lengths (Å)	0.003		0.005	0.004
Bond angles (°)	0.67		0.68	0.73
Ramachandran plot (%)[d]				
Total favored	99.65		99.24	98.26
Total allowed	0.35		0.68	1.74
PDB code	6XRB		6XRR	6XRF

[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).

Table 2.5. Strains used for the work presented in chapter II.

Organism	Genotype	Description	Reference
P. protegens Pf-5	wild-type		(Paulsen et al., 2005)
	ΔPFL_6095	*eagR1* deletion strain	This study
	ΔPFL_6099	*eagT2* deletion strain	This study
	ΔPFL_6209	*tne2* deletion strain	(Tang et al., 2018)
	ΔPFL_6096	*rhsA* deletion strain	(Tang et al., 2018)
	ΔPFL_6079	*pppA* deletion strain	This study
	ΔPFL_6094	*vgrG1* deletion strain	This study
	ΔPFL_6096 ΔPFL_6097 attB::*lacZ*, Tet^R	*rhsA rhsI* deletion strain, constitutive *lacZ* expression, Tet^R	This study
	ΔPFL_6079 ΔPFL_6096 ΔPFL_6097 attB::*lacZ*, Tet^R	*pppA rhsA rhsI* deletion strain, constitutive *lacZ* expression, Tet^R	This study
	ΔPFL_6079 ΔPFL_6209 ΔPFL_6210 attB::*lacZ*, Tet^R	*pppA tne2 tni2* deletion strain, constitutive lacZ expression, Tet^R	This study
	His_{10}-PFL_6096	Expresses RhsA with a N-terminal His_{10} tag	This study
	ΔPFL_6095 His_{10}-PFL_6096	*eagR1* deletion strain expressing His_{10}-RhsA	This study
	ΔPFL_6099 His_{10}-PFL_6096	*eagT2* deletion strain expressing His_{10}-RhsA	This study
	His_{10}-PFL_6209-VSV-G	Expresses Tne2 with a N-terminal His_{10} tag and a C-terminal VSV-G tag	This study
	ΔPFL_6095 His_{10}-PFL_6209-VSV-G	*eagR1* deletion strain expressing His_{10}-Tne2-VSV-G	This study
	ΔPFL_6099 His_{10}-PFL_6209-VSV-G	*eagT2* deletion strain expressing His_{10}-Tne2-VSV-G	This study
	ΔPFL_6095 His_{10}-PFL_6096_Δ2-74	*eagR1* deletion strain expressing His_{10}-RhsA lacking its N-terminal TM region	This study
	ΔPFL_6081	*tssM* deletion strain	(Tang et al., 2018)
	PFL_6096_Δ2-74	Expresses RhsA lacking its N-terminal TM region	This study
	FLAG-PFL_6094 His_{10}-PFL_6096	Expresses VgrG1 with a N-terminal FLAG tag and His_{10}-RhsA	This study
	FLAG-PFL_6094 His_{10}-PFL_6096_Δ2-74	Expresses VgrG1 with a N-terminal FLAG tag and His_{10}-$RhsA_{\Delta NT}$	This study

P. aeruginosa PAO1	Wild-type		(Stover et al., 2000)
	ΔPA4856	*retS* deletion strain	(Goodman et al., 2004)
	ΔPA4856 attB::*lacZ*	Constitutive *lacZ* expression strain, Tet^R	(Whitney et al., 2014)
	ΔPA4856 ΔPA0092 ΔPA0093 attB::*lacZ*	*tse6 tsi6* deletion strain constitutively expressing *lacZ*, Tet^R	(Whitney et al., 2014)
	ΔPA4856 ΔPA0090	*clpV1* deletion strain	(Whitney et al., 2014)
	ΔPA4856 ΔPA0091	*retS vgrG1a* deletion strain	(Whitney et al., 2014)
	ΔPA4856 ΔPA0093	*retS tse6* deletion strains	(Whitney et al., 2014)
	ΔPA4856 ΔPA0094	*retS eagT6* deletion strain	(Whitney et al., 2015)
	ΔPA4856 PA0094-V	Expresses EagT6 with a C-terminal VSV-G tag	This study
	ΔPA4856 ΔPA0090 PA0094-V	*clpV1* deletion strain expressing EagT6 with a C-terminal VSV-G tag	This study
	ΔPA4856 PA0094-V_L3Q	Expresses $EagT6^{L3Q}$ variant	This study
	ΔPA4856 PA0094-V_Q20A	Expresses $EagT6^{Q20A}$ variant	This study
	ΔPA4856 PA0094-V_I22Q	Expresses $EagT6^{I22Q}$ variant	This study
	ΔPA4856 PA0094-V_R34A	Expresses $EagT6^{R34A}$ variant	This study
	ΔPA4856 PA0094-V_V39Q	Expresses $EagT6^{V39Q}$ variant	This study
	ΔPA4856 PA0094-V_D48A	Expresses $EagT6^{D48A}$ variant	This study
	ΔPA4856 PA0094-V_K64A	Expresses $EagT6^{K64A}$ variant	This study
	ΔPA4856 PA0094-V_T115Q	Expresses $EagT6^{T115Q}$ variant	This study
	ΔPA4856 PA0094-V_H125A	Expresses $EagT6^{H125A}$ variant	This study
	ΔPA4856 PA0094-V_R140A	Expresses $EagT6^{R140A}$ variant	This study
	ΔPA4856 PA0093_Δ2-61	Expresses Tse6 variant lacking the N-terminal TMD1-containing region ($Tse6^{\Delta NT}$)	This study
	ΔPA4856 PA0093_Δ180-222	Expresses Tse6 variant lacking TMD2	This study
	ΔPA4856 PA0093_Δ2-61, Δ180-222	*retS* deletion strain expressing	This study

		Tse6$^{\Delta NT/TMD2}$ (Tse6$^{\Delta prePAAR,\ \Delta TMDs}$)	
	ΔPA4856 PA0093_Δ16-61_Δ180-222	*retS* deletion strain expressing Tse6$^{\Delta TMDs}$	This study
	ΔPA4856 ΔPA0094 PA0093_Δ2-61, Δ180-222	*eagT6* deletion strain expressing Tse6$^{\Delta NT/TMD2}$	This study
	ΔPA4856 ΔPA4427	*retS* and *sspB* deletion strain	(Whitney et al., 2015)
	ΔPA4856 ΔPA4427 *tsi6*-D4	*retS* and *sspB* deletion strain expressing Tsi6 with a C-terminal DAS+4 tag	(Whitney et al., 2015)
	ΔPA4856 ΔPA4427 PA0093_Δ2-61_Δ180-222	*sspB* deletion strain expressing Tse6$^{\Delta NT/TMD2}$	This study
	ΔPA4427 ΔPA4856 *tsi6*-D4 PA0093_Δ2-61_Δ180-222	*sspB* deletion strain expressing Tsi6 with a C-terminal DAS+4 tag and Tse6$^{\Delta NT/TMD2}$	This study
	ΔPA4856 PA0093_A6L	*retS* deletion strain expressing Tse6^{A6L}	This study
	ΔPA4856 PA0093_A7L	*retS* deletion strain expressing Tse6^{A7L}	This study
	ΔPA4856 PA0093_R8A	*retS* deletion strain expressing Tse6^{R8A}	This study
	ΔPA4856 PA0093_D11A	*retS* deletion strain expressing Tse6^{D11A}	This study
	ΔPA4856 PA0093_H15A	*retS* deletion strain expressing Tse6^{H15A}	This study
	ΔPA4856 PA0093_D11A_H15A	*retS* deletion strain expressing Tse6$^{D11A,\ H15A}$	This study
	ΔPA4856 ΔPA2684	*retS tse5* deletion strain	(Whitney et al., 2014)
	ΔPA4856 ΔPA0093	*retS tse6* deletion strain	(Whitney et al., 2014)
	ΔPA4856 ΔPA2684 ΔPA0093	*retS tse5 tse6* deletion strain	This study
	ΔPA4856 ΔPA2684 PA0093_Δ2-61_Δ180-222	*retS tse5* deletion strain expressing Tse6$^{\Delta NT/TMD2}$	This study
E. coli SM10 λpir	*thi thr leu tonA lac Y supE recA*::RP4-2-Tc::Mu	Conjugation strain	BioMedal LifeScience
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	Novagen
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 2.6. Plasmids used for the work presented in chapter II.

Plasmid	Relevant features	Reference
pETDuet-1	Co-expression vector with *lacI*, T7 promoter, N-terminal His_6 tag in MCS-1, Amp^R	Novagen
pRSETA	Expression vector with *lacI*, T7 promoter, N-terminal His_6 tag and a HRV 3C protease cleavage site, Amp^R	Life Technologies
pET29b	Expression vector with *lacI*, T7 promoter, C-terminal His_6 tag, Kan^R	(Rietsch et al., 2005)
pEXG2	Allelic replacement vector containing *sacB*, Gm^R	(Baynham et al., 2006)
pSW196	MiniCTX1 plasmid, Tet^R	(Mougous et al., 2006)
pSCrhaB2-CV	Expression vector with *PrhaB*, Tmp^R	(Cardona and Valvano, 2005)
pPSV39-CV	Expression vector with *lacI*, *lacUV5* promoter, C-terminal VSV-G tag, Gm^R	This study
pSW196::*lacZ*	*lacZ* in miniCTX1 plasmid	This study
pETDuet-1::His_6-ECL_01567-FLAG ::ECL_01568	Vector used to co-express N-terminal His_6 and C-terminal FLAG tagged RhsA and untagged RhsI from *E. cloacae*	This study
pETDuet-1::His_6-SF0266-FLAG	Vector used to express N-terminal His_6 tagged class I prePAAR effector SF0266 from *S. flexneri*	This study
pETDuet-1::His_6-SL1344_0286-FLAG :: SL1344_0286a	Vector used to co-express N-terminal His_6 and C-terminal FLAG tagged Rhs1 and untagged RhsI from *S.* Typhimurium	This study
pETDuet-1::His_6-PA0093-FLAG ::PA0092	Vector used to co-express N-terminal His_6 and C-terminal FLAG tagged Tse6 and untagged Tsi6 from *P. aeruginosa*	This study
pETDuet-1::His_6-Spro_3017_FLAG::Spro_3018	Vector used to co-express N-terminal His_6 and C-terminal FLAG tagged Tre1 and untagged Tri1 from *S. proteamaculans*	This study
pETDuet-1::His_6-PFL_6096::PFL_6097	Vector used to co-express N-terminal His_6-tagged RhsA and RhsI from *P. protegens*	This study
pETDuet-1::His_6-PFL_6096_Δ2-74::PFL_6097	Vector used to co-express N-terminal His_6-tagged $RhsA_{\Delta NT}$ and untagged RhsI from *P. protegens*	This study
pETDuet-1::PA0093_1-61-His_6::PA0094-VSV-G	Vector used to co-express C-terminal His_6 tagged $Tse6^{NT}$ and C-terminal VSV-G tagged EagT6	This study
pETDuet-1:: SL1344_0286_1-59-His_6:: SL1344_0285-VSV-G	Vector used to co-express C-terminal His_6 tagged $Rhs1^{NT}$ and C-terminal VSV-G tagged SciW	This study
pETDuet-1::His_6-PFL_6209::PFL_6210	Vector used to co-express N-terminal His_6-tagged Tne2 and Tni2 from *P. protegens*	This study
pETDuet-1::His_6-SL1344_0286_Δ1-59-FLAG :: SL1344_0286a	Vector used to co-express N-terminal His_6 and C-terminal FLAG tagged $Rhs1_{\Delta NT}$ and untagged RhsI from *S.* Typhimurium	This study

pETDuet-1::His_6-PA0093::PA0092	Vector used to co-express N-terminal His_6-tagged Tse6 and untagged Tsi6 from *P. aeruginosa*	This study
pETDuet-1::His_6-PA0093_D11A::PA0092	Vector used to co-express N-terminal His_6-tagged $Tse6^{D11A}$ and untagged Tsi6 from *P. aeruginosa*	
pETDuet-1::His_6-PA0093_H15A::PA0092	Vector used to co-express N-terminal His_6-tagged $Tse6^{H15A}$ and untagged Tsi6 from *P. aeruginosa*	This study
pETDuet-1::His_6-PA0093_D11A_H15A::PA0092	Vector used to co-express N-terminal His_6-tagged $Tse6^{D11A, H15A}$ and untagged Tsi6 from *P. aeruginosa*	This study
pETDuet-1::FLAG-PA0091	Vector used to express N-terminal FLAG tagged VgrG1 from *P. aeruginosa*	This study
pETDuet-1::PFL_6096_1-74-VSV-G::PFL_6095-His6	Expression vector for C-terminal VSV-G tagged $RhsA_{NT}$ and N-terminal His6-tagged EagR1	This study
pETDuet-1::PA0091	Vector used to express VgrG1a	(Whitney et al., 2015)
pETDuet-1::PA0093::PA0092	Vector used to co-express N-terminal His_6-tagged Tse6 and untagged Tsi6	(Whitney et al., 2015)
pETDuet-1::PA0093_ Δ2-61::PA0092	Vector used to co-express N-terminal His_6-tagged $Tse6^{\Delta NT}$ and untagged Tsi6	This study
pETDuet-1::PA0093_ Δ180-222::PA0092	Vector used to co-express N-terminal His_6-tagged $Tse6^{\Delta TMD2}$ and untagged Tsi6	This study
pETDuet-1::PA0093_ Δ2-61_Δ180-222::PA0093	Vector used to co-express N-terminal His_6-tagged $Tse6^{\Delta NT/TMD2}$ and untagged Tsi6	This study
pETDuet-1::PA0093-His_8::PA0093	Vector used to co-express C-terminal His_8 tagged Tse6 and untagged Tsi6	(Whitney et al., 2015)
pRSFDuet-1::PA0094	Vector used to express C-terminal VSV-G tagged EagT6	(Whitney et al., 2015)
pRSFDuet-1::PA0091::PA0094	Vector used to co-express untagged VgrG1a and C-terminal VSV-G tagged EagT6	(Whitney et al., 2015)
pET29b::PA0092	Vector used to express Tsi6	(Whitney et al., 2015)
pRSETA::SL1344_0285	Vector used to express SciW (for crystallization)	This study
pET29b::ECL_01566-VSV-G	Vector used to express C-terminal VSV-G tagged $EagR_A$ from *E. cloacae*	This study
pET29b::SF0260a-VSV-G	Vector used to express C-terminal VSV-G tagged SF0260a (Eag) from *S. flexneri*	This study
pET29b:: SL1344_0285-VSV-G	Vector used to express C-terminal VSV-G tagged SciW from *S.* Typhimurium	This study
pET29b::PA0094-VSV-G	Vector used to express C-terminal VSV-G tagged EagT6 from *P. aeruginosa*	This study
pET29b::Spro_3016-VSV-G	Vector used to express C-terminal VSV-G tagged EagT6 from *S. proteamaculans*	This study
pET29b::PFL_6095-VSV-G	Vector used to express C-terminal VSV-G tagged EagR1 from *P. protegens*	This study
pET29b::PFL_6099-VSV-G	Vector used to express C-terminal VSV-G tagged EagT2 from *P. protegens*	This study
pET29b::FLAG-PFL_6094	Vector used to express N-terminal FLAG tagged VgrG1 from *P. protegens*	This study

pET29b::PA0093_75-162-FLAG	Vector used to express C-terminal FLAG tagged PAAR domain of Tse6	This study
pEXG2::ΔPFL_6095	*eagR1* deletion construct	This study
pEXG2::ΔPFL_6099	*eagT2* deletion construct	This study
pEXG2::ΔPFL_6209	*tne2* deletion construct	This study
pEXG2::ΔPFL_6096	*rhsA* deletion construct	This study
pEXG2::ΔPFL_6079	*pppA* deletion construct	This study
pEXG2::ΔPFL_6096 ΔPFL_6097	*rhsA-rhsI* effector-immunity pair deletion construct	This study
pEXG2::ΔPFL_6209 ΔPFL_6210	*tne2-tni2* effector-immunity pair deletion construct	This study
pEXG2::ΔPFL_6094	*vgrG1* deletion construct	This study
pEXG2::His_{10}-PFL_6096	N-terminal His_{10}-*rhsA* fusion construct	This study
pEXG2::His_{10}-PFL_6096*	N-terminal His_{10}-*rhsA* fusion construct compatible with a strain lacking *eagR1*	This study
pEXG2::FLAG-PFL_6094	N-terminal FLAG-*vgrG*1 fusion construct	This study
pEXG2::His_{10}-PFL_6209	N-terminal His_{10}-*tne2* fusion construct	This study
pEXG2::PFL_6209-VSV-G	VSV-G	This study
pEXG2::PFL_6096_Δ2-74	$RhsA_{NT}$ deletion construct	This study
pEXG2::His_{10}-PFL_6096_Δ2-74	$RhsA_{NT}$ deletion construct compatible in a strain with an N-terminal His_{10}-*rhsA* fusion	This study
pEXG2::His_{10}-PFL_6096_Δ2-74*	$RhsA_{NT}$ deletion construct compatible in a strain with an N-terminal His_{10}-*rhsA* fusion and lacking *eagR1*	This study
pEXG2::ΔPA0090	*clpV1* deletion allele in pEXG2	(Mougous et al., 2006)
pEXG2::ΔPA0091	*vgrG1a* deletion allele in pEXG2	(Hood et al., 2010)
pEXG2::PA0092_DAS+4	For generating strains encoding Tsi6 with a C-terminal DAS+4 tag	(Whitney et al., 2015)
pEXG2::PA0093_Δ2-61	For generating strains encoding Tse6 lacking its N-terminal TMD1-containing region ($Tse6^{\Delta NT}$)	This study
pEXG2::PA0093_Δ180-222	For generating strains encoding Tse6 lacking TMD2	This study
pEXG2::PA0093_ Δ2-61_Δ180-222	For generating strains encoding Tse6 lacking its N-terminal TMD1 containing region and TMD2	This study
pEXG2::PA0093_Δ16-61, Δ180-222	For generating strains encoding Tse6 lacking TMD1 and TMD2	This study
pEXG2::ΔPA0094	*eagT6* deletion allele in pEXG2	(Whitney et al., 2015)
pEXG2::PA0094_L3Q	For generating strains encoding the EagT6 L3Q point mutant	This study
pEXG2::PA0094_Q20A	For generating strains encoding the EagT6 Q20A point mutant	This study
pEXG2::PA0094_I22Q	For generating strains encoding the EagT6 I22Q point mutant	This study
pEXG2::PA0094_R34A	For generating strains encoding the EagT6 R34A point mutant	This study
pEXG2::PA0094_V39A	For generating strains encoding the EagT6 V39A point mutant	This study

pEXG2::PA0094_D48A	For generating strains encoding the EagT6 D48A point mutant	This study
pEXG2::PA0094_K64A	For generating strains encoding the EagT6 K64A point mutant	This study
pEXG2::PA0094_T115Q	For generating strains encoding the EagT6 T115Q point mutant	This study
pEXG2::PA0094_H125A	For generating strains encoding the EagT6 H125A point mutant	This study
pEXG2::PA0094_R140A	For generating strains encoding the EagT6 R140A point mutant	This study
pEXG2::ΔPA2684	*tse5* deletion allele in pEXG2	This study
pEXG2::ΔPA2685	*vgrG4* deletion allele in pEXG2	This study
pEXG2::ΔPA4427	*sspB* deletion allele in pEXG2	(Silverman et al., 2013)
pEXG2::ΔPA4856	*retS* deletion allele in pEXG2	(Goodman et al., 2004)
pEXG2::PA0093_A6L	For generating strains encoding the $Tse6^{A6L}$ point mutant	This study
pEXG2::PA0093_A7L	For generating strains encoding the $Tse6^{A7L}$ point mutant	This study
pEXG2::PA0093_R8A	For generating strains encoding the $Tse6^{R8A}$ point mutant	This study
pEXG2::PA0093_D11A	For generating strains encoding the $Tse6^{D11A}$ point mutant	This study
pEXG2::PA0093_H15A	For generating strains encoding the $Tse6^{H15A}$ point mutant	This study
pEXG2::PA0093_D11A_H15A	For generating strains encoding the $Tse6^{D11A,H15A}$ point mutant	This study
pPSV35-CV::PA0094	Vector used to express EagT6	Whitney et al., 2015
pPSV39-CV::PFL_6095-VSV-G	Vector used to express C-terminal VSV-G tagged EagR1	This study
pPSV38-CV::PA4427	Vector used to express SspB	(Castang et al., 2008)
pSCrhaB2-CV::PA0093_D396A	Vector used to express $Tse6^{D396A}$	This study
pSCrhaB2-CV::PA0093_Δ2-61, Δ180-222_D396A	Vector used to express $Tse6^{\Delta NT/TMD2,\ D396A}$	This study
pSCrhaB2-CV::PA0093_282-430_D396A	Vector used to express $Tse6_{tox}^{D396A}$	(Whitney et al., 2015)
pSCrhaB2-V::PFL_6096_D1404A	Vector used to express $RhsA_{D1404A}$	This study
pSCrhaB2-V::PFL_6096_Δ2-74_D1404A	Vector used to express $RhsA^{\Delta NT}{}_{D1404A}$	This study

Methods

Bacterial strains and growth conditions

Pseudomonas strains used in this study were derived from *P. aeruginosa* PAO1 and *P. protegens* Pf-5 (Table 2.5). Both organisms were grown in LB medium (10 g L^{-1} NaCl, 10 g L^{-1} tryptone, and 5 g L^{-1} yeast extract) at 37°C (*P. aeruginosa*) or 30°C (*P. protegens*). Solid media contained 1.5% or 3% agar. Media were supplemented with gentamicin (30 µg mL^{-1}) and IPTG (250 µM) as needed.

Escherichia coli strains XL-1 Blue, SM10 and BL21 (DE3) Gold or CodonPlus were used for plasmid maintenance and toxicity experiments, conjugative transfer and protein overexpression, respectively (Table 2.5). All *E. coli* strains were grown at 37°C in LB medium. Where appropriate, media were supplemented with 150 µg mL^{-1} carbenicillin, 50 µg mL^{-1} kanamycin, 200 µg mL^{-1} trimethoprim, 15 µg mL^{-1} gentamicin, 0.25-1.0 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), 0.1% (w/v) rhamnose or 40 µg mL^{-1} X-gal.

Bioinformatics

Genomic analyses of effector sequences in UniProtKB

For the analysis of all effectors in UniprotKB we used six iterations of *jackhmmer* (HmmerWeb v2.41.1) using the first 60 amino acids of Tse6 (PA0093) protein to obtain 2,378 sequences. We removed any UniProtKB deprecated sequences entries (324/2378, remaining: 2,054) and further filter, cluster, and

analyze the remaining 975 effector sequences as stated below (same as analysis in Figure 2.2A). In our PAAR motif search, using our first to fourth PAAR motif HMMs (see analysis below), we identified 734/975, 200/241, 30/41, and 8/11 sequences to respectively have PAAR motifs. The remaining 3 sequences that did not have PAAR motifs were found to be associated with a gene encoding a PAAR domain-containing protein. There were 7 sequences that did not have any predicted TM. All scripts and intermediate files can be found in: https://github.com/karatsang/effector_chaperone_T6SS/tree/master/UniProtKB.

Genomic analyses of effector sequences in T6SS-containing genera

The genome assemblies of *Pseudomonas*, *Burkholderia*, *Enterobacter*, *Escherichia*, *Salmonella*, *Serratia*, *Shigella*, *Vibrio* and *Yersinia* were downloaded from NCBI using ncbi-genome-download (https://github.com/kblin/ncbi-genome-download, v0.2.10). Protein coding genes were predicted using Prodigal (v2.6.3) and the `-e 1` option (Hyatt et al., 2010). We developed a Hidden Markov Model (HMM) for detecting effectors by using the first 61 amino acids of Tse6 (PA0093) protein and six iterations of *jackhmmer* (HmmerWeb v2.41.1). *hmmsearch* (v3.1b2) and the effector HMM were used to identify the effectors in all genome assemblies using the ` -Z 45638612 -E 1000` options and we further filtered for a bitscore greater than 40. We further filtered to include effectors that included the prePAAR (AARxxDxxxH) motif, which we searched for using regular expressions, identifying 6,129 prePAAR-containing sequences across 5,584 genomes. For the analysis for

Figure 2.1C, each genome with at least one effector had to also encode for an Eag chaperone which we searched for using Pfam's established DcrB HMM (http://pfam.xfam.org/family/PF08786#tabview=tab6) and hmmsearch with the same parameter and bitscore cutoff as the effector search. For Figure 2.2A, to reduce spurious effector predictions, we removed sequences with less than 100 amino acids. To reduce redundancy, we removed any sequences that were 100% identical and clustered sequences with 95% sequence similarity that were less than 50 amino acids different in length using CD-HIT (v4.8.1 with ` -c 0.95 -n 5 -S 50`), leaving 1,166 sequences for further analysis (Li and Godzik, 2006). The numbers of sequences before and after filtering for the UniprotKB and sequences isolated from the 8 genera listed above are indicated in Table 2.3. We identified the presence of a PAAR domain through an iterative process of generating a PAAR motif HMMs and using *hmmsearch* (as described above) to capture the known diversity of the PAAR motif. We started broadly by using Pfam's PAAR motif HMM (http://pfam.xfam.org/family/PF05488#tabview=tab4) to identify 895/1166 PAAR motif containing sequences. For the 271 sequences that were predicted to not have a PAAR motif, we then generated an HMM using three iterations *jackhmmer* and the PAAR motif of the Tse6 (PA0093) protein (L75 to G162) to identify 219/271 PAAR motifs. We generated a third PAAR motif HMM using 60-160 amino acids of a randomly selected sequence (GCF_001214785.1 in contig NZ_CTBP01000066.1) and two iterations of *jackhmmer* that were not identified to have a PAAR motif in the previous search but was identified to have a PAAR

domain using phmmer (HmmerWeb version 2.41.1). We identified 42/52 sequences had a PAAR domain using the third PAAR motif. For the fourth PAAR domain HMM, we used the 60-160 amino acid sequence of GCF_005396085.1 in the NZ_BGGV01000116 contig and three iterations of *jackhammer* to identify 8/10 sequences that had a PAAR motif. The remaining two sequences with no PAAR domain were manually analyzed and were determined to either be directly associated with a PAAR domain downstream (GCF_001425105.1) or directly beside T6SS machinery gene (GCF_001034685.1). We predicted the transmembrane (TM) helices in proteins first using TMHMM (v2.0), Phobius web server, and TMbase (https://embnet.vital-it.ch/software/) (Käll et al., 2007; Krogh et al., 2001). Using TMHMM, we defined a TM region to include TM helices that were less than or equal to 25 amino acids apart. Therefore, any TM helix that was greater than 25 amino acids apart from another TM helix would be considered part of a new TM region. Any effector considered to have no TM or three TM regions were analyzed with Phobius with the same criteria as with TMHMM. Any effector considered to have no TM or three TM regions using Phobius, were analyzed with TMbase where we used the strongly preferred model and only interpreted TM helices with a score greater than 1450. In this model, any TM helices within the first 120 amino acids is one TM region and any number of TM helices between 200 and 300 amino acids were another region. MAFFT (v7.455) was used to align the sequences using the `--auto` option and the alignment was then trimmed to remove gaps using trimAl (v1.4) and the `-gt 0.8 -cons 80` options (Capella-Gutiérrez et

al., 2009; Katoh and Standley, 2013). We constructed the maximum-likelihood phylogenetic tree using FastTree (v2.1.10) and the `-gamma` option (Price et al., 2010). The phylogenetic tree was visualized using ggtree (Yu, 2020). We removed any effectors that did not have a chaperone and we categorized the chaperones with its corresponding effectors TM prediction. Sequence logos in Figures 2.1B and 2.2B were created by using logo maker (v0.8) (Tareen and Kinney, 2020). All scripts and intermediate files can be found in: https://github.com/karatsang/effector_chaperone_T6SS/tree/master/NCBI_8_Genera.

Screening for potential Zn^{2+}-binding residues

To collect orphan PAAR sequences, we used the Pfam database's information on the PAAR motif (PF05488, http://pfam.xfam.org/family/PF05488#tabview=tab1) and only obtained the 1,923 sequences with one PAAR motif architecture. We then aligned and trimmed the alignment of the 1,923 orphan PAAR sequences. We then used the previously mentioned 2,054 effector sequences from UniProtKB and filtered to only use 1765 sequences with an AARxxDxxxH motif. To identify Zn^{2+}-binding residues in orphan and prePAAR effector sequence logos, we used logo maker (v0.8) to create sequence logos for the first 200 amino acids (Tareen and Kinney, 2020). To explore the relationship between the orphan PAAR sequences and the PAAR domain sequences of the prePAAR effector, we generated two phylogenetic trees.

First, we truncated the prePAAR effector sequences to only the first 300 amino acids as an estimation of the PAAR domain location. Using the orphan PAAR sequences and the shortened prePAAR effector sequences, we created a phylogenetic tree (Figure 2.12B-C) using the methods described above. Next, we wanted to determine the relationship between orphan PAAR sequences and PAAR domain sequences in the prePAAR effectors. We built a hidden Markov model using *hmmbuild* (v 3.1b2) and eight PAAR domains annotated using Phyre2 (Kelley et al., 2015). We then searched for a PAAR domain using *hmmsearch* (v 3.1b2) in the first 300 amino acids of the prePAAR effector sequences. For each sequence, we identified an envelope boundary (amino acid coordinates) of where the PAAR domain is predicted to be located and we truncated the sequence at these positions. All scripts and intermediate files can be found in: https://github.com/karatsang/effector_chaperone_T6SS/tree/master/ZnBindingResidues

DNA manipulation and plasmid construction

Primers were synthesized and purified by Integrated DNA Technologies (IDT). Phusion polymerase, restriction enzymes and T4 DNA ligase were obtained from New England Biolabs (NEB). Sanger sequencing was performed by Genewiz Incorporated.

Plasmids used for heterologous expression were pETDuet-1, pET29b and pSCrhaB2-CV. Mutant constructs were made using splicing by overlap-extension

PCR and standard restriction enzyme-based cloning procedures were subsequently used to ligate PCR products into the plasmid of interest.

In-frame chromosomal deletion mutants in *P. aeruginosa* and *P. protegens* were made using the pEXG2 plasmid as described previously (Hmelo et al., 2015). Briefly, 500-600 bp upstream and downstream of target gene were amplified by standard PCR and spliced together by overlap-extension PCR. The resulting DNA fragment was ligated into the pEXG2 allelic exchange vector using standard cloning procedures (Table 2.6). Deletion constructs were transformed into *E. coli* SM10 and subsequently introduced into *P. aeruginosa* or *P. protegens* via conjugal transfer. Merodiploids were directly plated on LB (lacking NaCl) containing 5% (w/v) sucrose for *sacB*-based counter-selection. Deletions were confirmed by colony PCR in strains that were resistant to sucrose, but sensitive to gentamicin. Chromosomal point mutations or tags were constructed similarly with the constructs harboring the mutation or tag cloned into pEXG2. Sucrose-resistant and gentamicin-sensitive colonies were confirmed to have the mutations of interest by Sanger sequencing of appropriate PCR amplicons.

Bacterial toxicity experiments

This section pertains to the *E. coli* plates in Figure 2.5A and C. For Figure 2.5A, it was previously shown that a D1404A mutation was sufficient to attenuate, but not abolish, the toxicity of RhsA and allows for the cloning of this toxin in the absence of its immunity gene (Tang et al., 2018). Therefore, to assess the toxicity

of the full-length effector and a truncated variant, we cloned RhsA^{D1404A} or RhsA$_{\Delta NT}$^{D1404A} into the rhamnose-inducible pSCrhaB2-CV vector. For Figure 2.5C, plasmids containing Tse6$_{\Delta NT/TMD2}$ were not tolerated under non-inducing conditions. Therefore, we cloned a variant (D396A) that we previously demonstrated reduces, but does not abolish, the toxic NADase activity of Tse6 (Whitney et al., 2015). To allow for the pairwise comparison of toxicity levels between strains, this amino acid substitution was also introduced into the plasmids expressing Tse6 and Tse6$_{tox}$ (see Table 2.6). The pSCrhaB2-CV expressing RhsA and Tse6 variants were co-transformed into *E. coli* XL-1 Blue with IPTG-inducible pPSV39-CV containing EagR1 or pPSV35-CV containing EagT6, respectively (see Table 2.6). Stationary-phase overnight cultures containing these plasmids were serially diluted 10^{-6} in 10-fold increments and each dilution was spotted onto LB agar plates containing 0.1% (w/v) L-rhamnose, 250 µM IPTG, trimethoprim 250 µg mL^{-1} and 15 µg mL^{-1} gentamicin. Photographs were taken after overnight growth at 37°C.

Cell fraction preparation and secretion assays

Stationary-phase overnight cultures of *E. coli* (DE3) BL21 CodonPlus, *P. aeruginosa* Δ*retS* or *P. protegens* were inoculated into 2 mL or 50 mL LB at a ratio of 1:500, respectively. Cultures were grown at 37 °C (*E. coli* and *P. aerugionsa*) or 30 °C (*P. protegens*) to OD 0.6-0.8. Upon reaching the desired OD, all samples were centrifuged at 7, 600 x *g* for 3 min. The secreted fraction in *P. aeruginosa* or

P. protegens samples was prepared by isolating the supernatant and treating it with TCA (final conc: 10% (v/v)) as described previously (Whitney et al., 2015). The cell pellet was resuspended in 60 µL PBS, treated with 4X laemmli SDS loading dye and subjected to boiling to denature and lyse cells. For experiments examining the stability of Tse6-VgrG1a complexes, *P. aeruginosa* cells were resuspended in 60 µL PBS and subjected to six freeze-thaw cycles or brief sonication (2 second pulse, 5x, amplitude 30%) prior to mixing with 2X laemmli SDS loading dye. For preparation of *P. protegens* and *E. coli* cell fractions containing His-tagged complexes, cells were resuspended in lysis buffer containing 50 mM Tris-HCl (pH 8.0), 250 mM NaCl, 10 mM imidazole and purified according to the protocol described below (see *Protein expression and purification*).

Tsi6 depletion assays

A sequence encoding the C-terminal DAS + 4 degradation tag (DNA: 5'-GCCGCCAACGACGAGAACTACAGCGAGAACTACGCCGACGCCAGC-3'; protein: AANDENYSENYADAS, (McGinness et al., 2006)) was fused to the 3' end of the native *tsi6* locus in *P. aeruginosa* strains lacking the native *sspB* gene and expressing either wild-type Tse6 or $Tse6_{\Delta NT/TMD2}$. An IPTG-inducible plasmid containing *sspB* was used to stimulate controlled degradation of Tsi6-DAS + 4 (Tsi6-D4). The SspB protein recognizes DAS+4 tagged proteins and delivers them to the ClpXP protease for degradation. Strains harbouring this plasmid were streaked on LB agar supplemented with 500 µM IPTG.

Competition assays

A tetracycline-resistant, *lacZ*-expression cassette was inserted into a neutral phage attachment site (*attB*) of recipient *P. aeruginosa* and *P. protegens* strains to differentiate these strains from unlabeled donors. *P. protegens* recipient strains also contain a Δ*pppA* mutation to stimulate T6SS effector secretion to induce a T6SS 'counterattack' from *P. protegens* donor strains (Basler et al., 2013).

For intraspecific competitions between *P. aeruginosa* or *P. protegens* donors against isogenic recipient that lack the indicated effector-immunity pairs, stationary-phase overnight cultures were mixed in a 1:1 (v/v) ratio.

Initial ratios of donors:recipients were counted by plating part of the competition mixtures on LB agar containing 40 μg mL^{-1} X-gal. The remainder of each competition mixture was spotted (10 μL per spot) in triplicate on a 0.45 μm nitrocellulose membrane overlaid on a 3% LB agar plate and incubated face up at 37 °C for 20-24 h. Competitions were then harvested by resuspending cells in LB and counting colony forming units by plating on LB agar containing 40 μg mL^{-1} X-gal. The final ratio of donor:recipient colony forming units were normalized to the initial ratios of donor and recipient strains.

Protein expression and purification

All plasmids used for co-purification experiments (chaperone-effector pairs shown in Figure 2.2, tagged variants of *P. protegens* and tagged variants of *P.*

aeruginosa proteins), the Tse6-VgrG1-EagT6$_2$-EF-Tu-Tsi6 complex used for cryo-EM, the RhsA-RhsI-EagR1-VgrG complex for negative-stain EM, Hcp (PFL_6089) and RhsA$_{\Delta NT}$ used for antibody development or the SciW, EagT6-Tse6$_{NT}$ complex and the SciW-Rhs1$_{NT}$ complex used for crystallization were expressed in *E. coli* BL21 (DE3) CodonPlus or *E. coli* BL21 (DE3) Gold cells. Differences in the expression strategy used are indicated below. Additional details about protein tags, selection markers and plasmids are indicated in Table 2.6.

Co-purification experiments, *preparation of cryo-EM and negative stain EM samples and preparation of samples for antibody development*

Chaperone-effector pairs (e, effector; c, chaperone) from: *Pseudomonas aeruginosa* (e: PA0093, c: PA0094), *Salmonella* Typhimurium (e: SL1344_0286, c: SL1344_0285), *Shigella flexneri* (e: SF0266, c: SF3490), *Enterobacter cloacae* (e: ECL_01567, c: ECL_01566) and *Serratia proteamaculans* (e: Spro_3017, c: Spro_3016) were co-expressed using pET29b containing the predicted chaperone and pETDuet-1 harboring the full-length effector and its predicted immunity determinant. The effector and immunity in the RhsA-RhsI-EagR1-VgrG1 complex and the Tse6-VgrG1-EagT6$_2$-EF-Tu-Tsi6 were co-expressed in pETDuet-1 (see Table 2.6 for further details). The EagR1 chaperone for the RhsA-RhsI-EagR1-VgrG1 complex was cloned into a pET29b vector and co-expressed with the effector-immunity pair, whereas the VgrG1 protein was expressed using a pET29b vector. Both EagT6 and VgrG1a for the Tse6-VgrG1-EagT6$_2$-EF-Tu-Tsi6 complex

were co-expressed in pRSFDuet-1 and co-transformed with the pETDuet-1 vector containing Tse6 and Tsi6. To test EagT6 and VgrG binding, Tse6 and the Tse6 variants (ΔNT, ΔTMD2, ΔNT/TMD2, ΔTMDs, prePAAR mutants) were co-expressed with Tsi6 in pETDuet-1. The EagT6 chaperone was cloned into either pPSV35-CV or pET29b and co-expressed with Tse6 variants. VgrG1a was expressed in isolation using pETDuet-1. Hcp (PFL_6089) was expressed in pET29b. For *P. protegens*, all purified proteins were expressed from their native loci.

For the expression of chaperone-effector pairs in Figure 2.2D and the Tse6 ΔNT, ΔTMD2, ΔNT/TMD2, ΔTMDs and prePAAR mutants, a 1 mL overnight culture of expression strains was diluted in 50 mL or 100 mL of LB broth and grown at 37°C (*E. coli*) until OD 0.6-0.8. 40 mL overnight cultures were grown for all other of expression strains and were diluted into 2 L of LB broth and grown to OD_{600} 0.6-0.8 in a shaking incubator at 37°C. For most samples, protein expression was induced by the addition of 1 mM IPTG and cells were further incubated for 4.5 h at 37°C. Expression of other protein complexes such as the chaperone-effector pairs from *Salmonella* and *Enterobacter,* $RhsA_{\Delta NT}$-RhsI and RhsA-RhsI-EagR1-VgrG1 complexes were induced using 1 mM IPTG at 18 °C and incubated at this temperature overnight. One millilitre overnight cultures of *P. protegens* strains expressing the desired tagged protein was diluted in 50 mL of LB broth and grown at 30°C (*P. protegens*) until OD 0.8. Both *E. coli* and *P. protegens* cells were harvested by centrifugation at 9,800 *g* for 10 min following incubation. For the

RhsA-EagR1-VgrG1 complex and the experiments containing Tse6 N-terminal truncations or prePAAR mutants, the cultures expressing the cognate VgrG were co-pelleted with cells expressing the effector-immunity-chaperone. VgrG-expressing cultures were distributed equally among all samples prior to co-pelleting and ranged from 50%-100% of the total OD/mL of the effector-immunity-chaperone-expressing culture used for the pellet.

Final cell pellets from 50 mL to 100 mL cultures were resuspended in 3.5 mL lysis buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM imidazole), whereas those from 2 L of culture were resuspended in 25 mL of lysis buffer prior to rupture by sonication (6 x 30 second pulses, amplitude 30%). Cell lysates were cleared by centrifugation at 39,000 *g* for 60 min and the soluble fraction was loaded onto a gravity flow Ni-NTA column that had been equilibrated in lysis buffer. Ni-NTA-bound complexes were washed twice with 25 mL of lysis buffer followed by elution in 10 mL of lysis buffer containing 400 mM imidazole. The Ni-NTA purified complexes was further purified by gel filtration using a HiLoad 16/600 Superdex 200 column equilibrated in 20 mM Tris-HCl pH 8.0 150 mM NaCl or phosphate buffered saline (for samples used for antibody development only). Samples were run on SDS-PAGE gels followed by TGX or Coomassie Brilliant Blue R250 staining.

The purification of the Tse6-VgrG1-EagT6$_2$-EF-Tu-Tsi6 complex used for cyro-EM and liposome experiments was conducted similarly, with a few key changes. *E. coli* expressing Tse6, Tsi6, EagT6 and VgrG1a (pETDuet-1 and

pRSFDuet-1 system described above) were grown at 37 °C to mid-log phase and protein expression was induced by the addition of 1 mM IPTG. Following incubation at 30 °C for an additional 5–6 h, cells were pelleted and lysed in either detergent-free lysis buffer (same as above) or buffer containing 0.2% β-D-decylmaltopyranoside (depending on the downstream application). The VgrG1-Tse6-EagT62-EF-Tu-Tsi6 complexes were then further purified by Ni-NTA chromatography and size exclusion chromatography as described above. For the detergent-bound complex, a second purification step using a Superose 6 Increase 5/150 column was performed.

Preparation of samples for crystallization

sciW (SL1344_0285) was synthesized with codon optimization for *E. coli* and cloned into the vector pRSETA with the restriction sites NdeI/HindIII (Life Technologies). This construct includes an N-terminal His_6 tag and an HRV 3C protease cleavage site (MGSSHHHHHHSSDLEVLFQGPLS). SciW-$Rhs1_{NT}$ and EagT6-$Tse6_{NT}$ complexes were co-expressed using pETDUET-1. Note that the EagT6 construct has a C-terminal VSV-G tag (see Table 2.6). Cells were grown in LB broth to OD_{600} 0.6 at 37°C at which point protein expression was induced by the addition of 1mM IPTG. The temperature was reduced to 20°C and cultures were allowed to grow overnight. Cells were harvested by centrifugation and resuspended in lysis buffer followed by lysis with an Emulsiflex-C3 (Avestin). The lysate was cleared by centrifugation at 16,000 rpm for 30 minutes and the

supernatant passed over a nickel NTA gravity column (Goldbio) followed by washing with 50 column volumes of chilled lysis including PMSF, DNase I, and $MgCl_2$. Proteins were eluted with 5 column volumes elution buffer then purified by gel filtration using an SD75 16/60 Superdex gel filtration column equilibrated in gel-filtration buffer (GF) with an AKTA pure (GE Healthcare). For SciW, after affinity purification the protein was dialyzed in GF buffer O/N at 4°C and the His-tag removed during dialysis using HRV 3C protease. The digested SciW was passed over a nickel NTA gravity column and the flow through was collected. SciW was further purified using an SD75 16/60 Superdex gel filtration column equilibrated in GF buffer.

The buffers used were as follows: SciW lysis buffer (20mM Tris pH 7.5, 500mM NaCl, 20mM imidazole); SciW elution buffer (20mM Tris pH 7.5, 500mM NaCl, 500mM imidazole); SciW GF buffer (20 mM Tris pH 7.5, 250mM NaCl, 1mM 2-Mercaptoethanol); SciW-$Rhs1_{NT}$ and EagT6-$Tse6_{NT}$ complexes lysis buffer (20 mM Tris pH 8.0, 150 mM, 25 mM imidazole); elution buffer (20 mM Tris pH 8.0, 150 mM, 500 mM imidazole); and GF buffer (20 mM Tris pH 8.0, 150 mM NaCl, 1mM 2-Mercaptoethanol).

Crystallization and structure determination

SciW was concentrated to 7, 14 and 22 mg mL^{-1} for initial screening using commercially available screens (Qiagen) by sitting-drop vapor diffusion using a Crystal Gryphon robot (Art Robbins Instruments). The crystallization conditions for

SciW were 22 mg mL^{-1} with a 1:1 mixture of 0.1 M Tris HCL pH 8.5, 25% (v/v) PEG 550 MME at 4°C. EagT6-$Tse6_{NT}$ complex was concentrated to 5, 10 and 20 mg mL^{-1} and screened for crystallization conditions as per SciW. The final crystallization conditions were 20 mg mL^{-1} with a 1:1 mixture of 0.2M Magnesium chloride, 0.1M Bis-Tris pH 5.5, and 25% (w/v) PEG 3350 at 4°C. SciW-$Rhs1_{NT}$ complex was concentrated to 15, 20 and 25mg mL^{-1} and screened for crystallization as per SciW. The crystallization conditions were 25 mg mL^{-1} protein with a 1:1 mixture of 0.2M Ammonium sulfate, 0.1M Bis-Tris pH 5.5, and 25% (w/v) PEG 3350 at 4°C.

Diffraction data from crystals of SciW and EagT6-$Tse6_{NT}$ complex were collected in-house at 93K using a MicroMax-007 HF X-ray source and R-axis 4++ detector (Rigaku). Diffraction data from SciW-$Rhs1_{NT}$ crystals were collected at the Canadian Light Source at the Canadian Macromolecular Crystallography Facility Beam line CMCF-ID (08ID-1). SciW crystals were prepared by cryo-protection in mother liquor plus 38% PEG 550 MME and flash freezing in liquid nitrogen. Crystals of EagT6-$Tse6_{NT}$ and SciW-$Rhs1_{NT}$ complexes were prepared in the same manner with increasing the concentration of PEG3350 to 35-38%. All diffraction data were processed using XDS (Kabsch, 2010). Phases for SciW were determined by the molecular replacement-single anomalous diffraction (MR-SAD) technique. A home-source data set was collected from SciW crystals soaked in cryo-protectant containing 350 mM NaI for one-minute before flash freezing. EagT6 (PDB: 1TU1) was used as a search model and phases were improved by SAD

using the Phenix package (Adams et al., 2010). Phases for both the EagT6-Tse6$_{NT}$ and SciW-Rhs1$_{NT}$ complexes were obtained by molecular replacement using EagT6 (PDB: 1TU1) and SciW as search models, respectively, with the Phenix package. Initial models were built and refined using Coot, Refmac and the CCP4 suite of programs, Phenix, and TLS refinement (Emsley et al., 2010; Murshudov et al., 1997; Winn et al., 2011; Winn et al., 2001). Data statistics and PDB codes are listed in Table 2.4. The ligands identified included polyethylene glycol (PEG) for SciW and sulfate ions for the SciW-Rhs$_{NT}$ complex. Additionally, the side-chain of residue C11 in SciW chain B was observed to be covalently bound to 2-Mercaptoethanol to form CME (S,S-(2-hydroxyethyl)thiocysteine). The coordinates and structure factors have been deposited in the Protein data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NY (www.rcsb.org). Molecular graphics and analysis were performed using Pymol (Schrödinger, LLC) and UCSF Chimera (Pettersen et al., 2004).

Electron microscopy and image analysis

Sample vitrification

Three microlitres VgrG1-Tse6-EagT6$_2$-EF-Tu-Tsi6 complex, at a concentration of 0.015 mg ml^{-1}, was applied onto freshly glow-discharged Quantifoil 2/1 (with 2 nm additional carbon layer) cryo-EM grids, automatically blotted and plunged into liquid ethane using a CryoPlunge3 (Gatan) at a humidity between 90 and 100%. To improve ice quality and distribution, 0.01% Tween-20

was added during this step. Grid quality was assessed before data collection using a JEOL JEM-1400 or JEOL JEM-3200FSC equipped with a FEG and an in-column energy filter, operating at 300 kV. Grids were kept in liquid nitrogen for long-term storage.

Cryo-EM and image processing

Two cryo-EM datasets for the PFC complex were collected on a Titan Krios electron microscope (FEI) equipped with a Cs-Corrector, operated at 300 kV acceleration. Micrographs were recorded on a Falcon II direct electron detector (FEI) at a ×122,807 magnification (×59,000 nominal magnification), corresponding to a pixel size of 1.14 Å. Twenty-four frames taken at intervals of 62.5 ms were collected during each exposure, resulting in a total exposure time of 1.5 s and a total electron dose of 60 $e^-/Å^2$. Using the automated data collection software EPU (FEI), two datasets, with a defocus range of 1.7–4.2 µm, were automatically collected featuring 5,822 and 5,820 micrographs, respectively. The 24 frames were aligned and summed with the help of MotionCor2 (3 × 3 patches) (Tang et al., 2007; Zheng et al., 2017). Furthermore, dose-weighted and unweighted full-dose images were calculated. Data processing was performed using the software package SPHIRE/EMAN2 (Moriya et al., 2017). Unweighted full-dose images were used for defocus and astigmatism determination with CTER (SPHIRE). After merging both datasets and visually inspecting the integrated images, 8,744 micrographs were selected for subsequent processing. A combination of manual and automated

particle selection, as well as several rounds of 2D classification, yielded 137,906 'clean' dose-weighted and drift-corrected particles that were extracted with a final window size of 360 × 360 pixels (see Supplementary Fig. 2a,c at: https://static-content.springer.com/esm/art%3A10.1038%2Fs41564-018-0238-z/MediaObjects/41564_2018_238_MOESM1_ESM.pdf). Two-dimensional classification was performed using the iterative stable alignment and clustering (ISAC) algorithm implemented in SPHIRE. The particle stack was subjected to sxmeridien (3D refinement in SPHIRE) with imposed C3 symmetry, resulting in a 3.6 Å map of the C3 symmetric VgrG1 (top) part of the complex, estimated by the 'gold standard' criterion of Fourier shell correlation (FSC) = 0.143. The determined 3D projection parameters for each particle were subsequently used to create a symmetrized particle stack. This new stack contained three copies of each original particle with projection parameters rotated by 120° along the (C3-) symmetry axis. The ensuing 3D classification resulted in three volumes that were rotated 120° to each other (see Supplementary Fig. 4 at https://static-content.springer.com/esm/art%3A10.1038%2Fs41564-018-0238-z/MediaObjects/41564_2018_238_MOESM1_ESM.pdf). As anticipated, the three copies of the original particle were evenly distributed to these classes. However, given that classification procedures are not perfect in reality, we further confirmed that no more than one copy of the original particle is present in each class. Finally, one of the classes, containing 55,000 particles, was selected and subjected to a new local 3D refinement without imposing symmetry. This resulted in the 4.2 Å

density map of the PFC, where the resolution of the EM density decreases towards the periphery of the map (see Supplementary Fig. 1b at https://static-content.springer.com/esm/art%3A10.1038%2Fs41564-018-0238-z/MediaObjects/41564_2018_238_MOESM1_ESM.pdf). Global resolutions were calculated between two independently refined half maps at the 0.143 FSC criterion, local resolution was calculated using sxlocres of SPHIRE. Final densities were filtered to an estimated average resolution. To visualize local resolution gradients within the map, it was coloured according to the local resolution in Chimera (Pettersen et al., 2004). Graphical rendering of 3D average and variability was calculated using 'sx3dvariability' of the SPHIRE software package and filtered for illustrative purposes. The final electron density map allowed for the placement of crystal structures of EagT6 (PDB:1TU1) and the homology model of the $Tse6_{PAAR}$ domain (Phyre2 web server) (Kelley et al., 2015) using the 'Rigid-body Fit-in-Map' tool of Chimera. Rosetta was used to perform a relaxation of the known crystal structure of VgrG1 (PDB:4MTK) into the obtained cryo-EM density map (Wang et al., 2016).

Angular distribution plots for all structures, as well as beautified 2D class averages, were calculated using SPHIRE (see Supplementary Figs 2c and 3 at https://static-content.springer.com/esm/art%3A10.1038%2Fs41564-018-0238-z/MediaObjects/41564_2018_238_MOESM1_ESM.pdf).

Negative stain sample preparation

Four microlitres of each protein sample at a concentration of approx. 0.01 mg mL^{-1} was applied onto glow-discharged carbon-coated copper grids. After 45 s of incubation at room temperature, excess liquid was blotted away using Whatman No. 4 filter paper, followed by two washing steps with GF buffer. Samples were then stained with 1 % (w/v) uranyl formate solution and grids stored at RT until usage.

Data collection and image analysis

Images were recorded manually with a JEOL JEM-1400 microscope, equipped with a LaB_6 cathode and 4k x 4k CMOS detector F416 (TVIPS), operating at 120 kV. For VgrG1, $RhsA_{\Delta NT}$, the EagR1-RhsA complex and EagR1-RhsA-VgrG1 complex, a total of 99, 140, 100 and 120 micrographs, respectively, were collected with a pixel size of 2.26 Å. Particles for the VgrG1, $RhsA_{\Delta NT}$, EagR1-RhsA complex and EagR1-RhsA-VgrG1 complex were selected automatically with crYOLO using individually pre-trained models, resulting in 18676, 23907, 32078 and 31409 particles, respectively (Wagner et al., 2019). Subsequent image processing was performed with the SPHIRE software package (Moriya et al., 2017). Particles were then windowed to a final box size of 240 x 240 pixel. Reference-free 2-D classification was calculated using the iterative stable alignment and clustering algorithm (ISAC) implemented in SPHIRE, resulting in 2-D class averages of each respective complex (Yang et al., 2012). Distance

measurement were performed with the e2display functionality in EMAN2 (Tang et al., 2007). The placement of the crystal structure into the electron density map (EMD-0135) was done using rigid-body fitting in Chimera (Pettersen et al., 2004). Here, Tse6-TMD and EagT6 of the EagT6-TMD crystal structure were fitted independently as rigid bodies to better describe the density. Due to the distinct shape of the PAAR domain, three different orientations were possible in the docking step, each rotated by 120°. Docking of Tse6-TMD into the density embraced by the second EagT6 described this density less well.

Complex reconstitution into nanodiscs

The Tse6-VgrG1-EagT6$_2$-EF-Tu-Tsi6 solubilized in detergent (β-D-decylmaltopyranoside) was mixed with preformed nanodiscs (Cube Biotech), containing MSP1D1-ΔH5 and 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC), at a molar ratio of 1:4 and dialysed against ND-buffer (20 mM Tris-HCl, pH 8.0; 300 mM NaCl) for 96 h at room temperature. The ND-buffer was exchanged with fresh buffer after 24 h. Subsequently, size-exclusion chromatography with a Superose 6 Increase 10/300 GL column (GE Healthcare) was used to separate aggregates and/or empty nanodiscs, as well as non-reconstituted complexes.

Western blot analyses

Western blot analyses of protein samples were performed as described previously for rabbit anti-Tse6 (diluted 1:5000, Genscript), anti-Tse1 (diluted

1:5,000; Genscript), rabbit anti-FLAG (diluted 1:5,000; Sigma), rabbit anti-VSV-G (diluted 1:5,000; Sigma), rabbit anti-Hcp1 (*P. protegens*) (diluted 1:5,000, Genscript) and mouse anti-His_6 (diluted 1:5000, THE™ His Tag Antibody, Genscript) detected with anti-rabbit or anti-mouse horseradish peroxidase-conjugated secondary antibodies (diluted 1:5,000; Sigma) (Hood et al., 2010). Western blots were developed using chemiluminescent substrate (Clarity Max, Bio-Rad) and imaged with the ChemiDoc Imaging System (Bio-Rad).

Data Availability

All data supporting the findings of this work are available within this thesis and the associated appendix. Additional data and information can be found using the links below. The cryo-EM density maps of the Tse6-VgrG1-EagT6_2-EF-Tu-Tsi6 complex are deposited into the Electron Microscopy Data Bank with the accession numbers EMD-0135. X-ray crystallographic coordinates and structure factor files are available from the PDB: SciW (PDB 6XRB), SciW-Rhs1_{NT} (PDB 6XRR), EagT6-Tse6_{NT} (PDB 6XRF). Relevant data and details of plasmids and strains are available upon request from Dr. John Whitney.

Links to publisher websites for additional information:

https://www.nature.com/articles/s41564-018-0238-z#Sec27

https://elifesciences.org/articles/62816#s4

CHAPTER III – An interbacterial toxin inhibits target cell growth by synthesizing (p)ppApp

Preface

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

Ahmad, S.*, Wang, B*., Walker, M.D., Tran, H.R., Stogios, P.J., Savchenko, A., Grant, R.A., McArthur, A.G., Laub, M.T., and Whitney, J.C. (2019). An interbacterial toxin inhibits target cell growth by synthesizing (p)ppApp. Nature *575*, 674-678.

*indicates equal contribution

Author contributions: S.A., B.W., M.T.L. and J.C.W. conceived experiments and designed research. S.A. and B.W. performed cloning, bacterial competition assays, protein purification, biochemical experiments and protein crystallization. P.J.S. and R.A.G. conducted X-ray data collection and analyses. H.-K.R.T. and A.G.M. performed bioinformatics analyses for Figure 3.1B. M.D.W provided assistance with cloning, purification and crystallization of Tas1–Tis1 complex. S.A., B.W., M.T.L. and J.C.W. conducted figure design, manuscript writing and editing. M.T.L. and J.C.W. supervised the project. A.S., R.A.G., A.G.M., M.T.L. and J.C.W. provided funding.

Steinchen, W., **Ahmad, S.**, Valentini, M., Eilers, K., Majkini, M., Altegoer, F., Lechner, M., Filloux, A., Whitney, J.C., and Bange, G. (2021). Dual role of a (p)ppGpp- and (p)ppApp-degrading enzyme in biofilm formation and interbacterial antagonism. Mol Microbiol.

Author contributions: W.S., S.A., A.F., J.C.W., and G.B. designed research. W.S., S.A., M.V., K.E., and M.M. performed research. W.S., S.A., M.V., K.E., F.A., and M.L. analyzed data. W.S., S.A., A.F., J.C.W., and G.B. prepared the manuscript with input from the other authors.

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

Abstract

Bacteria have evolved sophisticated mechanisms to inhibit the growth of competitors (Granato et al., 2019). One such mechanism involves type VI secretion systems, which bacteria can use to directly inject antibacterial toxins into neighboring cells. Many of these toxins target cell envelope integrity, but the full range of growth inhibitory mechanisms remains to be determined (Russell et al., 2014a). Here, we discover a novel type VI secretion effector, Tas1, in the opportunistic pathogen *Pseudomonas aeruginosa*. A crystal structure of Tas1 reveals similarity to enzymes that synthesize (p)ppGpp, a broadly conserved signaling molecule in bacteria that modulates cell growth rate, particularly in response to nutritional stress(Hauryliuk et al., 2015). Strikingly, however, we find that Tas1 does not synthesize (p)ppGpp, and instead pyrophosphorylates adenosine nucleotides to produce (p)ppApp at rates of nearly 180,000 per min. Consequently, delivery of Tas1 into competitor cells drives the rapid accumulation of (p)ppApp, depletion of ATP, and widespread dysregulation of essential metabolic pathways, resulting in target cell death. Collectively, our findings reveal a new mechanism for interbacterial antagonism and demonstrate, for the first time, a physiological role for the metabolite (p)ppApp in bacteria.

Main

Identification of a novel type VI secretion system effector

Effectors that transit the bacterial type VI secretion system (T6SS) are often encoded adjacent to structural components of the secretion apparatus (Wexler et

al., 2016). In *Pseudomonas aeruginosa* strain PAO1, the Tse6 effector is found next to the bacteria-targeting haemolysin-coregulated protein secretion island I T6SS (H1-T6SS) (Whitney et al., 2014). We noted that in the more virulent, clinical isolate PA14, a unique domain is encoded by PA14_01140 instead of the well-characterized C-terminal NAD^+ glycohydrolase toxin domain of Tse6 (Figure 3.1A) (Quentin et al., 2018; Whitney et al., 2015). Orthologues of PA14_01140 are found in many PA14-related strains of *P. aeruginosa* as well as several other species of Proteobacteria (Figure 3.1B-C and Appendix Dataset A3.1). An additional open reading frame, PA14_01130, immediately downstream of PA14_01140 may encode a cognate immunity protein as T6SS effector-immunity genes are typically found adjacent to one another.

We hypothesized that the unique toxin encoded by PA14_01140 could contribute to the fitness of PA14 when co-cultured with PAO1 under contact-promoting conditions that facilitate T6SS attack. Indeed, we found that a PA14 strain lacking PA14_01140 displayed an approximately 40-fold decrease in competitive index against PAO1 (Figure 3.1D, F). Conversely, a variant of PAO1 lacking *tse6* exhibited a 7-fold decrease in co-culture fitness versus PA14 (Figure 3.1E). Though PA14 possesses a homologue of the Tse6-specific immunity determinant *tsi6*, this gene was not protective against Tse6 (Figure 3.1G).

To test the proposed immunity function of PA14_01130, we deleted the PA14_01140-PA14_01130-*tsi6* gene cluster from a PA14 recipient and found that it was outcompeted by its parental donor strain in a PA14_01140-dependent

manner (Figure 3.2A-B). The fitness defect of this recipient was restored by expressing PA14_01130 but not *tsi6* (Figures 3.2A and 3.3B). We also confirmed that PA14_01140, like Tse6, is secreted from cells in a H1-T6SS-dependent manner, as it requires the structural T6SS VgrG1 spike protein and the ClpV1 ATPase for secretion (Figure 3.2C-D). Together, our data demonstrate that PA14_01140 is a unique H1-T6SS effector with PA14_01130 functioning as its cognate immunity protein.

We next sought to determine how PA14_01140 inhibits the growth of bacterial cells. Conventional homology searches were inconclusive, although more sensitive hidden Markov model-based algorithms indicated weak similarity of its C-terminal domain to proteins harboring RelA-SpoT Homolog (RSH) domains (Figure 3.3A) (Johnson et al., 2010). These domains are highly conserved across bacteria and usually synthesize the bacterial alarmones guanosine penta- and tetraphosphate, (p)ppGpp, by transferring pyrophosphate from ATP to either GDP or GTP (Atkinson et al., 2011). Intracellular levels of (p)ppGpp tune growth rate in response to nutritional conditions (Potrykus et al., 2011). We found that expressing the C-terminal domain of PA14_01140 ($PA14\_01140_{tox}$) inhibited the growth of *E. coli*, even at levels of approximately three copies per cell, indicating that this domain is sufficient for toxicity (Figure 3.3B-D).

To determine if $PA14\_01140_{tox}$ is an RSH enzyme, we determined its structure in complex with the PA14_01130 immunity protein to a resolution of 2.2 Å (Figure 3.4A, Table 3.1). This structure revealed strong similarity to the

(p)ppGpp-synthetase domains of RelQ from *Bacillus subtilis* and RelP from *Staphyloccoccus aureus* (Figure 3.4B-D). Structural overlay of PA14_01140$_{tox}$ with RelQ revealed a highly conserved three-dimensional positioning of residues known to interact with the pyrophosphate donor ATP (Figure 3.4C). Mutating any of these residues drastically reduced toxicity when expressed in *E. coli* (Figure 3.4E). In contrast to the ATP binding site, the predicted guanosine nucleotide binding site of PA14_01140$_{tox}$ is substantially distorted relative to the catalytically competent position in the Rel enzymes. In our co-crystal structure, two α-helices in PA14_01140$_{tox}$ predicted to form this acceptor site are rotated by approximately 30° relative to the equivalent helices in the Rel proteins (Figure 3.4D). This rotation likely arises from binding of the immunity protein, PA14_01130, which may neutralize PA14_01140-mediated toxicity by inducing a structural rearrangement in the acceptor nucleotide binding site.

A toxin that synthesizes (p)ppApp

To assess the enzymatic activity of PA14_01140$_{tox}$, we used an assay that couples AMP production to the depletion of NADH, which can be monitored at 340nm (Wang et al., 2019). Incubating purified PA14_01140$_{tox}$ with ATP and GTP led to a dose-dependent decrease in A_{340} over time indicating the production of AMP (Figure 3.5A). Surprisingly, however, AMP production by PA14_01140$_{tox}$ did not require GTP (Figure 3.5B). This finding suggested that PA14_01140$_{tox}$ can transfer a pyrophosphate from ATP to an adenosine nucleotide acceptor. To test

this hypothesis, we incubated purified PA14_01140$_{tox}$ with ATP alone, ATP+ADP, or ATP+AMP and used anion-exchange chromatography to examine the products. In these reactions, PA14_01140$_{tox}$ produced pppApp, ppApp and pApp, respectively, with the identities of these molecules verified by mass spectrometry and by ^{1}H and ^{31}P NMR (Figure 3.5C-E, Table 3.2). In the presence of ATP alone, we also observed pApp formation, which suggests that the pppApp initially produced can subsequently be used to pyrophosphorylate AMP, producing two pApp molecules (Figure 3.6D-E). Collectively, these results demonstrated that PA14_01140$_{tox}$ is a pyrophosphate kinase for adenosine nucleotides. We therefore renamed this effector Tas1 for type VI secretion effector (p)ppApp synthetase 1 and its cognate immunity protein Tis1 for type VI secretion immunity to (p)ppApp synthetase 1.

We next examined the catalytic rate of pppApp production by Tas1. Strikingly, one molecule of Tas1 was found to pyrophosphorylate 180,000 molecules of ATP per minute (Figure 3.6A). This catalytic rate is two orders of magnitude higher than characterized (p)ppGpp synthetases and likely reflects the role of Tas1 as an interbacterial toxin rather than an enzyme involved in growth rate control (Beljantseva et al., 2017; Gaca et al., 2015). Rapid turnover was also observed when ADP or AMP were used as pyrophosphate acceptors (Figure 3.6B). In addition to being unable to use GTP as a pyrophosphate donor or acceptor, Tas1 was unable to use dATP as an acceptor, though this deoxynucleotide could serve as a suboptimal pyrophosphate donor (Figure 3.6C). Alanine substitution of

a conserved glutamate residue known to bind the pyrophosphate donor ATP in RSH enzymes abolished Tas1 activity in vitro and toxicity during interbacterial competition (Figures 3.4F and 3.6A).

The remarkable catalytic rate of Tas1 predicts that T6SS-dependent delivery of one toxin molecule into a 1 μm^3 target bacterium would reduce ATP concentration by approximately 0.6 mM per minute. This calculation led us to hypothesize that Tas1 intoxicates cells in part by depleting essential adenosine nucleotides. To test this idea, we first examined nucleotide levels in *E. coli* cells expressing $Tas1_{tox}$. Within 30 minutes of $Tas1_{tox}$ expression, we observed a profound reduction in cellular AMP, ADP, and ATP levels that coincided with a substantial increase in pApp, ppApp and pppApp (Figure 3.7A-B).

We also examined nucleotide levels in *P. aeruginosa* cells natively expressing full-length Tas1 and depleted of the Tis1 immunity protein (Figure 3.8A) and observed a similarly large and rapid drop in ADP and ATP levels along with robust formation of pppApp and ppApp (hereafter called (p)ppApp) (Figure 3.8B-C). ADP and ATP levels were not completely abolished, suggesting that intoxicated cells attempt to compensate for the loss of these essential nucleotides by altering their metabolism. AMP levels remained unchanged in Tis1-depleted *P. aeruginosa* cells suggesting that at physiologically relevant concentrations of Tas1, ADP and ATP are the preferred adenosine nucleotide acceptors. Finally, we detected (p)ppApp during interbacterial competition between a PA14 donor and a Δ*tas1* Δ*tis1* recipient strain in a manner that was dependent on a functional T6SS in donor

cells (Figure 3.8E). Collectively, these results demonstrate that T6SS-delivered Tas1 depletes ADP and ATP in target bacteria by synthesizing (p)ppApp.

To compare the effects of (p)ppGpp and (p)ppApp production, we assessed the viability of *E. coli* cells expressing either a constitutively active fragment of the (p)ppGpp synthetase RelA (RelA′) or $Tas1_{tox}$. Even though the expression of both enzymes results in growth arrest, only cells undergoing Tas1-mediated intoxication showed a significant reduction in viability (Figures 3.7C and 3.8D, F-G). This difference likely arises because, in contrast to (p)ppApp production, (p)ppGpp production does not significantly reduce ATP levels and results in only a two-fold reduction in cellular GTP (Figure 3.8H-I). In line with these findings in *E. coli*, we also observed a substantial decrease in the viability of Tis1-depleted *P. aeruginosa* cells and during interbacterial competition with a Tas1-expressing donor strain (Figure 3.8J-K). These results indicate that the production of (p)ppGpp is bacteriostatic whereas (p)ppApp production by Tas1 is bactericidal.

(p)ppApp kills target cells in multiple ways

Our findings suggest that (p)ppApp affects target cell physiology by depleting ADP and ATP, which would have pleiotropic consequences for many cellular processes. In particular, ADP is an essential regulator of energy production due to its role in dissipating the proton motive force (pmf) via ATP synthase-catalyzed ATP production. Consequently, reduced levels of ADP following Tas1 delivery may produce excessive electrostatic potential across the inner membrane.

Consistent with this notion, we found that addition of sub-lethal levels of the pmf uncoupling ionophore CCCP reduced the toxicity of $Tas1_{tox}$ (Figure 3.9A-C). We also sought to test if Tas1-intoxicated cells can regenerate ADP by hydrolyzing ppApp. In Proteobacteria, the bifunctional RSH enzyme SpoT can cleave the 3′ pyrophosphate of ppGpp to produce GDP (Sarubbi et al., 1989). We found that the ppGpp-hydrolyzing domain of SpoT was substantially less active on ppApp than ppGpp *in vitro* (Figure 3.9D). Furthermore, expression of SpoT during interbacterial competition did not result in a change in ppApp levels (Figure 3.9E). Together, these data suggest that SpoT cannot alleviate Tas1 toxicity by regenerating ADP from ppApp.

In contrast to ADP, ATP is required for virtually all anabolic and catabolic pathways in bacteria. To examine the impact of Tas1-dependent ATP depletion, we performed metabolic profiling of *P. aeruginosa* cells depleted of the Tis1 immunity protein. These Tas1-intoxicated cells displayed a dramatic decrease in metabolites belonging to many essential pathways including glycolysis, TCA cycle, and the pentose-phosphate pathway, as well as decreases in intermediates of lipid, amino acid, pyrimidine and purine biosynthesis (Figure 3.10A and Table 3.3). Additionally, the levels of mononucleotide triphosphates and nucleotide-activated precursors involved in cell wall biosynthesis were substantially depleted (Figure 3.10A, Table 3.4). Thus, our results suggest that (p)ppApp production by Tas1 is bactericidal due to a decrease in ADP and ATP levels, leading to a dysregulation of the pmf and depletion of numerous metabolites required for cell viability.

We also considered the possibility that (p)ppApp itself contributes to Tas1-mediated toxicity by binding directly to protein targets. As with (p)ppGpp, (p)ppApp accumulation resulted in the reduction of *de novo* purine biosynthesis intermediates (Figure 3.10B). (p)ppGpp blocks the dedicated step of purine synthesis by competitively inhibiting PurF (Wang et al., 2019). Given its similarity to (p)ppGpp, we hypothesized that (p)ppApp could also inhibit PurF. Indeed, we found that (p)ppApp binds to and inhibits PurF from both *E. coli* ($PurF^{EC}$) and *P. aeruginosa* ($PurF^{PA}$) at concentrations of (p)ppApp achieved in Tas1-intoxicated cells (Figure 3.10C-F). To determine if the mode of PurF inhibition by the two nucleotides is similar, we determined the crystal structure of $PurF^{EC}$ in complex with ppApp to a resolution of 2.5 Å (Figure 3.11A-B, Table 3.1). Our structure indicated that despite differences in the purine rings, ppGpp and ppApp bind PurF in a similar manner, and mutation of an arginine residue required for ppGpp binding to $PurF^{EC}$ also ablated the ability of ppApp to bind and inhibit PurF (Figures 3.10E). Despite these similarities, our structural data reveal that the guanine base in ppGpp exhibits additional hydrogen bonding interactions with PurF compared to the adenine in ppApp (Figure 3.11C-D). This likely provides additional rotational flexibility ppApp, making it a slightly weaker inhibitor of PurF than ppGpp (Figure 3.10C-E). These data indicate that (p)ppApp directly inhibits purine biosynthesis via PurF and likely targets many, if not most, of the more than 50 proteins targeted by (p)ppGpp (Wang et al., 2019), further enhancing the toxicity of (p)ppApp that results from depletion of ADP and ATP.

Discovery and characterization of a bifunctional (p)ppGpp and (p)ppApp hydrolase

Most T6SS effectors that have been characterized, including Tas1, interact with immunity proteins that structurally block their active site and neutralize toxicity, preventing self-intoxication and permitting kin discrimination (Hood et al., 2010; LeRoux et al., 2015; Russell et al., 2014a). Recently, bifunctional immunity proteins for ADP-ribsoyltransferase toxins have been characterized (Ting et al., 2018). These immunity proteins not only directly inhibit their target effector by occluding the active site of their cognate toxin but also possess an ADP-ribosylhydrolase enzymatic domain that removes ADP-ribose modifications made by the toxin on target proteins (Ting et al., 2018). Because they do not require a large protein-protein interaction interface, enzymatic immunity proteins can confer generalized protection against a family of toxins, rather than inhibit a cognate toxin alone. Other more general immunity mechanisms have also been reported, typically involving changes to macromolecular cellular structures such as the cell envelope, which in turn, confers protection against toxins that target molecules within these structures (Hersch et al., 2020; Le et al., 2020).

Our analysis of Tas1-mediated toxicity in *P. aeruginosa* suggests that (p)ppApp production rapidly collapses the metabolic potential of the cell and because (p)ppApp is not turned over by any known mechanism, these nucleotides also inhibit essential downstream targets. However, unlike (p)ppApp, (p)ppGpp hydrolysis is a well-known function of many enzymes, most notably by the stringent response enzyme SpoT (Aravind and Koonin, 1998). Despite the similarity of

(p)ppApp and (p)ppGpp and some of their shared protein targets (Figure 3.11), we found that SpoT cannot cleave (p)ppApp. Small alarmone synthetases (SAS) and hydrolases (SAH) are a recently identified group of proteins that typically function in isolation from classical hydrolases like SpoT because they operate independent of the ribosome (Steinchen and Bange, 2016) (Figure 3.12A). Though the physiological roles of these proteins are not well understood, recent studies have reported that some SAS and SAH proteins can produce or hydrolyze (p)ppGpp and (p)ppApp, respectively (Fung et al., 2020; Jimmy et al., 2020a). To extend our analysis on the Tas1-Tis1 effector-immunity pair we characterized a novel SAH protein from *P. aeruginosa* to determine if it exhibits bifunctional activity toward (p)ppApp and (p)ppGpp and if so, whether this activity confers some degree of protection against Tas1 during interbacterial competition.

We identified *Pa*SAH using homology-based searches in *P. aeruginosa* PAO1 (Atkinson et al., 2011; Jimmy et al., 2020a). A multiple sequence alignment of *Pa*SAH with other (p)ppGpp hydrolase domain-containing proteins revealed that *Pa*SAH harbors five of the six catalytic “HD” motifs required for substrate recognition and hydrolysis (Figure 3.12B). However, *Pa*SAH lacks the conserved HD1 motif, which has been shown to coordinate the guanine base in (p)ppGpp and is essential for (p)ppGpp hydrolysis in other hydrolases (Hogg et al., 2004; Sun et al., 2010; Tamman et al., 2020). To determine if *Pa*SAH is structurally and functionally similar to known (p)ppGpp hydrolases, we determined its crystal structure a resolution of 1.85 Å (Figure 3.12C and Table 3.1). The structure shows

a monomer of *Pa*SAH consisting of a bundle of nine anti-parallel α-helices that form a deep cavity along one side (Figure 3.12C). We also noted electron density present within the central cavity of *Pa*SAH that we modelled as a Mn^{2+} ion with 50% occupancy, which is coordinated by two histidine (H40, H70) and two aspartate residues (D71, D131) (Figure 3.12D). The Mn^{2+} ion is essential for (p)ppGpp hydrolysis and its coordination is similar to that observed in the crystal structures of the MESH1 (PDB: 3NR1, 3NQW), Lmo0812 (PDB: 4YF1) and Rel (PDB: 6S2V, 1VJ7) (p)ppGpp hydrolases. To determine if *Pa*SAH functions similar to other (p)ppGpp hydrolases, we tested its enzymatic activity in vitro using (p)ppGpp as a substrate. Similar to known hydrolases, *Pa*SAH cleaves the 3'-pyrophosphate moiety of ppGpp to form GDP and its activity is dependent on Mn^{2+}-coordinating residues (H70, D71) (Figure 3.12E). Kinetic assays on both ppGpp and pppGpp confirmed that *Pa*SAH exhibits similar affinities and maximal velocities for both substrates (Figure 3.12F).

We next compared substrate affinity and reaction velocities of *Pa*SAH on (p)ppApp. Remarkably, we found that both ppApp and pppApp are degraded by *Pa*SAH at rates 5- to 10-fold higher than their ppGpp and pppGpp counterparts (Figure 3.13A). Based on these data, we hypothesized that *Pa*SAH may be able to provide protection against Tas1 intoxication. To assess this, we measured the growth of *E. coli* cells expressing the Tas1 toxin in presence of Tis1, *Pa*SAH or the inactive *Pa*SAH$^{E74A/D75A}$ variant. Co-expression with *Pa*SAH resulted in a marked increase in *E. coli* growth rate compared to cells expressing Tas1 alone, though,

this protective effect was intermediate to Tas1 and Tas1-Tis1 expressing cells (Figure 3.13B). By contrast, *Pa*SAH$^{E74A/D75A}$ was unable to restore *E. coli* growth indicating that *Pa*SAH is detoxifying Tas1 enzymatically via (p)ppApp degradation. We previously demonstrated that Tas1 reduces the viability of actively growing *E. coli* by approximately three orders of magnitude 60 minutes after inducing toxin expression (Figure 3.8G). To assess the impact of *Pa*SAH on cell viability of actively growing *E. coli*, we measured the viability of *E. coli* expressing Tas1 in the presence of Tis1 and *Pa*SAH. Similar to our growth rate measurements, the expression of *Pa*SAH had an intermediate protective effect, reducing the Tas1-dependent loss in cell viability by ~1000-fold compared to cells expressing the toxin in isolation. This substantial protective effect is ~10-fold lower than the Tis1-expressing cells, which do not show any reduction in viability compared to a control strain (Figure 3.13C).

Having established a protective effect for *Pa*SAH against Tas1 in a heterologous system, we hypothesized that *Pa*SAH may also partially protect against Tas1 during interbacterial competition by providing enzymatic immunity. As shown previously, *P. aeruginosa* PA14 has a significant Tas1-dependent growth advantage over PAO1 because the latter strain lacks the Tas1-Tis1 effector-immunity pair (Figure 3.2A). *Pa*SAH is found in both *P. aeruginosa* PAO1 and PA14, however, having structurally and biochemically characterized the protein from PAO1, we initiated our experimentation using interstrain competitions between PAO1 and PA14. Co-culture competition assays between a T6SS-active

strain of PA14 with a *Pa*SAH-deficient mutant of PAO1 resulted in a 12-fold increase in competitive index relative to a competition involving the PAO1 parent strain, suggesting that the *Pa*SAH mutant strain is more susceptible to T6SS-delivered Tas1 (Figure 3.13D). Plasmid-borne expression of *Pa*SAH, but not *Pa*SAH$^{E74A/D75A}$, in the *Pa*SAH-deficient mutant abolished the Tas1-dependent fitness advantage of PA14 over PAO1. Competition of a PA14 donor strain harbouring a Tas1-inactivating mutation (E382A) with PAO1 did not reveal any *Pa*SAH-dependent fitness advantages demonstrating that *Pa*SAH specifically provides protection against Tas1 and not other T6SS effectors (Figure 3.13D). We also noted that adjusting the PAO1:PA14 ratio in favour of PA14 in this assay reduced the protective effect of *Pa*SAH against PA14 attackers (Figure 3.13E-F). This likely occurs because *Pa*SAH is an enzymatic form of immunity against Tas1 and based on our work in *E. coli*, does not provide complete protection against Tas1. Therefore, as increased amounts of PA14 are introduced to the competition and the frequency of Tas1-injection events increase, the number of *Pa*SAH copies in PAO1 cells become insufficient to overcome (p)ppApp-induced toxicity.

To directly compare the protective effects of Tis1, the cognate immunity to Tas1, and *Pa*SAH, we next conducted intraspecific competition assays using *P. aeruginosa* PA14 strains lacking these genes. Co-culture of PA14 donors with a strain lacking only the PA14 *Pa*SAH homolog did not have an impact on the growth of donor strains (Figure 3.13G, first bar) and deletion of the Tas1-Tis1 pair in recipients increased the donor competitive index by approximately ~7-fold, as

previously shown (Figures 3.2A and 3.13G, second bar). Remarkably, deletion of *Pa*SAH in the Δ*tas1* Δ*tis1* parent background resulted in a 150-fold increase in the competitive index of donors and resulted in a dramatic loss of recipient cell viability (Figure 3.13G, fourth bar and Figure 3.13H). Plasmid-borne expression of Tis1 or *Pa*SAH improved the competitive fitness of the recipient strain, but neither provided protection to the same extent as Tis1 expression in the Δ*tas1* Δ*tis1* recipient. Collectively, these results suggest that *Pa*SAH reduces the toxicity of T6SS-delivered Tas1 in *P. aeruginosa* strains lacking Tis1 by enzymatically degrading (p)ppApp.

Discovery of Tas1 homologs that are not associated with a T6SS

Enzymes that synthesize (p)ppGpp are highly conserved across bacteria and play an important role in bacterial survival during nutritional stress (Hauryliuk et al., 2015). Comparatively, the role of (p)ppApp in physiology is not well known. Previous work from several decades ago showed that (p)ppApp initiates sporulation in *Bacillus* species and inhibits spore germination in some *Streptomyces* species (Hamagishi et al., 1980; Oki et al., 1975; Rhaese and Groscurth, 1976). Multiple enzymes from *Streptomyces* were also found to synthesize (p)ppApp, however, none of these proteins were shown to be directly involved with the inhibition of spore germination (Oki et al., 1976). Other reports also suggest that some previously characterized (p)ppGpp synthetases are bifunctional enzymes that are capable of synthesizing both (p)ppGpp and (p)ppApp

synthesis activity (Fung et al., 2020). Additionally, (p)ppApp-synthesizing toxins were recently found to be associated with toxin-antitoxin systems in several bacterial species (Jimmy et al., 2020a). Collectively, these findings suggest that (p)ppApp synthetases exist in other bacteria. Having determined that Tas1 is the first known monofunctional (p)ppApp synthetase, we used the sequence of its toxin domain as a seed to identify homologs in diverse bacteria that may also possess this enzymatic activity.

We used *jackhammer* to generate a sequence alignment hidden Markov model (HMM) for the toxin domain of Tas1 using an iterative search procedure that queried the UniprotKB database (Johnson et al., 2010). Two iterations yielded 2,468 sequences that were similar in size to Tas1 (~200 residues) and contained all the catalytic residues predicted to be necessary for (p)ppApp synthesis (Figure 3.4). More iterations yielded thousands of hits that were primarily annotated as the (p)ppGpp synthetases RelA and SpoT, which are highly conserved across bacteria and significantly larger (~700 residues) than Tas1 and thus, these hits were not used for our analysis. The candidate Tas1-like sequences were used to construct a maximum likelihood phylogenetic tree in which sequences with greater than 24.5% homology were grouped together (Cohen et al., 2019). A total of 14 groups were identified. Five homologs from each group were manually analyzed using secondary structure, signal sequence and transmembrane domain prediction servers (Käll et al., 2007; Kelley et al., 2015). The genomic context (10 kb upstream and downstream) of each homolog was also analyzed to determine potential

association to a T6SS or other pathway. Remarkably, this analysis provided clear distinction of three broad protein families: (p)ppGpp synthetases, bifunctional (p)ppGpp/(p)ppApp synthetases and monofunctional (p)ppApp synthetases (Figure 3.14A). Predictions for bifunctional synthetases were made based on the recently identified bifunctional synthetases in *B. subtilis* and *Streptococcus pyogenes* (Fung et al., 2020). Approximately 1,300 sequences belonged to the (p)ppGpp synthetase or (p)ppGpp/(p)ppApp bifunctional synthetase families, however, the remaining ~900 sequences all showed high homology to Tas1 and appeared to be more evolutionary distant from (p)ppGpp synthetases than Tas1 (Figure 3.14A). Notably, only a small subset of these predicted (p)ppApp synthetases were found to be associated with T6SSs. In total, we identified predicted (p)ppApp synthetases that belonged to six distinct groups: 1) T6SS-associated, 2) T7SS-associated, 3) *Vibrio* group, 4) MuF prophage-associated, 5) *Streptomyces* group and 6) T3SS-associated.

Predicted (p)ppApp synthetases from each group also possess notable sequence signatures that are conserved in Tas1 (R330, N370, E428, Y427), but absent in the (p)ppGpp synthetases RelA and Rel_{Seq} or the bifunctional (p)ppGpp synthetases RelQ and RelP, confirming the evolutionary distinction of these proteins from (p)ppGpp synthetases (Figure 3.14B) (Hogg et al., 2004; Manav et al., 2018; Steinchen et al., 2015). Mapping these conserved residues to the structure of $Tas1_{tox}$ and structurally aligning this structure with RelP in complex with AMPCPP and GTP reveals that all four conserved residues are present within

active site of $Tas1_{tox}$. Two residues in Tas1, Y427 and E428, are directly facing the GTP acceptor molecule in the RelP structure and thus may be participating in important π-stacking and hydrogen bonding interactions, respectively, that favour the binding of adenosine rather than guanosine acceptor nucleotides (Figure 3.14C).

While some of the Tas1 homologs identified from our phylogenetic analysis appear to be associated with a type III or type VII secretion system and likely function as toxins, many homologs from the *Vibrio*, prophage (MuF) and the *Streptomyces* groups are not associated with a known specialized secretion system suggesting that these proteins may have functions that are distinct from Tas1 and extend beyond inhibiting target cell growth (Figure 3.15A). The *Vibrio* sequences are highly conserved in this group and are annotated as 'RelV', an enzyme thought to function as a (p)ppGpp synthetase (Das et al., 2009; Dasgupta et al., 2014). While the in vitro evidence from these studies suggests that RelV is active on a mixture of ATP and GTP, the products of these reactions were analyzed only by thin layer chromatography and not subject to further quantitative analyses that better differentiates between adenosine versus guanosine containing products. Further, RelV possesses a Sec-signal sequence at its N-terminus, which was not described in the previous studies and is a feature that has not been shown to be associated with (p)ppGpp synthetases (Hauryliuk et al., 2015). While *Vibrio* species do possess an active T6SS (Fridman et al., 2020; Pukatzki et al., 2006; Santoriello et al., 2020), it is unlikely that RelV is a T6SS effector because of its N-

terminal signal sequence, as all previously characterized T6SS effectors are recruited to the apparatus from the cytoplasm (Klein et al., 2020). The Tas1 homologs found to be associated with prophage gene clusters contain the poorly characterized MuF/gp7 domain at their N-terminus (Vinga et al., 2006). Previous in silico analyses identified several proteins containing MuF domains fused to polymorphic C-terminal toxin domains, but no toxin in this family has been biochemically or phenotypically characterized (Jamet et al., 2017). Lastly, the *Streptomyces* Tas1 homologs are highly conserved across different species within this genus and are primarily annotated as ATP 3'-pyrophosphokinases. As noted above, previous work has shown that enzymes in *Streptomyces* synthesize these nucleotides (Oki et al., 1976), however, no direct effect of these enzymes on the physiology of these organisms has been identified. Similar to the *Vibrio* homologs, we identified an N-terminal signal sequence with the *Streptomyces* proteins that has not been previously described. These analyses collectively suggest that Tas1 homologs from these three broad groups do not function intracellularly and thus likely do not function as (p)ppGpp synthetases.

To begin our studies of these putative (p)ppApp synthetases, we first conducted biochemical assays to determine if representative Tas1 homologs from each group synthesize (p)ppApp as predicted by our informatics analyses. To start, we selected Tas1 homologs from *Vibrio parahaemolyticus* (*Vp*, *Vibrio* group), *Bacteroides caccae* (*Bc*, MuF group) and *Streptomyces albidoflavus* (*Sa*, *Streptomyces* group) for further characterization (Figure 3.15A). These sequences

were selected based on genetic tractability of their native organism, conservation with other sequences in a group and representation of the diversity of other sequences in the group. As shown previously, Tas1 rapidly kills *E. coli* shortly after its expression is induced (Figure 3.7). To determine the effects of each homolog on *E. coli* growth, we overexpressed regions of each protein that exhibited the greatest resemblance to Tas1 (Figure 3.15A, blue regions) in *E. coli* using a rhamnose-inducible vector. Dilution plating overnight cultures expressing each protein on rhamnose-containing agar resulted in significant growth inhibition for all homologs tested, suggesting that these proteins can also function as toxins when overexpressed in *E. coli* (Figure 3.15B). Monitoring cell growth and cell viability at various time points following induction revealed that only the *Sa* toxin, like Tas1, reduces the viability of *E. coli* (Figure 3.15C-D). Both the *Vp* and *Bc* toxins are instead bacteriostatic, which is similar effect to that observed following RelA' expression in *E. coli* (Figure 3.8G). To determine if these proteins produce (p)ppApp, we purified each homolog following recombinant expression in *E. coli* and incubated each toxin with ATP to test for pppApp/pApp synthesis or ATP and ADP to assay for ppApp synthesis. In line with our sequence predictions, all three Tas1 homologs produced both pppApp and ppApp in vitro (Figure 3.15E, top: pppApp, bottom: ppApp). We did not find any in vitro activity of these enzymes when they were incubated with GTP or GDP in the presence of ATP suggesting that these proteins do not function as (p)ppGpp synthetases. In accordance with the newfound catalytic activity of these enzymes, we renamed these proteins to

Adenosine 3'-pyrophosphokinases (Apk) 1-4, with Tas1 being the founding member Apk1 (*Bc* homolog is Apk2; *Sa* homolog is Apk3; *Vp* homolog is Apk4) (Figure 3.15A). The different effects of these toxins on *E. coli* growth suggest that these homologs produce varying amounts of (p)ppApp when expressed in cells. To determine whether this is the case, we extracted metabolites from *E. coli* expressing either Apk2 and Apk3 or the predicted catalytic mutants of both homologs based on the alignment with Tas1 (Figure 3.14B). In line with our in vitro experiments, both Apk2 and Apk3 produce (p)ppApp in vivo and not (p)ppGpp (Figure 3.15F). Importantly, mutants of both homologs showed a significant reduction in (p)ppApp synthesis, suggesting that these proteins require similar catalytic residues as Tas1 to bind donor and acceptor nucleotides. The Apk3 toxin from *Streptomyces* produces significantly greater amounts of pApp in the cell compared to the Apk2 toxin. We previously showed that pApp synthesis can be catalyzed by Tas1 through a pyrophosphate transfer from pppApp to AMP (Figure 3.6E). The Apk3 toxin is the only toxin of the three Tas1 homologs tested that kills *E. coli*, thus it is possible that this bactericidal effect is due to pApp synthesis, which would consume the available 5'-pyrophosphate pool in the cell, causing metabolic dysregulation and cell death, similar to Tas1. A comparison of the kinetics of these proteins will provide valuable insight into their functions in cells. It is also clear that each enzyme produces different amounts of pppApp, ppApp and pApp and future experiments should aim to characterize the total metabolic changes taking place in cells following expression of these proteins.

We also showed that to prevent self-intoxication, Tas1 from *P. aeruginosa* is co-expressed with and neutralized by its cognate immunity protein, Tis1 (Figure 3.2). The toxicity of Apk2-Apk4 prompted us to determine whether there are immunity genes associated with any of these toxins. The genomic context of the Apk3 and Apk4 proteins from *S. albidoflavus* and *V. parahaemolyticus*, respectively, did not suggest co-association with an immunity gene for either toxin (Figure 3.15A). However, both toxins do possess predicted N-terminal Sec signal sequences, which may mitigate the requirement for a cytosolic immunity protein as the Sec general secretion machinery typically transports proteins co-translationally into the periplasm (Costa et al., 2015). By contrast, the Apk2 toxin from *B. caccae* is encoded beside two small co-directional genes (Figure 3.16A). The gene directly downstream of *apk2* encodes for a Tis1-homologous immunity protein (referred to as *Bc*Tis1) followed by another gene that encodes for a homolog of the (p)ppGpp hydrolase MESH-1 from *Drosophila melanogaster* (referred to as *Bc*SAH) (Sun et al., 2010). To determine if either of these proteins protects against the activity of Apk2, we co-expressed each gene, *Bc*Tis1 or *Bc*SAH, with Apk2 in *E. coli*. Surprisingly, given our findings with *P. aeruginosa* Tis1, the MESH-1 hydrolase-like protein conferred significantly greater protection against Apk2 toxicity than the *Bc*Tis1 protein (Figure 3.16B). Interestingly, however, expression of the complete *Bc*Tis1-*Bc*SAH bicistron with Apk2 resulted in complete protection against the toxin (Figure 3.16B). Similar to observations made for the bifunctional (p)ppGpp/(p)ppApp hydrolase *Pa*SAH, *Bc*SAH only confers partial protection

against this toxin (Figure 3.13E). To determine if *Bc*SAH could broadly protect against other (p)ppApp synthetases, we co-expressed it with $Tas1_{tox}$. Similar to the observations made for *Pa*SAH, we found that *Bc*SAH confers partial protection against $Tas1_{tox}$ (Figure 3.16C). We also tested whether the homologous Tis1 proteins may confer cross-protection against Tas1 and Apk2. Expression of *Bc*Tis1 with Tas1 or Tis1 with Apk2 did not confer any protection, suggesting that these proteins are highly specific for their cognate toxins (Figure 3.16B-C). Together, these data suggest that *Bc*Tis1 is a structural immunity protein, while *Bc*SAH is an enzymatic immunity protein. We next tested if *Bc*SAH, like *Pa*SAH, also hydrolyzes (p)ppApp to protect against Tas1. In line with our growth assays, *Bc*SAH readily converts pppApp and ppApp to ATP and ADP, respectively (Figure 3.16D). However, in contrast to *Pa*SAH and other known (p)ppGpp hydrolases, *Bc*SAH exhibits no activity towards ppGpp, suggesting that we have identified the first example of a monofunctional (p)ppApp hydrolase enzyme (Figure 3.16D). Because *Bc*SAH confers partial protects against Tas1, we wondered if this hydrolase also protects *E. coli* against the other identified Apk enzymes. Intriguingly, we found that *Bc*SAH confers protection against Apk2 but not Apk3 or Apk4 (Figure 3.16E). We also tested *Pa*SAH with these enzymes and similarly noted that it protects only against Apk2 (Figure 3.16E). Though our in vitro data and metabolite profiling experiments data suggest that the Apk enzymes produce (p)ppApp (Figure 3.15E-F), it is also possible that these enzymes produce other toxic nucleotides or other toxic products in the cell that cannot be cleaved by the

hydrolases described. Characterization of the complete metabolome of Apk-expressing cells will be necessary to address this possibility.

To determine the molecular basis for the specificity of the *Bc*SAH hydrolase for (p)ppApp, we solved its structure to a resolution of 2.3 Å (Figure 3.17A and Table 3.1). The structure contains eight α-helices and a manganese ion within its predicted active site that is stabilized by histidine-aspartate (HD) motifs. Both the manganese cofactor and the HD motifs are necessary for the function of other small alarmone hydrolases, including *Pa*SAH (Figure 3.12) (Hogg et al., 2004; Ruwe et al., 2018; Tamman et al., 2020). To understand more about the specificity of this hydrolase for adenosine nucleotides, we overlayed its structure with the previously characterized bifunctional (p)ppGpp/(p)ppApp hydrolase *Pa*SAH (Figures 3.12 and 3.13) and the previously characterized (p)ppGpp hydrolase *Se*Rel from *Streptococcus equisimilis*. The two enzymes exhibit significant structural similarity with *Bc*SAH (Cα root mean square deviation (r.m.s.d.) of 1.1 Å over 70 equivalent positions with *Pa*SAH and 1.2 Å across 67 equivalent positions with *Se*Rel), corroborating our homology-based findings that *Bc*SAH is a hydrolase of the MESH-1 superfamily (Figure 3.17A-B) (Aravind and Koonin, 1998). The structural alignment also revealed that *Bc*SAH possesses fewer total helices than both *Pa*SAH and *Se*Rel which contain 9 and 10 α-helices, respectively, whereas *Bc*SAH possesses only 8 (Figure 3.17B). With the exception of *Pa*SAH, all known (p)ppGpp hydrolases possess 10 α-helices (Pausch et al., 2020; Ruwe et al., 2018; Sun et al., 2010; Tamman et al., 2020). It is possible that all 10 helices are required

for (p)ppGpp hydrolysis and the absence of one or more of these helices instead favours (p)ppApp binding by *Pa*SAH and *Bc*SAH. The additional helices in *Se*Rel and *Pa*SAH are not known to contain catalytic residues necessary for function so their functional significance is hard to predict. How the absence of these α-helices affects the specificity of *Bc*SAH for (p)ppApp is a possible avenue of future investigation.

To determine other differences that may exist between (p)ppApp hydrolases and (p)ppGpp hydrolases, we next compared the sequence of *Bc*SAH with known and predicted hydrolases. We manually identified two other predicted (p)ppApp hydrolases from *Streptococcus pneumoniae* (*Sp*SAH) and *Capnocytophaga haemolytica* (*Ch*SAH) that share a similar genomic context with *Bc*SAH. Other sequences used for comparison included *Pa*SAH and (p)ppGpp hydrolases identified *Corynebacterium glutamicum* (*Cg*SAH) (Ruwe et al., 2018), *D. melanogaster* (*Dm*MESH) (Sun et al., 2010), *H.* sapiens (*Hs*MESH) (Sun et al., 2010), *Thermus thermophilus* (*Tt*Rel) (Tamman et al., 2020), *Se*Rel and the broadly conserved SpoT enzyme from *E. coli* (*Ec*SpoT). The alignment revealed highly conserved residues amongst (p)ppGpp hydrolases and several residues that were conserved between *Bc*SAH and the other MuF-associated *Ch*SAH and *Sp*SAH sequences (Figure 3.17C). Mapping these residues to the *Bc*SAH structure revealed that many of them appear distant form the active site (Figure 3.17D-E). However, Y90, D127 and K133 appear directly beside the active site and may have an important role in determining specificity of *Bc*SAH for adenosine containing

nucleotides through π-stacking interactions, hydrogen bonding with the unique amino group of adenine and stabilizing negatively charged phosphate groups, respectively. Mutagenesis experiments will be crucial in determining which residues are necessary for determining adenosine specificity in *Bc*SAH and other related hydrolases.

Discussion

Tas1 kills cells by synthesizing (p)ppApp

Our work demonstrates that Tas1 is a novel interbacterial toxin and represents the first case of an RSH protein that is delivered between bacterial cells. Tas1 is also, to our knowledge, the first (p)ppApp synthetase enzyme with a known role in bacterial physiology. All previously characterized RSH enzymes synthesize (p)ppGpp, which regulates cell growth rate and promotes bacterial survival. Although (p)ppApp is very similar to (p)ppGpp in structure, its physiological role differs because its production irreversibly alters the cellular metabolome, depleting existing pools of ATP and hindering the ability of intoxicated cells to synthesize ATP. As the *P. aeruginosa* H1-T6SS delivers a diverse payload of effectors into target cells, the (p)ppApp synthetase activity of Tas1 likely augments the activities of co-secreted cell wall and membrane targeting effectors because pathways involved in cell envelope biosynthesis and repair require ATP (Bugg et al., 2011; LaCourse et al., 2018; Raetz, 1978). Although reports from several decades ago linked (p)ppApp production to sporulation in *B. subtilis* and inhibition of spore

germination in *Streptomyces* spp. (Hamagishi et al., 1980; Oki et al., 1975; Rhaese and Groscurth, 1976; Rhaese et al., 1977), a physiological role for this molecule had never been elucidated. Our discovery of Tas1 now indicates that (p)ppApp is a physiologically relevant molecule that can serve as a potent cellular toxin.

(p)ppApp synthetases are broadly distributed in bacteria

Our characterization of Tas1 has enabled the identification of hundreds of putative (p)ppApp synthetases across several bacterial phyla. Many of these homologs appear in diverse genomic contexts and only a small fraction of the predicted synthetases appear to be associated with interbacterial antagonism pathways such as T6SSs or T7SSs. Our preliminary analyses of three Tas1 homologs found in temperate phages, *Vibrio* spp. and *Streptomyces* spp. indicate that these proteins synthesize (p)ppApp and inhibit target cell growth, however, they could also be synthesizing other toxic products that inhibit target cell growth. Future metabolite profiling experiments will be essential in determining the complete substrate range for these toxins. Most importantly, these enzymes will need to be studied in the context of their native organism to determine the conditions under which they are expressed, whether or not they are secreted, and the effects they may have on growth of both the synthetase-producing bacterium and on any potential target organisms.

(p)ppApp hydrolases confer immunity against Tas1

In addition to the identification of novel (p)ppApp synthetases, we also identify two non-structural immunity proteins that hydrolyze (p)ppApp and protect against Tas1-mediated intoxication. While the physiological roles of these proteins are not known, we demonstrate that both hydrolases confer a significant protective effect when co-expressed in cells expressing Tas1. (p)ppGpp hydrolases are well conserved in many bacteria and our structural work demonstrates that (p)ppApp hydrolases are structurally distinct. Mutagenesis and enzyme kinetics experiments will shed light on the specificity of these novel hydrolases for (p)ppApp.

Our work provides the first characterization of a T6SS-associated (p)ppApp synthetase toxin and suggests that (p)ppApp is an important metabolite in inhibiting target cell growth. However, the identification of several other (p)ppApp synthetases that appear in genomic contexts distinct from Tas1 suggests that RSH proteins that synthesize (p)ppApp may serve biological roles that extend beyond interbacterial antagonism and my preliminary findings described herein set the stage for the future exploration of the potentially novel roles that these novel nucleotides play in bacterial physiology.

Figures

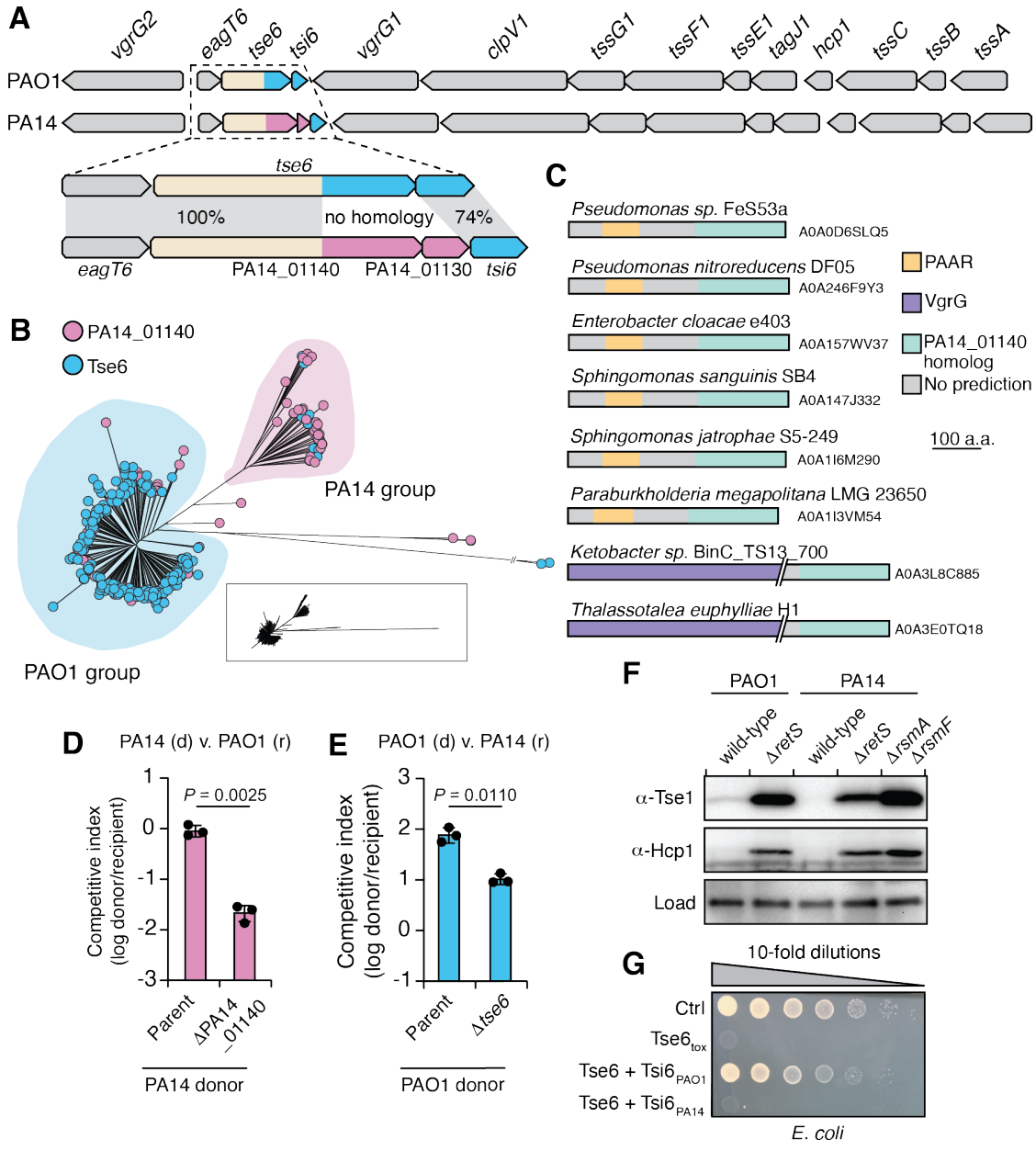


Figure 3.1. A unique T6SS effector-immunity pair is encoded within the H1-T6SS of *P. aeruginosa* strain PA14. A) Genomic context of *tse6-tsi6* and PA14_01140-PA14_01130 within the H1-T6SS gene clusters of *P. aeruginosa* strains PAO1 and PA14, respectively. Known toxin-immunity encoding regions of *tse6-tsi6* and predicted toxin-immunity encoding regions of PA14_01140-

PA14_01130 are shown in blue and pink, respectively. **B)** Phylogenetic distribution of PA14_01140 (pink) and *tse6* (blue) within 326 *P. aeruginosa* genomes based on whole-genome SNP maximum likelihood analysis. Circles denote individual *P. aeruginosa* strains. Each clade is labeled according to its representative member. Miniaturized tree depicts true branch distance between each clade. The full tree in Newick format, including bootstrap values, is provided as Appendix Dataset A3.1. **C)** Proteins containing a domain homologous to the C-terminus of PA14_01140 are found in several different species of Proteobacteria. Homologs were identified using the HMMER webserver and candidate T6SS effectors were selected based on the presence of predicted N-terminal domains known to facilitate export by the T6SS. The UniProtKB accession number for each identified protein is indicated. **D-E)** Outcome of growth competition assays between the indicated donor (d) and recipient (r) strains. The parental PA14 genotype is Δ*rsmA* Δ*rsmF* and the parental PAO1 genotype is Δ*retS*, both of which are mutations that stimulate H1-T6SS activity (Goodman et al., 2004; Marden et al., 2013). **F)** Expression levels of the conserved H1-T6SS effector Tse1 and the secreted H1-T6SS subunit Hcp1 are similar between *P. aeruginosa* PAO1 Δ*retS* and *P. aeruginosa* PA14 Δ*rsmA* Δ*rsmF*. Western blot analysis of Tse1 and Hcp1 in the indicated *P. aeruginosa* strains. A non-specific band that reacts with the α-Tse1 antiserum was used as a loading control. **G)** $Tsi6^{PA14}$ is not protective against Tse6-mediated intoxication. Viability of *E. coli* cells grown on solid media harboring inducible plasmids expressing $Tse6_{tox}$, $Tse6_{tox}$ + $Tsi6^{PAO1}$, $Tse6_{tox}$ + $Tsi6^{PA14}$, or an empty vector control. **D-E)** Data are mean ±SD for three biological replicates. *P* values are shown from two-tailed, unpaired *t*-tests.

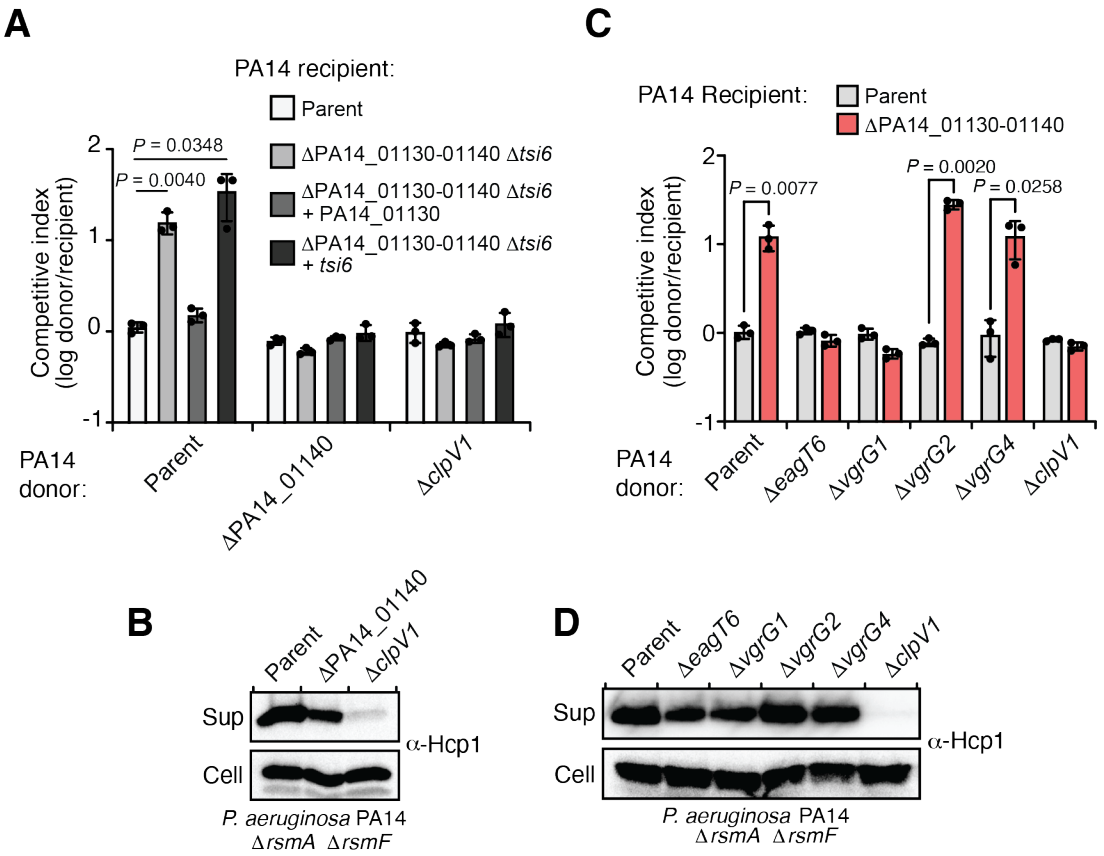


Figure 3.2. PA14_01140 inhibits target cell growth in a H1-T6SS-dependent manner. A) Outcome of intraspecific growth competitions between the indicated PA14 donor and recipient strains. The parental PA14 strain genotype is Δ*rsmA* Δ*rsmF.* The competitive index is normalized to starting ratios of donor/recipient. **B)** Mutational inactivation of PA14_01140 does not abrogate Hcp1 secretion. Western blot analysis of Hcp1 levels in the cell and supernatant (sup) fractions of the indicated *P. aeruginosa* PA14 strains. **C)** PA14_01140 delivery into recipient cells requires the H1-T6SS exported protein VgrG1 and the Tse6-specific chaperone EagT6. Intraspecific growth competition assay between indicated PA14 donor and recipient strains. The parental strain genotype is Δ*rsmA* Δ*rsmF*. **D)** Mutational inactivation of *eagT6*, *vgrG1*, *vgrG2* and *vgrG4* does not abrogate H1-T6SS function. Western blot analysis of Hcp1 levels in the cell and supernatant (sup) fractions of the indicated *P. aeruginosa* PA14 strains. **A, C)** Data are mean ±SD for three biological replicates. *P* values are shown from two-tailed, unpaired *t*-tests.

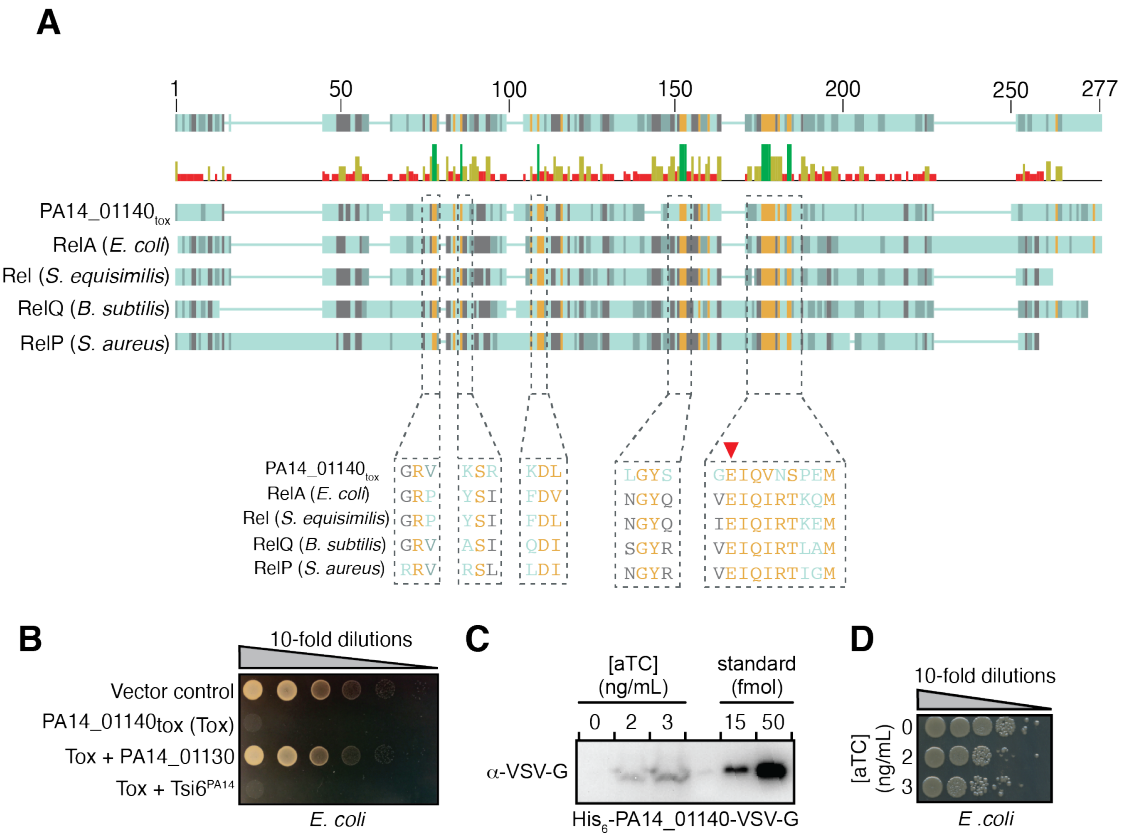


Figure 3.3. PA14_01140$_{tox}$ is a potent bacterial toxin that possesses remote homology to characterized (p)ppGpp synthetases. A) ClustalW alignment of PA14_01140$_{tox}$, the RSH domains of *E. coli* RelA (Magnusson et al., 2005) and *Streptococcus equisimilis* Rel (Hogg et al., 2004), and the small alarmone synthetases RelQ (Steinchen et al., 2015) and RelP (Manav et al., 2018) from *Bacillus subtilis* and *Staphylococcus aureus*, respectively. Dashed boxes represent regions of high sequence homology. The catalytic glutamic acid is indicated by a red triangle. **B)** PA14_01130 but not Tsi6^{PA14} inhibits PA14_01140$_{tox}$ mediated toxicity. Viability of *E. coli* cells grown on solid media harboring inducible plasmids expressing PA14_01140$_{tox}$, PA14_01140$_{tox}$ + PA14_01130, PA14_01140$_{tox}$ + Tsi6^{PA14}, or an empty vector control. **C-D)** PA14_01140$_{tox}$ is toxic to *E. coli*, even when expressed at approximately three copies per cell. Western blot analysis of pull-downs from *E. coli* expressing His$_6$-PA14_01140$_{tox}$-VSV-G in the presence of the indicated concentrations anhydrotetracyline (aTC) inducer **(C)**. See "Quantification of Tas1$_{tox}$ overexpression in *E. coli*" in Methods for details. Viability of *E. coli* cells expressing His$_6$-PA14_01140$_{tox}$-VSV-G in the presence of the indicated aTC concentrations for 15 minutes **(D)**.

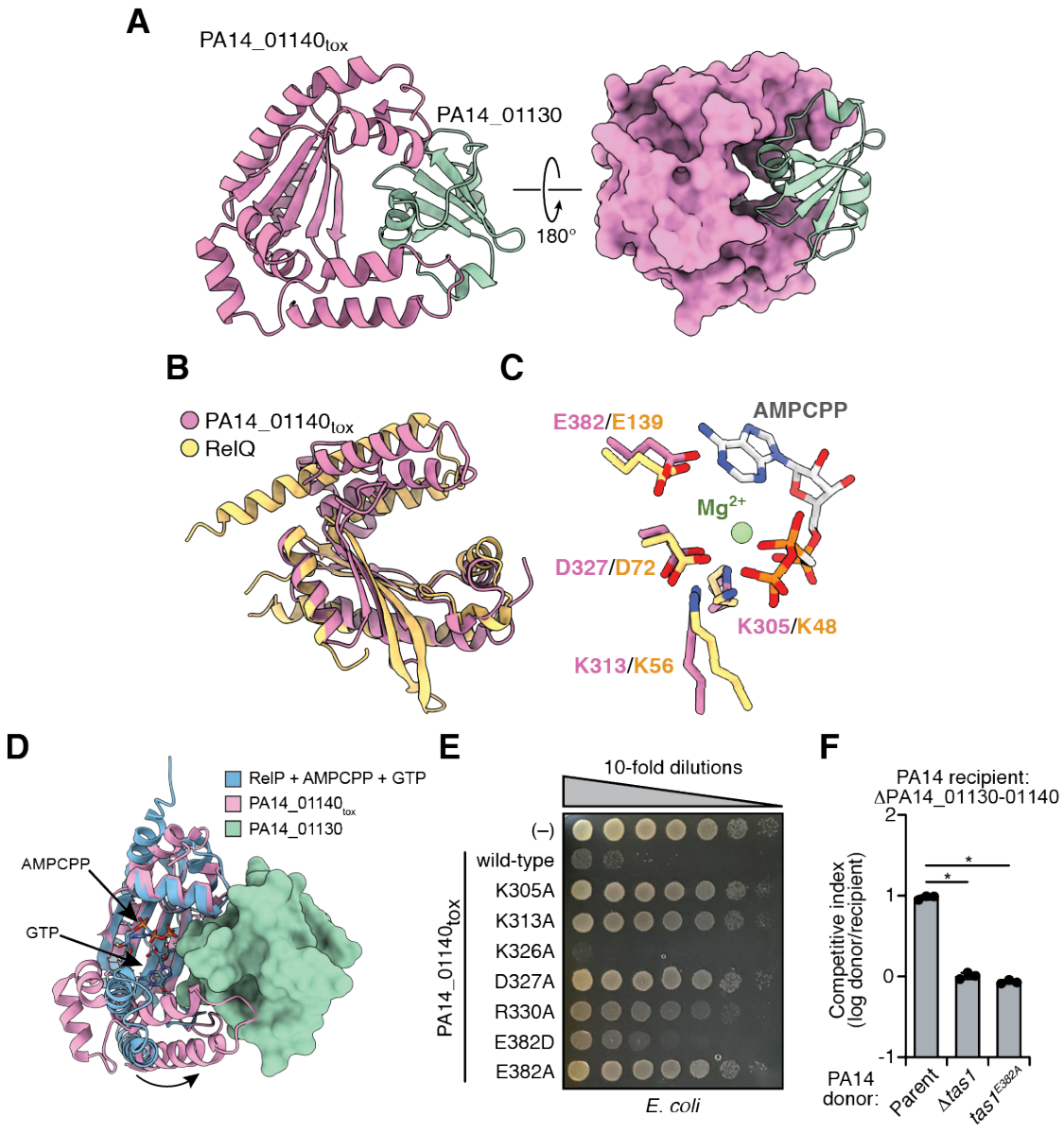
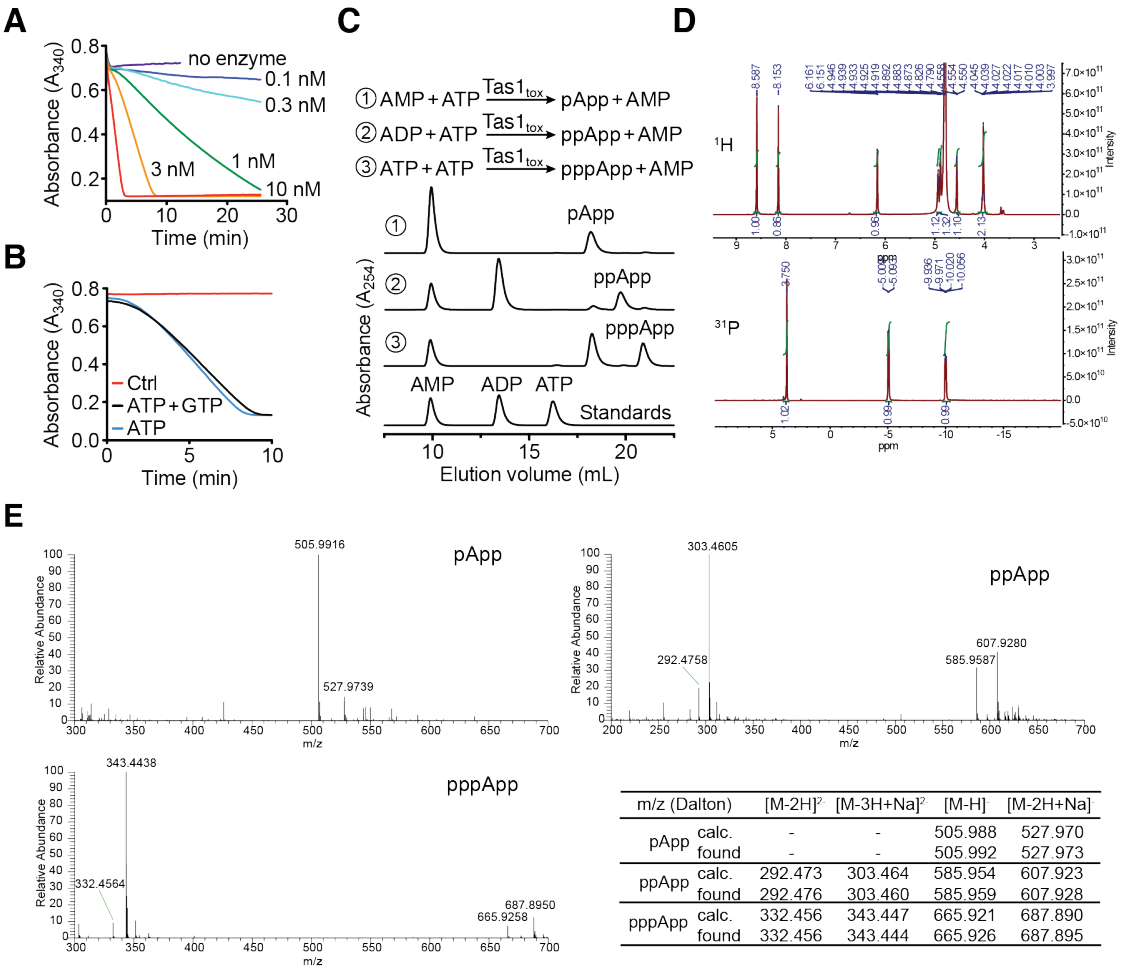


Figure 3.4. Tas1$_{tox}$ adopts a RelA-SpoT Homolog (RSH) fold found in enzymes that synthesize the bacterial alarmone (p)ppGpp. A) Overall structure of PA14_01140$_{tox}$ in complex with PA14_01130. Shown are ribbon (left) and space-filling (right) representations of PA14_01140$_{tox}$ (purple) in complex with ribbon representations of PA14_01130 (green). **B)** PA14_01140$_{tox}$ resembles (p)ppGpp synthetase enzymes. Structural overlay of PA14_01140$_{tox}$ and the small alarmone synthetase RelQ from *Bacillus subtilis* (PDB code 5DEC) (Steinchen et al., 2015). The structures superimpose with a Cα r.s.m.d. of 3.4Å over 145 equivalent positions. **C)** Structural alignment of the pyrophosphate donor ATP binding site of RelQ in complex with a magnesium ion and the non-hydrolyzable ATP analog AMPCPP (PDB code 5F2V) with the equivalent amino acid positions in

PA14_01140$_{tox}$. Amino acid side chains deriving from PA14_01140$_{tox}$ or RelQ and their corresponding labels are shown in purple and yellow, respectively. **D)** Interaction with PA14_01130 distorts the predicted nucleotide acceptor site of PA14_01140$_{tox}$. Structural alignment between PA14_01140$_{tox}$-PA14_01130 complex and the (p)ppGpp synthetase RelP bound to the non-hydrolysable ATP analog AMPCPP and a GTP acceptor nucleotide (PDB code 6EWZ) (Manav et al., 2018). Two C-terminal α-helices of PA14_01140$_{tox}$ that align with the GTP binding site of RelP are rotated approximately 30° as a consequence of their interaction with PA14_01130 (black arrow). Colours corresponding to each protein model are indicated. **E)** Amino acid residues in PA14_01140$_{tox}$ that structurally align with known pyrophosphate donor ATP interacting residues in RelQ are required for PA14_01140$_{tox}$-mediated toxicity. Viability of *E. coli* cells grown on solid media harboring inducible plasmids expressing PA14_01140$_{tox}$, each of the indicated PA14_01140$_{tox}$ point mutants or an empty vector control. Lysine 326 is a residue located within the PA14_01140$_{tox}$ active site that is not predicted to interact with the pyrophosphate donor ATP. **F)** Glutamate 382 is required for PA14_01140-based intoxication of susceptible recipient cells during interbacterial competition. Outcome of intraspecific growth competitions between the indicated PA14 donor strains and a ΔPA14_01130-1140 recipient. The parental PA14 strain genotype is Δ*rsmA* Δ*rsmF.* The competitive index is normalized to starting ratios of donor/recipient. Data are mean ±SD. *P* value range from two-tailed, unpaired *t*-tests is shown.



m/z (Dalton)		[M-2H]2-	[M-3H+Na]2-	[M-H]-	[M-2H+Na]-
pApp	calc.	-	-	505.988	527.970
	found	-	-	505.992	527.973
ppApp	calc.	292.473	303.464	585.954	607.923
	found	292.476	303.460	585.959	607.928
pppApp	calc.	332.456	343.447	665.921	687.890
	found	332.456	343.444	665.926	687.895

Figure 3.5. Tas1$_{tox}$ synthesizes (p)ppApp by pyrophosphorylating the 3' hydroxyl group of adenosine nucleotides. A) PA14_01140$_{tox}$ catalyzes the formation of AMP in a dose-dependent manner. Coupled enzyme assay of PA14_01140$_{tox}$-catalyzed AMP production as a function of NADH consumption over time. **B)** PA14_01140$_{tox}$ catalyzes the production of AMP from ATP in a GTP-independent manner. The control reaction lacks adenylate kinase, which is required for the initial step of the coupled assay. **C)** PA14_01140$_{tox}$ (Tas1$_{tox}$) is a (p)ppApp synthetase enzyme. Anion-exchange traces of ATP alone or with excess AMP or ADP after incubation with Tas1$_{tox}$. A standard trace for ATP, ADP and AMP is shown for comparison. **D)** ^{1}H (top) and ^{31}P NMR (bottom) spectra of pApp. See Table 3.2 for assignments. **E)** Negative mode electrospray mass spectra for pApp (top), ppApp (middle) and pppApp (bottom). Assignment of major peaks is shown in the table. **A-C)** Data are representative of two independent experiments.

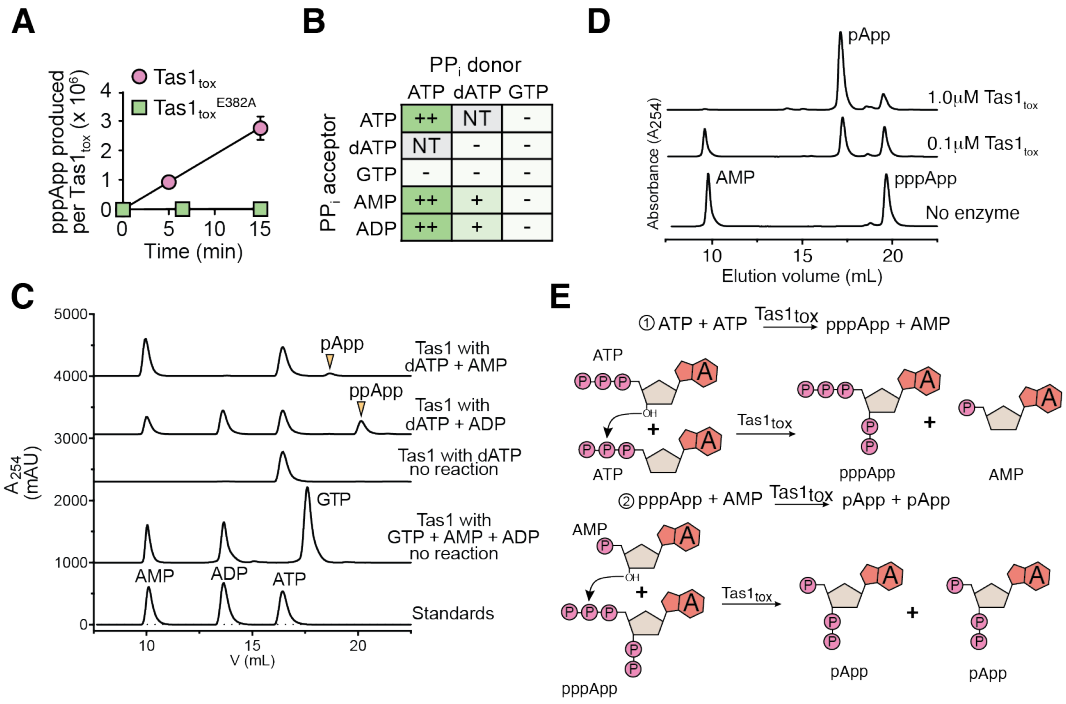


Figure 3.6. $Tas1_{tox}$ uses multiple adenosine donor nucleotides to rapidly synthesize (p)(p)pApp. A) Rate of pppApp production by $Tas1_{tox}$ or $Tas1_{tox}^{E382A}$. Reactions were performed at 37°C with 10 mM ATP and 1 nM $Tas1_{tox}$ or 1 μM $Tas1_{tox}^{E382A}$. Data are mean ±SD from three separate reactions. **B)** Specificity of $Tas1_{tox}$ towards pyrophosphate (PP_i) donors and acceptors. Indicated nucleotides (1 mM each) were incubated with 100 nM $Tas1_{tox}$ at room temperature for 10 minutes. Reactions that progressed to completion (++), made detectable product (+) or made no detectable product (-) are indicated. NT, not tested. **C)** Anion-exchange traces of $Tas1_{tox}$-catalyzed reactions with dATP or GTP as pyrophosphate donors. Arrowheads indicate 3′ pyrophosphorylation products. Purified $Tas1_{tox}$ can use pppApp as a pyrophosphate donor to pyrophosphorylate AMP resulting in pApp formation. **D)** Anion-exchange traces of pppApp and AMP after incubation with the indicated concentrations of $Tas1_{tox}$ for 30 min at room temperature. A control lacking $Tas1_{tox}$ is shown for comparison. **E)** Mechanism of quantitative conversion of ATP to pApp. Only heteroatoms that participate in the reaction mechanism of pApp formation are shown. **A-D)** Data are representative of two independent experiments.

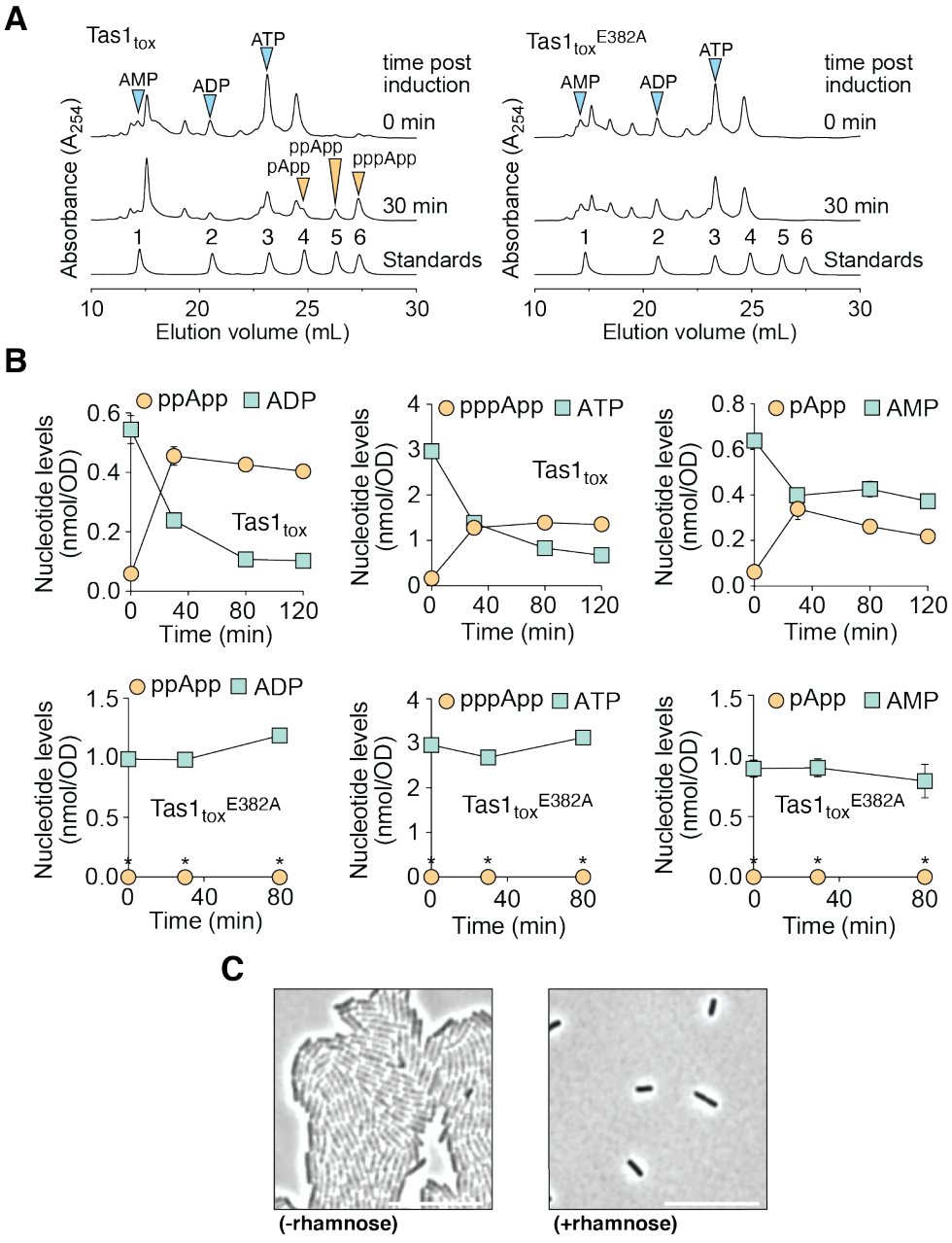


Figure 3.7. $Tas1_{tox}$ overexpression in *E. coli* leads to (p)(p)pApp accumulation and a reduction in cellular 5′ adenosine nucleotides. A) Anion exchange-chromatography traces of metabolites extracted from *E. coli* cells overexpressing $Tas1_{tox}$ (left) or $Tas1_{tox}^{E382A}$ (right) at the indicated time points. A trace generated from a mixture of standards containing an equimolar amount of AMP (1), ADP (2), ATP (3), pApp (4), ppApp (5) and pppApp (6) using the same gradient is shown for comparison. Peaks of adenosine 5′-nucleotides and (p)(p)pApp are indicated by blue and orange arrowheads, respectively. Traces are representative of three

biological replicates. **B)** Quantification of adenosine 5′-nucleotide and (p)(p)pApp levels in the *E. coli* strains from **A** as a function of time post induction. Data are mean ± SD for metabolites extracted from three separate cultures. Metabolites below the detection limit are indicated with an asterisk. **C)** Micrographs of *E. coli* harboring a rhamnose-inducible $Tas1_{tox}$ expression vector grown on an LB agarose pad with (right) or without (left) rhamnose after 6 hours. Scale bar is 10 μm.

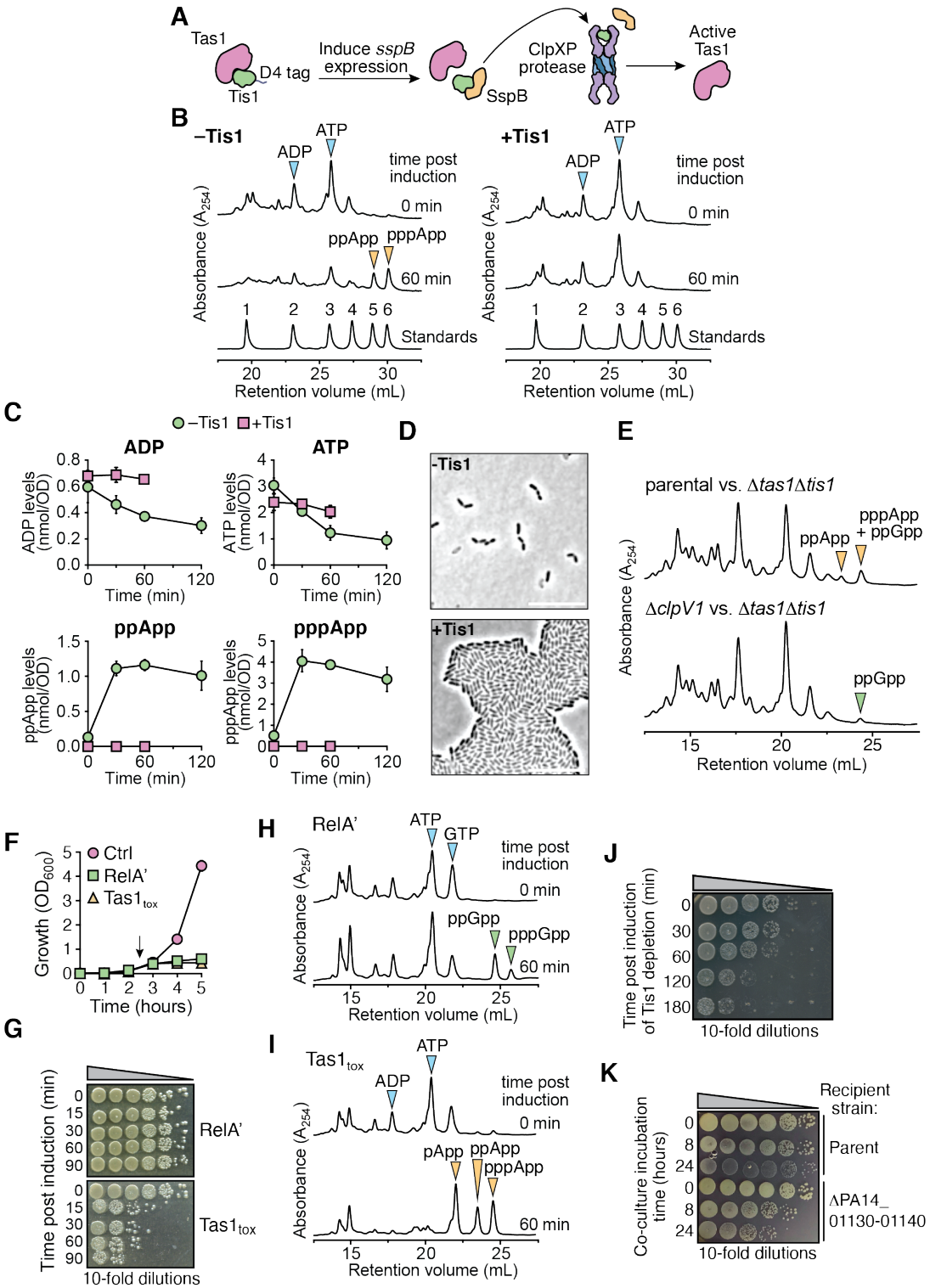


Figure 3.8. Tas1 intoxication depletes cellular ADP and ATP in *P. aeruginosa*. **A)** Schematic of the inducible Tis1 degradation system used to generate active

Tas1 in *P. aeruginosa* cells. Induction of *sspB* expression results in degradation of D4-tagged Tis1 by the ClpXP protease (McGinness et al., 2006). **B)** (p)ppApp accumulates in Tis1-depleted *P. aeruginosa* cells. Anion-exchange chromatography separated metabolites extracted from a *P. aeruginosa* PA14 parental strain (right, Δ*retS* Δ*sspB* pPSV39-CV::*sspB*) and a derivative expressing *tis1*-D4 (left, Δ*retS* Δ*sspB* PA14_01130-DAS+4 pPSV9-CV::*sspB*) before or 1-hour after induction of *sspB* expression. Blue and orange arrowheads indicate peaks of adenosine 5′-nucleotides and (p)ppApp, respectively. A standard trace of an equimolar mixture of 1-AMP, 2-ADP, 3-ATP, 4-pApp, 5-ppApp and 6-pppApp using the same gradient is shown for comparison. **C)** Absolute quantification of ADP, ATP and (p)ppApp levels in the *P. aeruginosa* strains from **A** as a function of time following induction of Tis1 depletion. **D)** Micrographs of *P. aeruginosa* expressing *tis1*-D4 pre-induction or 2 hours post-induction of *sspB* expression (as described in **B**) imaged following 6-hour growth on an LB agarose pad. **E)** Anion-exchange chromatography traces of metabolites extracted from growth competition experiments conducted on solid media for 2.5 hours. The parental strain is *P. aeruginosa* Δ*rsmA* Δ*rsmF*. **F)** Growth curves of *E. coli* cells expressing either the (p)ppGpp synthetase domain of RelA (RelA′), $Tas1_{tox}$ or a vector control (Ctrl). Arrow indicates time at which inducer was added to cultures. **G)** CFU plating of *E. coli* cells expressing the plasmids defined in **F**. Cells were plated either pre-induction (0 min) or at the indicated times post-induction on inducer-free agar. Representative CFU plates from three biological replicates are shown. **H-I)** Anion-exchange chromatography traces of metabolites extracted from *E. coli* expressing RelA' **(H)** or $Tas1_{tox}$ **(I)** either pre-induction or 1-hour post-induction. Arrows indicate relevant metabolites isolated from culture. **J)** Tis1-depleted cells exhibited a reduction in viability over time. CFU plating of *P. aeruginosa* PA14 Δ*retS* Δ*sspB* Tis1-D4 pPSV39::*sspB* cells at the indicated time points post induction of SspB expression. **K)** Tas1 reduces the viability of susceptible recipient cells during interbacterial competition. CFU plating of the indicated *P. aeruginosa* PA14 recipient strains after co-culture with a parental donor strain at the indicated time points. The parental PA14 strain genotype is Δ*rsmA* Δ*rsmF.*

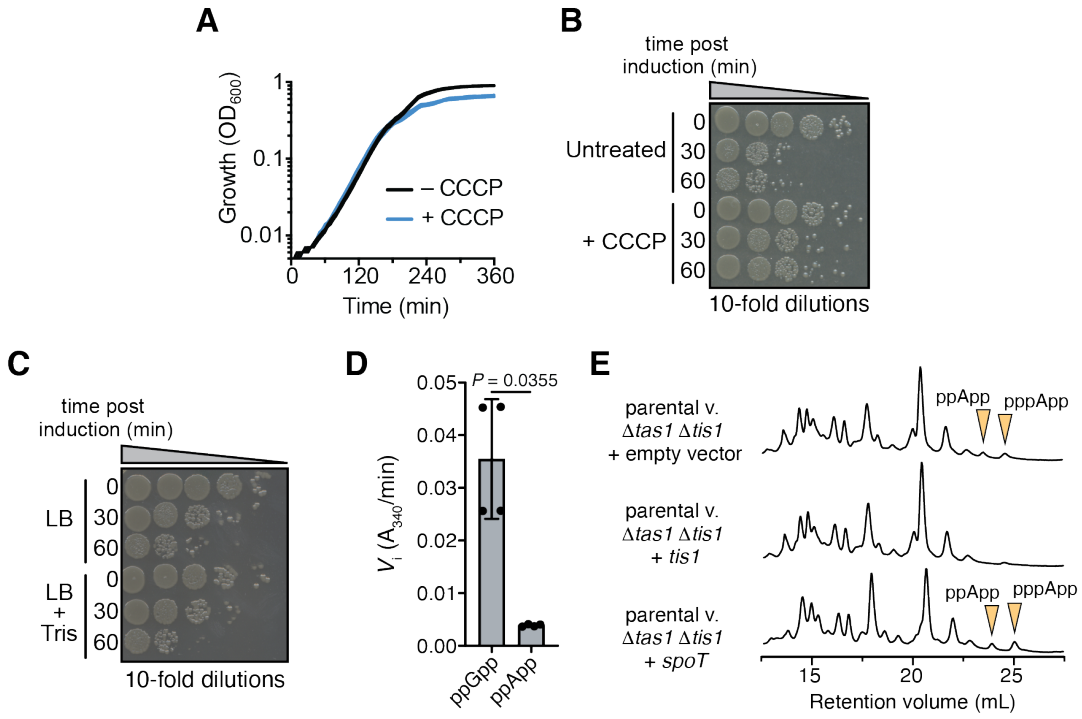


Figure 3.9. The pmf uncoupling ionophore CCCP but not the ppGpp-hydrolase domain of SpoT reduces the toxicity of Tas1$_{tox}$. A) Steady-state growth of *E. coli* is not substantially affected by the presence of carbonyl cyanide m-chlorophenyl hydrazine (CCCP). Growth curves of *E. coli* cells harboring the Tas1$_{tox}$ expression plasmid in LB medium with or without (CCCP). Results from three independent cultures were overlaid for each medium condition. **B)** Tas1$_{tox}$ toxicity is reduced in the presence of CCCP. Viability of *E. coli* cells following Tas1$_{tox}$ expression in the presence or absence of CCCP. Cells were plated either pre-induction or at the indicated times post-induction. **C**) Alkaline pH does not affect the ability of CCCP to reduce Tas1$_{tox}$-dependent toxicity indicating that the toxicity of Tas1$_{tox}$ likely arises from the generation of excessive membrane electrostatic potential. Cultures were untreated or conditioned to pH 8.0 using 25 mM Tris-HCl buffer immediately prior to induction. **D)** Activity of the ppGpp-hydrolase domain of SpoT against either ppGpp or ppApp. Initial velocities were normalized to hydrolase activity in the absence of either nucleotide. Data are mean ± SD for enzymatic activity from four technical replicates. *P* value from two-tailed, unpaired *t*-test is shown. **E)** Anion-exchange chromatography traces of metabolites extracted from growth competition experiments between the indicated strains conducted on solid media for 4 hours. The parental strain is Δ*rsmA* Δ*rsmF*. Traces are representative of three biological replicates. **B-C)** Representative plates of three biological replicates are shown.

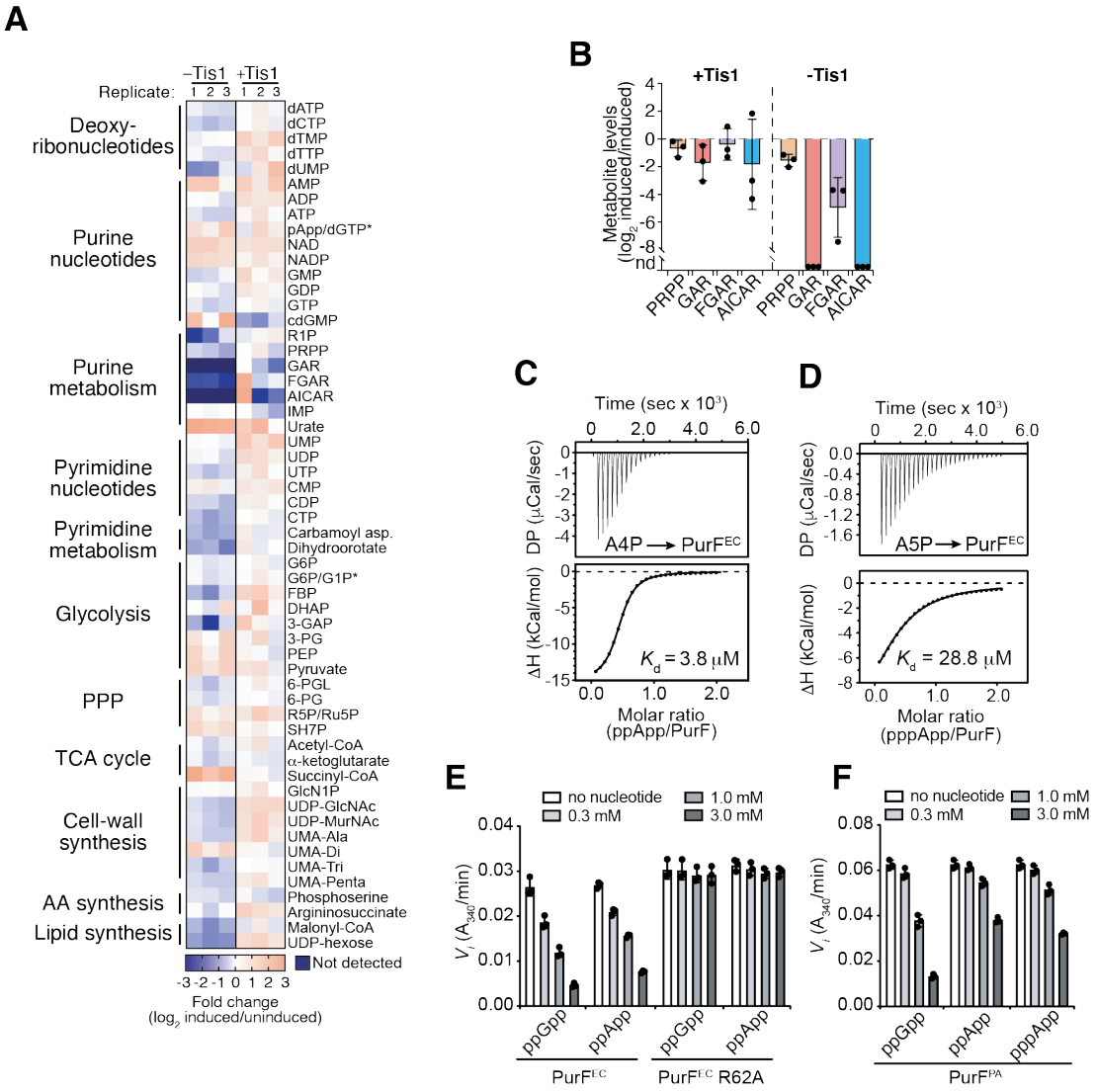


Figure 3.10. Rapid production of (p)ppApp by Tas1 causes dysregulation of central metabolism and leads to the direct inhibition of purine biosynthesis. **A)** Relative levels of metabolites from *P. aeruginosa* containing or lacking Tis1 (details about depletion system are provided in Figure 3.8A and in Methods). Heat map shows metabolite levels calculated for both the +Tis1 and -Tis1 strains as a $\log_2$ ratio for samples 1-hour post-induction relative to pre-induction of *sspB* expression. The asterisk indicates metabolites that are indistinguishable in the LC-MS analysis. The metabolic pathway or classification for each metabolite is shown. Data for three biological replicates are shown. **B)** Relative quantification of metabolites within the de novo purine biosynthesis pathway in *P. aeruginosa* strains containing or lacking Tis1 from **A**. Metabolite levels for both the +Tis1 and -Tis1 strains are shown as $\log_2$ ratios for samples 1-hour post-induction relative to pre-induction of *sspB* expression. nd, not detected. **C-D)** Isothermal calorimetry

traces (top) and fitted isotherms (bottom) for the titration of 100 μM $PurF^{EC}$ (monomer) with 1 mM ppApp **(C)** or pppApp **(D)**. Representative traces from two independent replicates are shown. **E-F)** Changes to the activity of $PurF^{EC}$ or a ppGpp-blind $PurF^{EC}$ variant **(E)** or $PurF^{PA}$ **(F)** in the presence of indicated concentrations of ppGpp or (p)ppApp (Wang et al., 2019). Data are mean ±SD for three reactions.

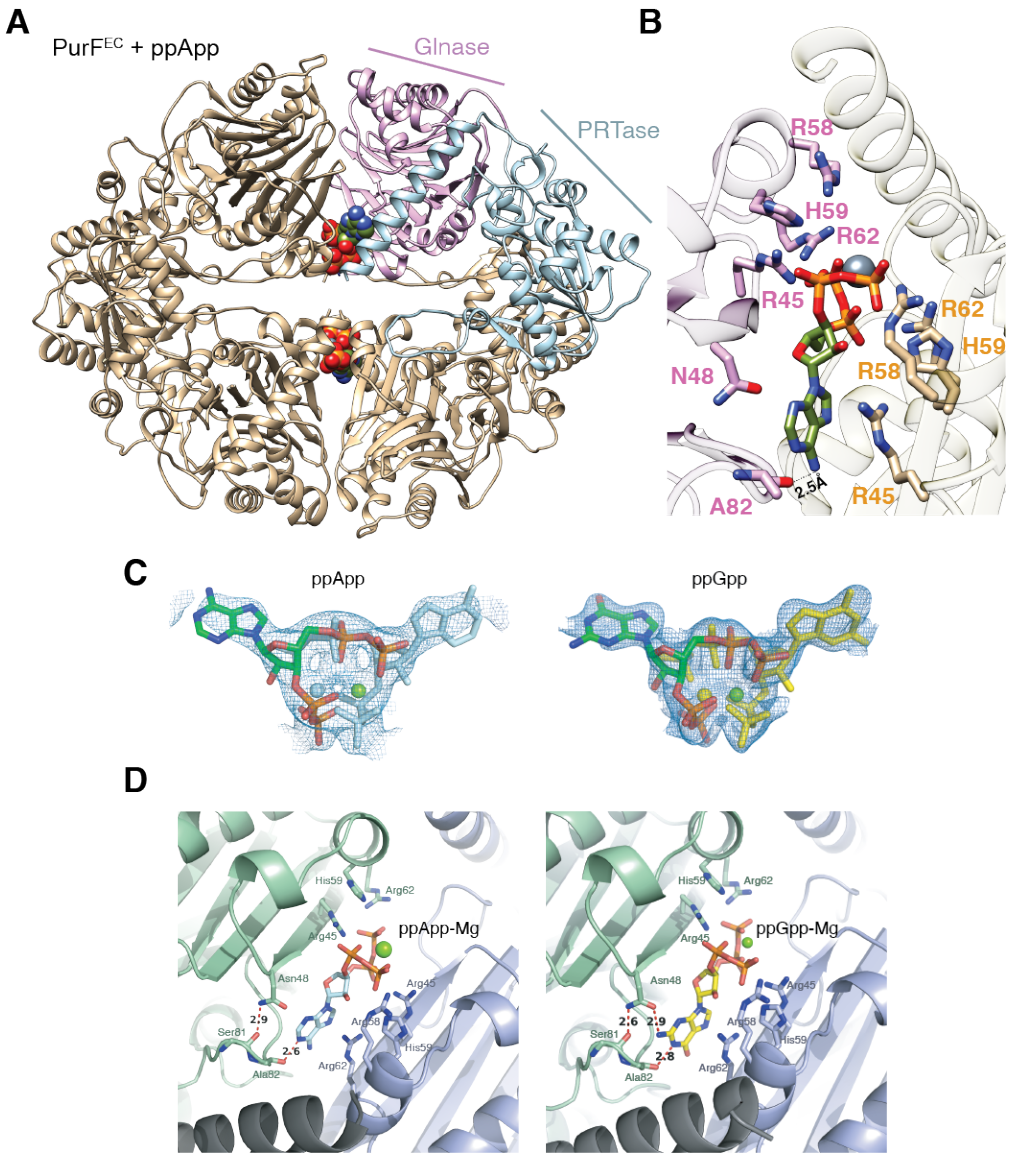


Figure 3.11. ppApp inhibits PurF in a manner similar to ppGpp. A) Ribbon diagram of the PurF^{EC} tetramer in complex with ppApp. A single PurF subunit is coloured by individual domains (GInase; glutaminase domain in pink, PRTase; phosphoribosyltransferase domain in blue), while the remaining subunits are coloured brown. ppApp and Mg^{2+} are shown in stick and sphere representations, respectively. **B)** Close-up view of the ppApp binding site between the glutaminase domains of two adjacent PurF^{EC} monomers. ppApp-interacting residues are shown as pink or orange sticks and hydrogen bonding between PurF^{EC} and the purine ring of ppApp are shown as black dashed lines. **C)** $2F_o$-F_c difference electron density maps of ppApp (left) and ppGpp (right, PDB code 6CZF) contoured at 0.4σ are shown in blue. Nucleotides are shown as stick models of two overlapping ppApp-

Mg^{2+} (coloured by heteroatom or light blue) or ppGpp-Mg^{2+} (coloured by heteroatom or yellow), related by a two-fold rotational axis. **D)** Comparison of ppGpp and ppApp binding configuration within $PurF^{EC}$. The nucleotide-Mg^{2+} complexes are modeled at 0.5 occupancy because they lie on a crystallographic two-fold rotational axis as shown in **C**. Relevant hydrogen bonding interactions and their distance in angstroms between $PurF^{EC}$ residues and the purine rings of ppApp (left) or ppGpp (right) are shown with red dashed line.

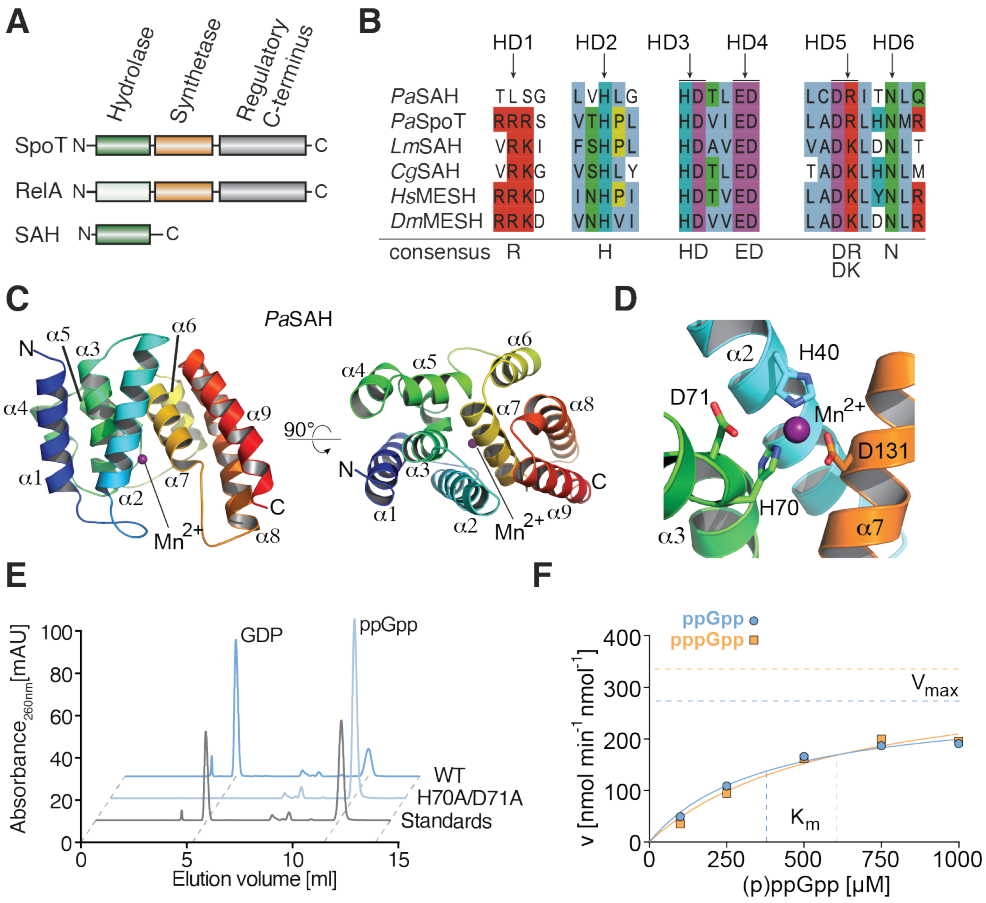


Figure 3.12. *Pa*SAH resembles known small alarmone hydrolases and hydrolyzes (p)ppGpp. A) Comparison of domain architectures for SpoT, RelA and general small alarmone hydrolase (SAH) enzymes. **B)** Multiple sequence alignment of *Pa*SAH with known (p)ppGpp hydrolases: *Drosophila melanogaster* (*Ds*) MESH1 (Sun et al., 2010), *Homo sapiens* (*Hs*) MESH1 (Sun et al., 2010), *Corynebacterium glutamicum* (*Cg*) SAH (Ruwe et al., 2018), *Listeria monocytogenes* (*Lm*) SAH (Lmo0812) (PDB: 4YF1), SpoT from *P. aeruginosa* (*Pa*SpoT). Consensus motif indicates residues with that are required for (p)ppGpp hydrolysis. **C)** Ribbon diagram of *Pa*SAH structure. Mn^{2+} ion is shown as a purple sphere. **D)** Close-up view of Mn^{2+} coordinating motif within the active site of *Pa*SAH. **E)** Chromatogram of an enzymatic reaction containing ppGpp and native *Pa*SAH (WT, dark blue trace) or a variant unable to coordinate manganese (H70A/D71A, light blue trace). GDP and ppGpp (gray trace) were used as standard. **F)** Velocity (v)/substrate curves of ppGpp (blue circles) and pppGpp (orange squares) hydrolysis by *Pa*SAH in presence of 2.5 mM $MnCl_2$. Data were fitted according to the Michaelis–Menten equation. Vertical and horizontal dashed lines indicate the Michaelis–Menten constants (K_m) and maximal velocities (V_{max}),

respectively. K_m (in μM) and V_{max} (in nmol/min $nmol^{-1}$) were: ppGpp (K_m: 378 ± 84 and V_{max}: 275 ± 24) and pppGpp (K_m: 604 ± 236 and V_{max}: 336 ± 65).

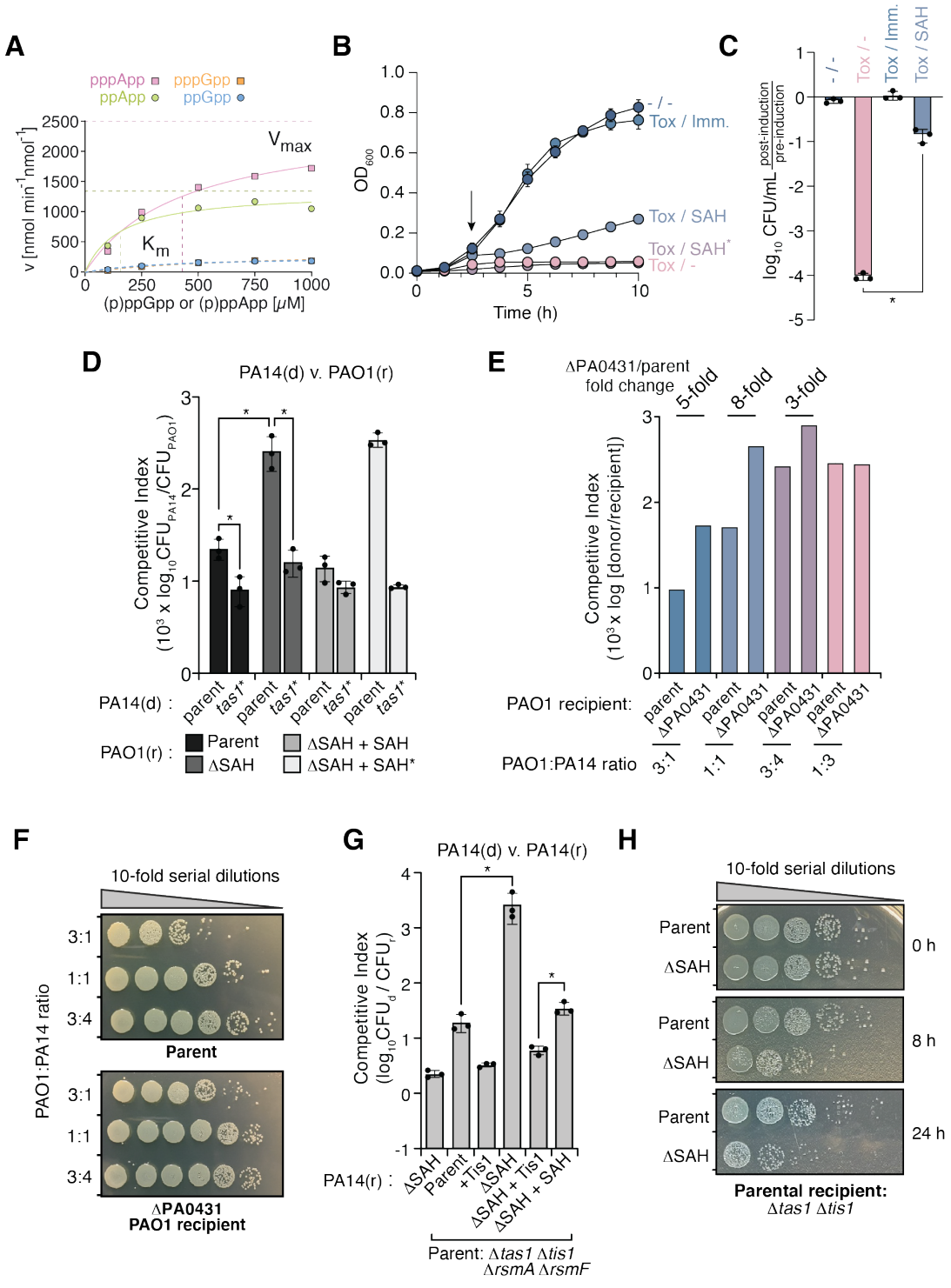


Figure 3.13. *Pa*SAH degrades (p)ppApp and confers protection to Tas1 during interbacterial competition. (a) Velocity (v)/substrate (S) curve of ppApp (green circles) and pppApp (pink squares) hydrolysis by *Pa*SAH in presence of 2.5

mM $MnCl_2$. Data were fitted according to the Michaelis–Menten equation. Vertical and horizontal dashed lines indicate the Michaelis–Menten constants (K_m) and maximal velocities (V_{max}), respectively. K_m (in μM) and V_{max} (in nmol/min $nmol^{-1}$) were: ppApp (K_m: 155 ± 58 and V_{max}: 1,336 ± 132) and pppApp (K_m: 429 ± 112 and V_{max}: 2,507 ± 274). The velocities for hydrolysis of ppGpp (blue circles) and pppGpp (orange squares) are shown for comparison (See Figure 3.12F). **B)** Growth of *E. coli* in liquid culture expressing the indicated proteins. The arrow indicates the addition of inducer after 2.5 hours. Data represent mean ± SD of *n* = 3 replicates. **C)** Ratio of *E. coli* colony forming units per millilitre (CFU/mL) 60 minutes postinduction to CFU/mL preinduction of the constructs used in **B**. CFU/mL was determined from LB agar plates lacking inducers. **D)** Growth competition assay between *P. aeruginosa* PA14 donor (d) and PAO1 recipient (r) strains. PA14 harboring an inactive Tas1 (E382A variant, denoted *tas1**) was used as a control. The parental backgrounds are Δ*retS* and Δ*rsmA* Δ*rsmF* for PAO1 and PA14, respectively; both of mutations stimulate H1-T6SS activity (Goodman et al., 2004; Marden et al., 2013). The PAO1 parent had a significant T6SS-independent growth advantage over PA14, and therefore, all ratios for the competitive index were multiplied by 10^3 to better visualize the changes in PA14 CFU. **E)** The outcome of a growth competition assay between PAO1 and PA14 is dependent on initial ratio. Growth competition assay between the indicated *P. aeruginosa* PAO1 recipient strains against a PA14 donor strain. The strains were mixed at the indicated ratios of PAO1:PA14 and incubated for 9 h. The parental strains are Δ*retS* and Δ*rsmA* Δ*rsmF* for PAO1 and PA14, respectively. PAO1 readily outcompetes PA14 in this assay and so all data were multiplied by 10^3 for visualization purposes (similar to **E**). Data are representative of *n* = 1 replicate. **F)** Representative plates showing CFU changes in PA14 following competition shown in panel **E**. The PA14 donor contains a *lacZ* reporter and a tetracycline-resistance cassette for blue-white screening and selection. **G)** Growth competition assay between *P. aeruginosa* PA14 donor and recipient strains. The donor strain genotype is Δ*rsmA* Δ*rsmF*. All recipient strains, except the ΔSAH strain, were generated from a Δ*tas1* Δ*tis1* Δ*rsmA* Δ*rsmF* parental strain. The competitive index denotes the $\log_{10}$ of the donor/recipient CFU ratio after co-cultivation of both strains for 24 hours. **H)** Viability of *P. aeruginosa* PA14 recipient strains after co-cultivation with a PA14 donor at the indicated times. The PA14 donor genotype is Δ*rsmA* Δ*rsmF*. The recipient strains (underlined in panel **G**) were generated from a Δ*tas1* Δ*tis1* Δ*rsmA* Δ*rsmF* parental strain. **C-D, G)** Data represent mean ± SD of *n* = 3 biological replicates. Asterisks indicates pairs that are significantly different.

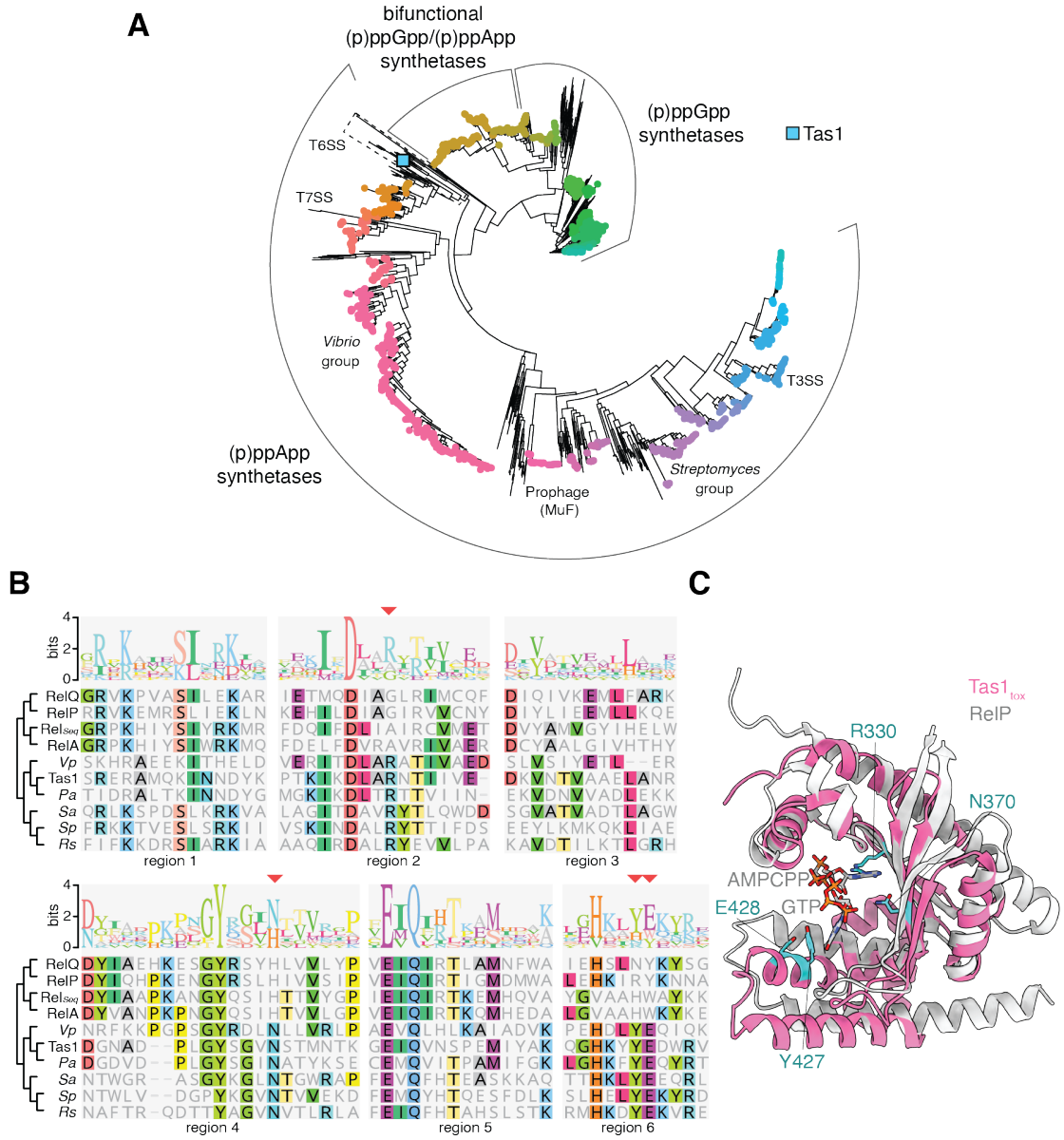


Figure 3.14. Hundreds of predicted (p)ppApp synthetases are not associated with type VI secretion systems. A) Phylogenetic distribution of 2,079 $Tas1_{tox}$ homologs identified in UniprotKB. Similarly coloured dots indicate sequences with high sequence similarity. Dots of the same shade exhibit greater than 25% sequence similarity. Outer rings broadly encompass regions with predicted (p)ppGpp synthetases, bifunctional (p)ppGpp/(p)ppApp synthetase and (p)ppApp synthetases. The $Tas1_{tox}$ sequence is indicated by the blue square. **B)** Multiple sequence alignment and associated sequence logo of $Tas1_{tox}$, predicted (p)ppApp synthetases from *Vibrio parahaemolyticus* (*Vp*, *Vibrio* group), *Pseudomonas aeruginosa* (*Pa*), *Streptomyces albidoflavus* (*Sa*, *Streptomyces* group), *Ralstonia solanacearum* (*Rs*, T3SS-associated) and previously characterized (p)ppGpp

synthetases RelA from *E. coli*, RelQ from *B. subtilis* (Steinchen et al., 2015), RelP *S. aureus* (Manav et al., 2018) and Rel_{Seq} from *Streptococcus equisimilis* (Hogg et al., 2004). Residues that are conserved in predicted (p)ppApp synthetases and $Tas1_{tox}$, but not enzymes that synthesize (p)ppGpp, are indicated by the red arrows. Six regions with greatest number of conserved residues are shown. A phylogenetic tree is shown to the left of the sequences to depict evolutionary relationship between each protein. **C)** Structural alignment of $Tas1_{tox}$ (pink) and RelP in complex with GTP and non-hydrolyzable ATP (AMPCPP) (PDB: 6EWZ). Residues that are conserved in predicted (p)ppApp synthetases (shown in panel **B**) are mapped to the $Tas1_{tox}$ structure and shown in cyan.

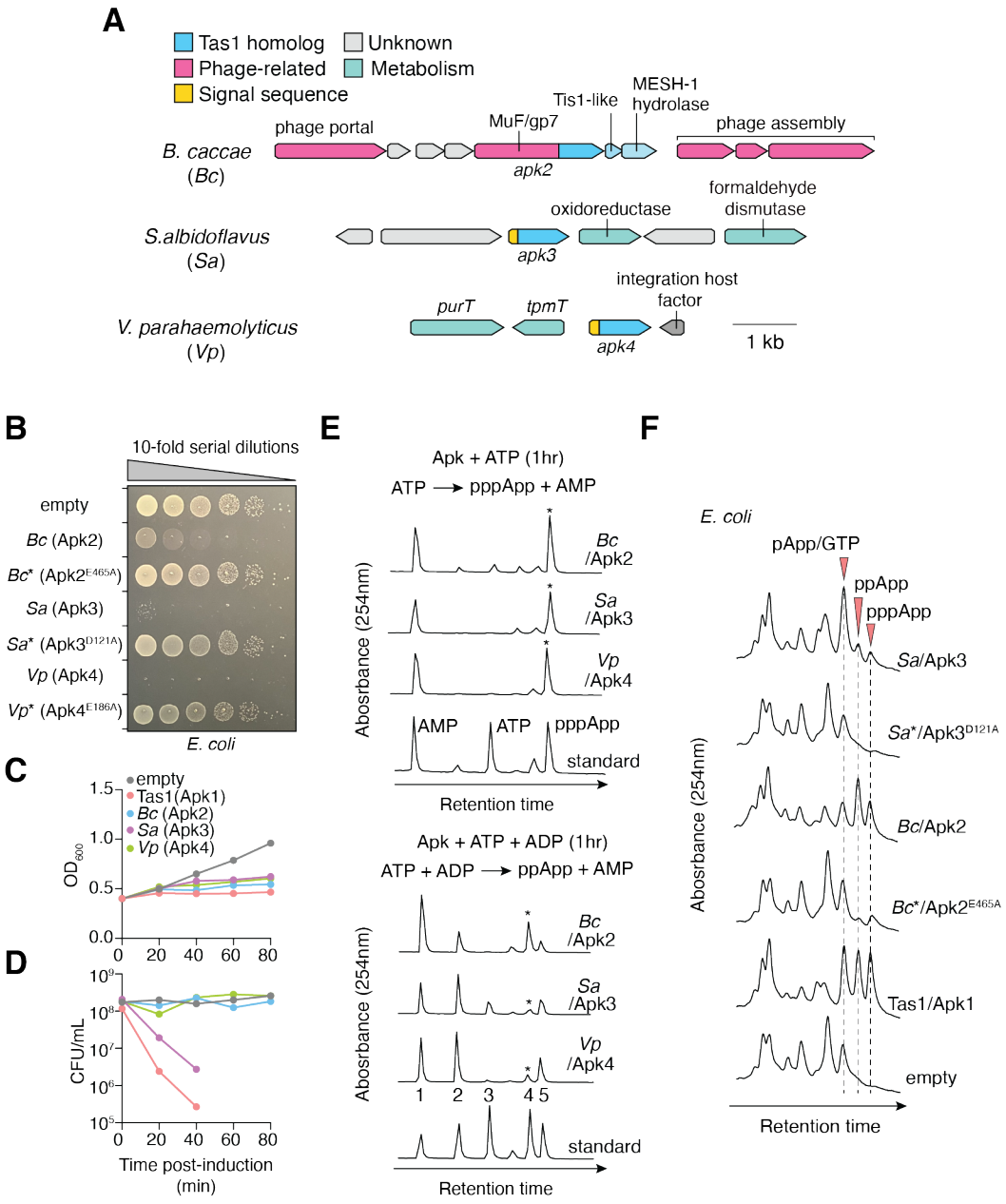


Figure 3.15. Tas1 homologs are adenosine 3'-pyrophokinases (Apk) that synthesize (p)(p)pApp. A) Genomic context of three Tas1 homologs selected for downstream experimentation. Regions that are homologous to $Tas1_{tox}$ are shown in blue, genes associated with phage are shown in pink, predicted signal sequences are shown in yellow, genes associated with metabolism are shown in aquamarine and genes without any predicted function are shown in gray. *Bacteroides caccae* (*Bc*) contains Apk2, *Streptomyces albdioflavus (Sa*) contains Apk3, *V. parahaemolyticus* (*Vp*) contains Apk4. Scale bar indicates 1 kilobase pairs. **B)** Overnight growth of *E. coli* expressing either empty vector or a vector

containing the indicated homolog or a catalytically inactive mutant (denoted with asterisk) from the listed organism on inducer-containing LB agar. The blue regions in each *apk* homolog in panel **A** indicate the region that was used for expression (predicted Apk2 toxin domain and the signal sequence lacking variants of Apk3 and Apk4). **C-D)** Optical density (OD_{600}, **C**) measurements and colony-forming unit per millilitre enumeration (CFU/mL**, D**) of *E. coli* expressing the constructs from **B** before and at the timepoints shown after induction. **E)** Anion-exchange traces of 50 nM Apk2-4 enzymes incubated with 1 mM ATP (top) or ATP and ADP (bottom) for 1 hour. Standard traces are shown for comparison. Bottom standards: 1) AMP, 2) ADP, 3) ATP, 4) ppApp, 5) pppApp. Asterisks indicates pppApp or ppApp production. Overall reactions are shown for both traces. **F)** Anion-exchange traces of metabolites extracted from *E. coli* expressing the constructs from **B** 30 minutes after induction. Red arrows and dotted lines indicate peaks corresponding to pppApp, ppApp and pApp/GTP.

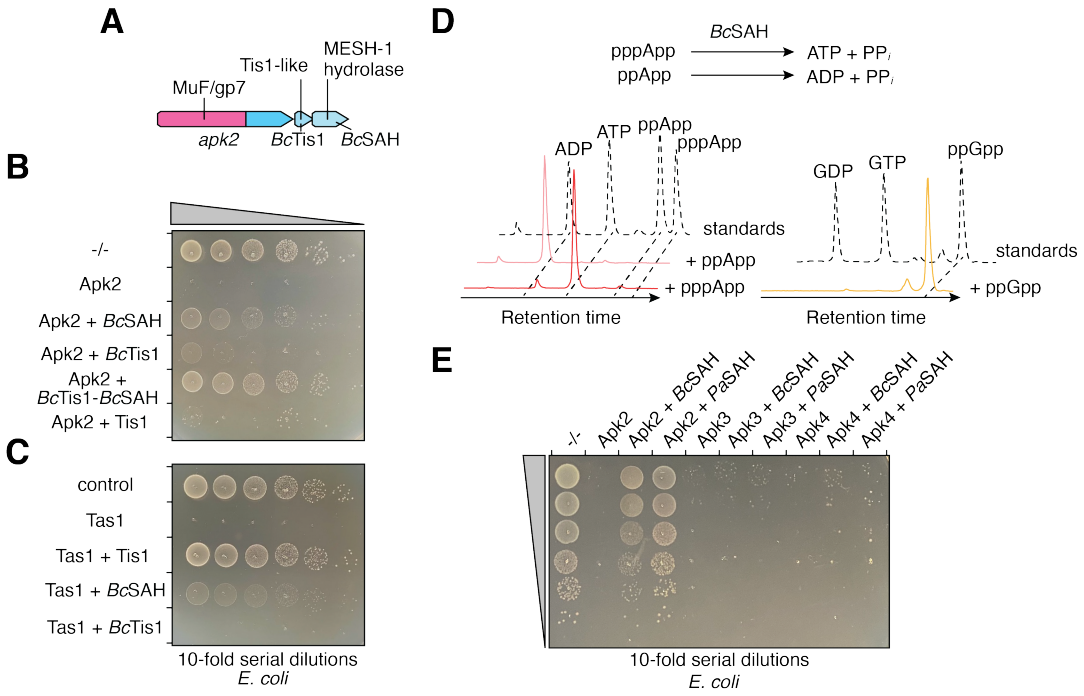


Figure 3.16. *Bc*SAH is a monofunctional (p)ppApp hydrolase. A) The *apk2* operon *B. caccae* (*Bc*) contains a predicted Tis1 like protein (*Bc*Tis1) and a predicted MESH-1 (p)ppGpp hydrolase (*Bc*SAH). The blue region indicates the predicted toxin domain with high homology to Tas1 that was used for downstream experiments. **B-C)** Overnight growth of *E. coli* expressing the indicated constructs on inducer-containing LB agar. "-/-" and "control" labels indicate cells expressing empty vectors. **D)** Anion-exchange traces of 500 nM *Bc*SAH incubated with 1 mM (p)ppApp (left, pppApp red line, ppApp light red line) or 1 mM ppGpp (right, yellow line) at 37°C for 1 hour. Overall reaction catalyzed by *Bc*SAH is shown above traces. Standards (dotted lines) contain (p)ppApp, ATP, ADP (left) or ppGpp, GTP, GDP (right). **E)** Same assay as that shown in **B** and **C**, except with the indicated constructs.

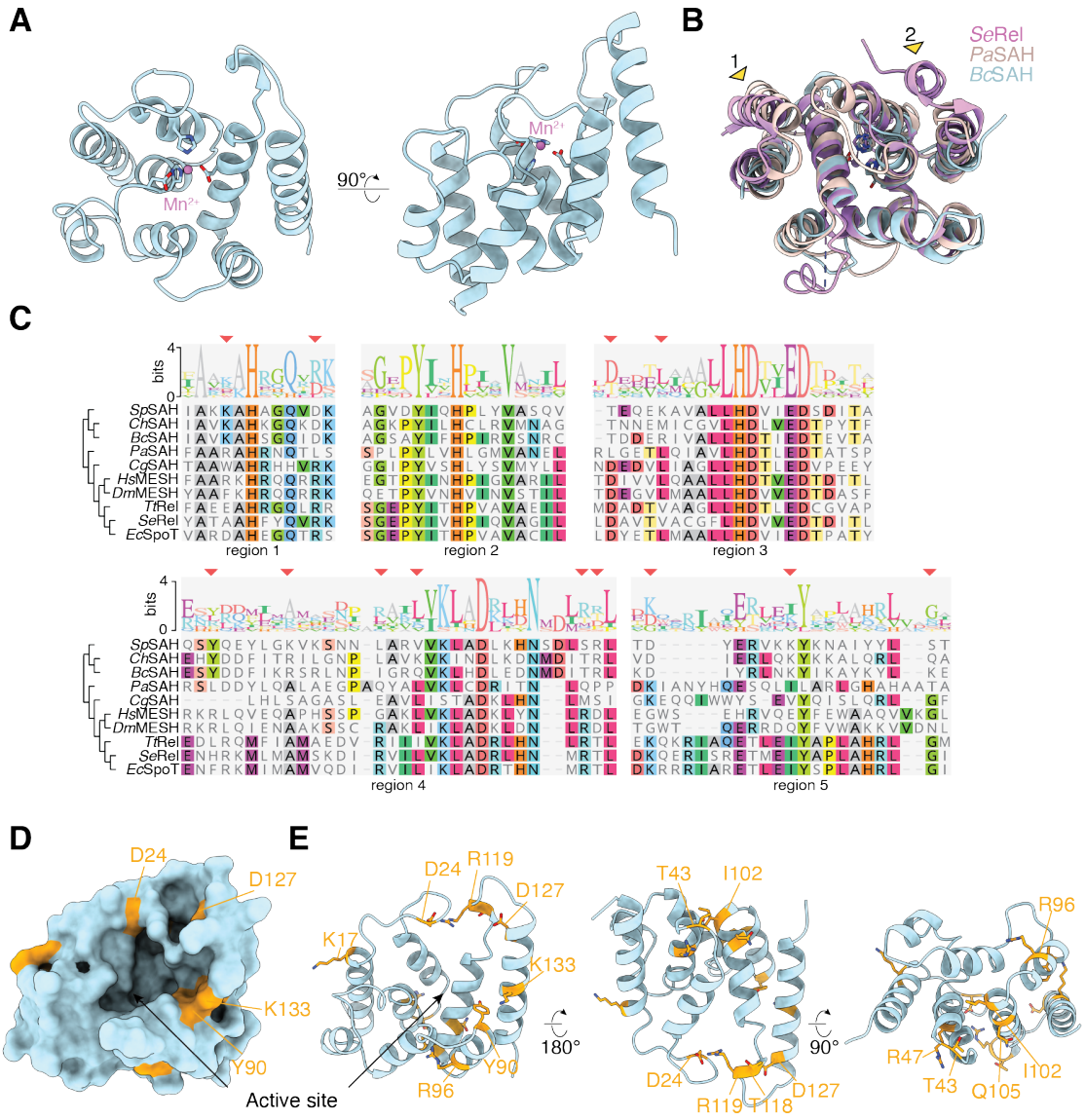


Figure 3.17. *Bc*SAH is structurally distinct from (p)ppGpp hydrolases. A) Overall structure of *Bc*SAH shown in ribbon representation. Manganese ion is shown in purple. **B)** Structural overlay of *Bc*SAH (blue) structure from **A** with *Pa*SAH (beige) structure (from Figure 3.12) and *Se*Rel (pink), a previously characterized (p)ppGpp synthetase/hydrolase from *Streptococcus equisimilis* (PDB: 1VJ7) (Hogg et al., 2004). Yellow arrows indicate helices missing in the *Bc*SAH structure. Arrow 1 points to a helix present in both *Pa*SAH and *Se*Rel, while arrow 2 points to a helix present only in *Se*Rel. **C)** Multiple sequence alignment and sequence logo of *Bc*SAH with previously characterized (p)ppGpp hydrolases: SpoT from *E. coli* (*Ec*SpoT) (Gentry and Cashel, 1996; Sarubbi et al., 1989), *Se*Rel (see description for panel **B**), Rel from *Thermus thermophilus* (*Tt*Rel) (Tamman et

al., 2020), MESH-1 from *Drosophila melanogaster* (*Dm*MESH), MESH-1 from *Homo sapiens* (*Hs*MESH) (Sun et al., 2010), SAH from *Corynebacterium glutamicum* (*Cg*SAH) (Ruwe et al., 2018), the bifunctional (p)ppGpp/(p)ppApp hydrolase *Pa*SAH (Figure 3.12) and two closely related homologs of *BcSAH* that are predicted (p)ppApp hydrolases identified manually: *Ch*SAH from *Capnocytophaga haemolytica* (*Ch*SAH) and *Sp*SAH from *Streptococcus pneumoniae*. Phylogenetic tree (left of alignment) shows the evolutionary relationship between each protein. Red arrows show conserved residues within *Bc*SAH and the predicted (p)ppApp hydrolases *Ch*SAH and *Sp*SAH, but not enzymes that hydrolyze (p)ppGpp. Five regions with highly conserved residues are shown. **D-E)** Surface **(D)** and ribbon **(E)** representations of *Bc*SAH structure with the conserved residues from **C** mapped in orange.

Tables

Table 3.1. X-ray data collection, phase and structure refinement statistics for the crystal structures presented in chapter III.

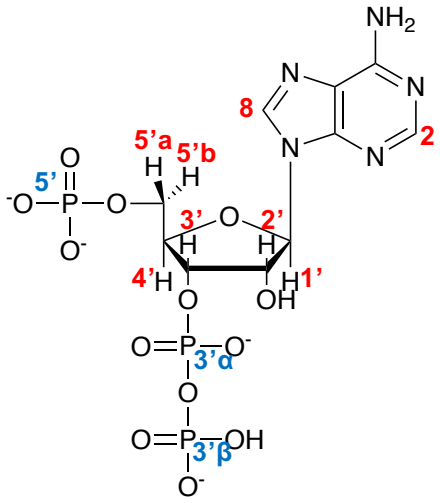
	Tas1$_{tox}$-Tis1 complex (6OX6)	**PurF^{EC}-ppApp complex** (6OTT)	***Pa*SAH** (6YVC)	***Bc*SAH**
Data Collection				
Space group	$P4_32_12$	$C222_1$	*C2*	$P6_1$
Cell dimensions				
a, b, c (Å)	66.4, 66.4, 147.9	115.2, 156.8, 107.5	85.33 85.38 122.17	144.1, 144.1, 63.5
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90 102.059 90	90.0, 90.0, 120.0
Wavelength (Å)	0.979	0.979	1.072	0.979
Resolution (Å)	50.00 – 2.17	47.6-2.55	45.78–1.85	47.64 – 2.3
R_{merge}[a]	0.083 (0.243)*	0.106 (1.27)	0.147 (0.622)	0.136 (1.42)
I / σ(I)	33.4 (6.2)	9.0 (1.0)	8.09 (1.92)	23.2 (1.0)
Completeness (%)	98.0 (96.4)	99.8 (100)	99.13 (99.08)	99.6 (93.2)
Redundancy	9.2 (4.1)	5.8 (6.2)	3.7 (3.8)	14.1 (7.1)
Refinement				
Resolution (Å)	49.6 – 2.17	47.6 – 2.55	48.75–1.85	47.48 – 2.3
No. reflections				
Total	18200	186253	72548 (7,192)	33470 (1573)
Unique	17596	32044	-	33668
Free (%)	5.00	6.23	-	5.00
R_{work} / R_{free} (%)[b]	18.5/23.1	22.3/26.1	0.19/0.22	49.6/51.0
No. atoms				
Protein	2119	15508	5632	1222
Ligand/ion	28	28	4	5
Water	316	0	896	144
B-factors (Å^2)				
Protein	40.8	74.6	25.69	62.5
Ligand/ion	72.0	61	28.81	62.3
Water	53.7	-	35.01	52.8
r.m.s deviations				
Bond lengths (Å)	0.006	0.02	0.004	0.003
Bond angles (°)	0.993	0.547	0.90	0.5

Single crystals were used to collect data for each structure. *values in brackets value refer to highest resolution shells.

[a] $R_{merge} = \Sigma_{hkl}\Sigma_j|I_{hkl,j} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\Sigma_j I_{hk,j}$, where $I_{hkl,j}$ and $\langle I_{hkl}\rangle$ are the jth and mean measurement of the intensity of reflection j.

[b] $R = \Sigma|F_p^{obs} - F_p^{calc}|/\Sigma F_p^{obs}$, where F_p^{obs} and F_p^{calc} are the observed and calculated structure factor amplitudes, respectively.

Table 3.2. ^{1}H and ^{31}P NMR assignments for pApp. The chemical shifts (in ppm) for pApp. α-proton of the pyrophosphate group was identified from its ddd pattern with a characteristic J-coupling constant with the proximal ^{31}P nuclei at 8.5Hz (0.014 ppm for protons or 0.036ppm for ^{31}P). This proton was assigned to the 3′ location by comparing proton chemical shifts to those observed for adenosine-5′-phosphate (AMP, BMRB database ID: BMSE000837) and a chemically synthesized adenosine-3′, 5′-bisdiphosphate (ppApp) under the same condition (Haas et al., 2019). Note that compared to a free hydroxyl group, pyrophosphorylation increases the chemical shift of the α-proton by 0.3-0.4 ppm.



		Base (^{1}H)		Ribose (^{1}H)					Phosphate (^{31}P)		
	Position	2	8	1′	2′	3′	4′	5′a / 5′b	5′α	3′α	3′β
pApp	Chemical Shift (ppm)	8.15	8.59	6.16	4.88	4.93	4.55	4.02	3.75	-10.0	-5.1
	Integration (multiplicity)	1H (s)	1H (s)	1H (d)	1H (dd)	1H (ddd)	1H (m)	2H (m)	1P (s)	1P (dd)	1P (d)
ppApp	Chemical Shift (ppm)	8.19	8.48	6.12	4.79	4.90	4.54	4.16	-	-	-
	Integration (multiplicity)	1H (s)	1H (s)	1H (d)	1H (dd)	1H (ddd)	1H (m)	2H (m)	-	-	-
AMP	Chemical Shift (ppm)	8.13	8.48	6.09	4.75	4.51	4.38	4.10	-	-	-
	Integration (multiplicity)	1H (s)	1H (s)	1H (d)	1H (dd)	1H (dd)	1H (m)	2H (m)	-	-	-

* *"-" indicates assignment not clear or data not available.*

Table 3.3. Relative metabolite concentrations of *P. aeruginosa* cells undergoing Tas1-mediated intoxication (*n* = 3 for Tis1-depleted and control cultures). The .xls file can be accessed at:

https://static-content.springer.com/esm/art%3A10.1038%2Fs41586-019-1735-9/MediaObjects/41586_2019_1735_MOESM4_ESM.xlsx

Table 3.4. Absolute nucleotide quantification of *P. aeruginosa* cells undergoing Tas1-mediated intoxication.

Metabolite	Concentration in Tis1 depleted cells (nmol/OD)				Concentration in Tis1 expressing cells (nmol/OD)		
	time (min)				time (min)		
	0	30	60	120	0	30	60
AMP							
R1	0.179	0.193	0.238	0.315	0.253	0.323	0.317
R2	0.151	0.172	0.207	0.236	0.369	0.326	0.342
R3	0.218	0.213	0.170	0.284	0.365	0.350	0.490
ADP							
R1	0.547	0.441	0.387	0.330	0.631	0.747	0.652
R2	0.507	0.407	0.379	0.232	0.700	0.634	0.644
R3	0.731	0.537	0.340	0.346	0.707	0.681	0.667
ATP							
R1	2.716	2.091	1.462	0.939	2.648	2.484	2.096
R2	3.212	2.054	1.299	0.614	2.042	2.281	2.000
R3	3.191	1.992	1.264	0.916	2.495	2.221	1.641
ppApp							
R1	0.128	1.020	1.186	1.154	n.d.	n.d.	n.d.
R2	0.126	1.095	1.073	0.773	n.d.	n.d.	n.d.
R3	0.154	1.223	1.105	1.226	n.d.	n.d.	n.d.
pppApp							
R1	0.355	3.588	3.964	3.768	n.d.	n.d.	n.d.
R2	0.576	3.986	3.763	2.609	n.d.	n.d.	n.d.
R3	0.588	4.628	3.184	3.909	n.d.	n.d.	n.d.
GMP							
R1	0.077	0.021	0.032	0.034	0.072	0.093	0.080
R2	0.048	0.019	0.022	0.003	0.090	0.094	0.068
R3	0.050	0.074	0.031	0.048	0.098	0.076	0.091
GDP							
R1	0.132	0.113	0.096	0.102	0.197	0.233	0.184
R2	0.141	0.115	0.085	0.065	0.203	0.219	0.173
R3	0.171	0.141	0.082	0.093	0.210	0.226	0.163

GTP								
	R1	0.874	0.748	0.490	0.340	1.095	0.972	0.749
	R2	1.060	0.742	0.410	0.249	0.785	0.907	0.652
	R3	0.928	0.703	0.433	0.335	0.916	0.799	0.526
UTP								
	R1	1.548	1.140	0.719	0.393	2.111	2.235	1.780
	R2	1.780	1.192	0.590	0.284	1.745	2.046	1.779
	R3	1.498	1.134	0.642	0.411	1.870	1.774	1.259
dATP								
	R1	0.147	0.126	0.088	0.063	0.114	0.108	0.082
	R2	0.164	0.120	0.078	0.041	0.093	0.090	0.080
	R3	0.163	0.116	0.073	0.057	0.102	0.088	0.065
dCTP								
	R1	0.233	0.152	0.100	0.057	0.206	0.189	0.146
	R2	0.258	0.159	0.082	0.040	0.167	0.168	0.143
	R3	0.224	0.157	0.087	0.057	0.189	0.158	0.108
dTTP								
	R1	0.264	0.220	0.168	0.090	0.241	0.228	0.222
	R2	0.306	0.224	0.150	0.072	0.202	0.199	0.216
	R3	0.297	0.222	0.143	0.106	0.228	0.195	0.178
NAD^+								
	R1	0.676	0.896	0.814	0.746	0.572	0.577	0.554
	R2	0.675	0.837	0.841	0.628	0.542	0.531	0.562
	R3	0.760	0.868	0.787	0.677	0.583	0.543	0.598
$NADP^+$								
	R1	0.272	0.359	0.309	0.331	0.241	0.216	0.196
	R2	0.292	0.328	0.310	0.264	0.213	0.222	0.210
	R3	0.308	0.355	0.310	0.338	0.223	0.199	0.185
UDP-GlcNAc								
	R1	0.313	0.215	0.154	0.121	0.364	0.394	0.372
	R2	0.353	0.204	0.135	0.117	0.326	0.390	0.354
	R3	0.432	0.187	0.140	0.108	0.343	0.404	0.389
IMP								
	R1	0.010	0.008	0.007	0.004	0.024	0.024	0.017
	R2	0.011	0.010	0.007	n.d.	0.031	0.016	0.014
	R3	0.011	0.015	0.008	n.d.	0.028	0.015	0.007
dGTP/pApp								
	R1	0.129	0.156	0.130	0.108	0.169	0.096	0.089
	R2	0.140	0.159	0.119	0.084	0.105	0.088	0.119
	R3	0.105	0.203	0.137	0.132	0.124	0.086	0.108

*NADH and NADPH were not detected due to instability during extraction procedure and storage conditions for extracts.

Table 3.5. Strains used for the work presented in chapter III.

Organism	Genotype	Description	Reference
P. aeruginosa PAO1	wild-type		(Stover et al., 2000)
	ΔPA4856	*retS* deletion strain	This study
	ΔPA4856 attB::*lacZ*	*retS* deletion strain, constitutive *lacZ* expression strain, Tet^R	This study
	ΔPA4856 ΔPA0093	*retS tse6* deletion strain	This study
	ΔPA4856 ΔPA0093 ΔPA0431	*retS Pa*SAH deletion strain	This study
P. aeruginosa PA14	wild-type		(Lee et al., 2006a)
	ΔPA14_52570	*rsmA* deletion strain	This study
	ΔPA14_52570 ΔPA14_68450	*rsmA rsmF* deletion strain	This study
	ΔPA14_52570 ΔPA14_68450 attB::*lacZ*	*rsmA rsmF* deletion strain, constitutive *lacZ* expression strain, Tet^R	This study
	ΔPA14_52570 ΔPA14_68450 ΔPA14_01100	*rsmA rsmF clpV1* deletion strain	This study
	ΔPA14_52570 ΔPA14_68450 ΔPA14_01140	*rsmA rsmF tas1* deletion strain	This study
	ΔPA14_52570 ΔPA14_68450 ΔPA14_01120	*rsmA rsmF tsi6* deletion strain	This study
	ΔPA14_52570 ΔPA14_68450 ΔPA14_01140 ΔPA14_01130 attB::*lacZ*	*rsmA rsmF tas1 tis1* deletion strain, constitutive *lacZ* expression strain, Tet^R	This study
	ΔPA14_52570 ΔPA14_68450 ΔPA14_01140 ΔPA14_01130 ΔPA14_01120 attB::*lacZ*	*rsmA rsmF tas1 tis1 tsi6* deletion strain, constitutive *lacZ* expression strain, Tet^R	This study
	ΔPA14_52570 ΔPA14_68450 ΔPA14_01150	*rsmA rsmF eagT6* deletion strain	This study
	ΔPA14_52570 ΔPA14_68450 ΔPA14_01110	*rsmA rsmF vgrG1* deletion strain	This study
	ΔPA14_52570 ΔPA14_68450 ΔPA14_01160	*rsmA rsmF vgrG2* deletion strain	This study
	ΔPA14_52570 ΔPA14_68450 ΔPA14_29390	*rsmA rsmF vgrG4* deletion strain	This study
	ΔPA14_64230	*retS* deletion strain	This study
	ΔPA14_64230 ΔPA14_57520	*retS sspB* deletion strain	This study

	ΔPA14_64230 ΔPA14_57520 PA14_01130-DAS+4	*retS sspB* deletion strain expressing Tis1 with a C-terminal DAS+4 tag	This study
	ΔPA14_52570 ΔPA14_68450 PA14_01140-VSV-G	*rsmA rsmF* deletion strain expressing Tas1 with a C-terminal VSV-G tag	This study
	ΔPA14_52570 ΔPA14_68450 PA14_01140-E382A	*rsmA rsmF* deletion strain expressing $Tas1^{E382A}$	This study
	ΔPA14_52570 ΔPA14_68450 ΔPA14_05600 attB::*lacZ*	*rsmA rsmF Pa*SAH deletion strain, constitutive *lacZ* expression strain, Tet^R	This study
	ΔPA14_52570 ΔPA14_68450 ΔPA14_01140 ΔPA14_01130 ΔPA14_05600 attB::*lacZ*	*rsmA rsmF tas1 tis1 Pa*SAH deletion strain, constitutive *lacZ* expression strain, Tet^R	This study
E. coli MG1655	wild-type		(Edwards and Palsson, 2000)
E. coli SM10 λpir	Km^R*, thi-1 thr leu tonA lac Y supE recA*::RP4-2-Tc::Mu, pir	Conjugation strain	BioMedal LifeScience
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

Table 3.6. Plasmids used for the work presented in chapter III.

Plasmid	Relevant features	Reference
pEXG2	Allelic replacement vector containing *sacB*, Gm^R	(Rietsch et al., 2005)
pPSV38-CV	Expression vector with *lacI*, *lacUV5* promoter, C-terminal VSV-G tag, Gm^R	(Castang et al., 2008)
pPSV39-CV	Expression vector with *lacI*, *lacUV5* promoter, C-terminal VSV-G tag, Gm^R	(Silverman et al., 2013)
pSCrhaB2-CV	Expression vector with *PrhaB*, Tmp^R	(Cardona and Valvano, 2005)
pETDuet-1	Co-expression vector with *lacI*, T7 promoter, N-terminal His_6 tag in MCS-1, Amp^R	Novagen
pET28b	Expression vector with *lacI*, T7 promoter, C-terminal His_6 tag, Amp^R	Novagen
pET29b	Expression vector with *lacI*, T7 promoter, C-terminal His_6 tag, Kan^R	Novagen
pCfa	Expression vector with *lacI*, T7 promoter, C-terminal Cfa-His_6 intein tag, Amp^R	(Wang et al., 2019)
pKSV45-Amp^R	Expression vector with *tetR* and *ptet* promoter, replaced kan^R with amp^R	(McKenna and Nielsen, 2011)
pALS13	Expression vector for (p)ppGpp synthetase fragment of RelA 1-455 (RelA′)	(Schreiber et al., 1991)
pSW196	MiniCTX1 plasmid, Tet^R	(Baynham et al., 2006)
pEXG2::ΔPA0905	*rsmA* deletion construct for PAO1	This study
pEXG2::ΔPA5182	*rsmF* deletion construct for PAO1	This study
pEXG2::ΔPA0093	*tse6* deletion construct	This study
pEXG2::ΔPA0090	*clpV1* deletion construct for PAO1	This study
pEXG2::ΔPA14_52570	*rsmA* deletion construct for PA14	This study
pEXG2::ΔPA14_68450	*rsmF* deletion construct for PA14	This study
pEXG2::ΔPA14_01100	*clpV1* deletion construct for PA14	This study
pEXG2::ΔPA14_01140	*tas1* deletion construct	This study
pEXG2::ΔPA14_01120	*tsi6* deletion construct	This study
pEXG2::ΔPA14_01140-01130	*tas1 tis1* deletion construct	This study
pEXG2::ΔPA14_01140-01120	*tas1 tis1 tsi6* deletion construct	This study
pEXG2::ΔPA14_01150	*eagT6* deletion construct	This study
pEXG2::ΔPA14_01110	*vgrG1* deletion construct	This study
pEXG2::ΔPA14_01160	*vgrG2* deletion construct	This study
pEXG2::ΔPA14_29390	*vgrG4* deletion construct	This study
pEXG2::ΔPA14_64230	*retS* deletion construct	This study
pEXG2::ΔPA14_57520	*sspB* deletion construct	This study
pEXG2::PA14_01130_DAS+4	For generating strains encoding Tis1 fused to a C-terminal DAS+4 tag (AANDENYSENYADAS)	This study

pEXG2::PA14_01140-VSV-G	For generating strains encoding Tas1 fused to a C-terminal VSV-G epitope tag (YTDIEMNRLGK)	This study
pEXG2::PA14_01140_E382A	For generating strains encoding Tas1^{E382A}	This study
pEXG2::ΔPA0431	*Pa*SAH deletion construct for *P. aeruginosa* PAO1	This study
pEXG2::ΔPA14_05600	*Pa*SAH deletion construct for *P. aeruginosa* PA14	This study
pPSV39-CV::PA14_01130	Vector used to express Tis1	This study
pPSV39-CV::PA14_01120	Vector used to express Tsi6	This study
pPSV39-CV::PA14_57520	Vector used to express SspB	This study
pPSV39-CV::PA5338	Vector used to express SpoT	This study
pPSV39-CV::PA0431	Vector used to express *Pa*SAH (PAO1)	This study
pPSV39-CV::PA0431_E74A_D75A	Vector used to express the inactive *Pa*SAH$^{E74A/D75A}$ (PAO1) variant	This study
pPSV39-CV::PA14_05600	Vector used to express *Pa*SAH (PA14)	This study
pPSV39-CV:: PA14_05600_E74A_D75A	Vector used to express the inactive *Pa*SAH$^{E74A/D75A}$ (PA14) variant	This study
pPSV39-CV:: BACCAC_01146	Vector used to express *Bc*SAH	This study
pPSV39-CV::BACCAC_01147	Vector used to express *Bc*Tis1	This study
pPSV39-CV::BACCAC_01146-7	Vector used to express *Bc*SAH-*Bc*Tis1	
pSCrhaB2-CV::PA14_01140_251-CT	Expression vector for the C-terminal toxin domain of Tas1 (Tas1$_{tox}$)	This study
pSCrhaB2-CV:: PA14_01140_251-CT_K305A	Vector used to express Tas1$_{tox}^{K305A}$	This study
pSCrhaB2-CV:: PA14_01140_251-CT_K313A	Vector used to express Tas1$_{tox}^{K313A}$	This study
pSCrhaB2-CV:: PA14_01140_251-CT_K326A	Vector used to express Tas1$_{tox}^{K326A}$	This study
pSCrhaB2-CV:: PA14_01140_251-CT_D327A	Vector used to express Tas1$_{tox}^{D327A}$	This study
pSCrhaB2-CV:: PA14_01140_251-CT_R330A	Vector used to express Tas1$_{tox}^{R330A}$	This study
pSCrhaB2-CV:: PA14_01140_251-CT_E382A	Vector used to express Tas1$_{tox}^{E382A}$	This study
pSCrhaB2-CV:: PA14_01140_251-CT_E382D	Vector used to express Tas1$_{tox}^{E382D}$	This study
pKSV45::His$_6$-Tas1-VSV-G	Vector used to express Tas1$_{tox}$ with N-terminal His$_6$ and C-terminal VSV-G epitope	This study
pSCrhaB2-CV:: BACCAC_01148_D339-CT-VSV-G	Vector used to express a C-terminal VSV-G tagged Apk2 toxin domain (D339-CT, Apk2$_{tox}$)	This study
pSCrhaB2-CV:: BACCAC_01148_D339-CT_E465A-VSV-G	Vector used to express Apk2$_{tox}^{E465A}$	This study
pSCrhaB2-CV:: Salbus254_1145_E34-CT-VSV-G	Vector used to express a C-terminal VSV-G tagged Apk3 variant lacking its N-terminal signal sequence (Apk3$_{\Delta SS}$)	This study

pSCrhaB2-CV:: Salbus254_1145_E34-CT_D121A-VSV-G	Vector used to express $Apk3_{\Delta SS}^{D121A}$	This study
pSCrhaB2-CV::VP1295_A23-CT-VSV-G	Vector used to express a C-terminal VSV-G tagged Apk4 variant lacking its N-terminal signal sequence ($Apk4_{\Delta SS}$)	This study
pSCrhaB2-CV::VP1295_A23-CT_E186A_VSV-G	Vector used to express $Apk4_{\Delta SS}^{E186A}$	This study
pETDuet-1::His_6_PA14_01140_251-CT:: PA14_01130	Co-expression vector for $Tas1_{tox}$ with N-terminal His_6 and Tis1	This study
pETDuet-1::His_6_PA14_01140_251-CT_VSV-G:: PA14_01130	Co-expression vector for $Tas1_{tox}$ with N-terminal His_6 and C-terminal VSV-G and Tis1	This study
pETDuet-1::PA5338_1-387_His_6	Expression vector for the $SpoT_{1-387}$ fragment with C-terminal His_6	This study
pETDuet-1::His_6-BACCAC_01148_D339-CT::BACCAC_01147	Vector used to express N-terminal His_6-tagged Apk2 toxin domain (D339-CT) with untagged *Bc*Tis1	This study
pET28b:: D8B36_07150	Expression vector for $PurF^{EC}$	(Wang et al., 2019)
pET28b:: D8B36_21350	Expression vector for $PurD^{EC}$	(Wang et al., 2019)
pET29b::Salbus254_1145-His_6	Vector used to express full-length Apk3	This study
pET29b::VP1295-His_6	Vector used to express full-length Apk4	This study
pCfa:: D8B36_07150	Expression vector for $PurF^{EC}$	(Wang et al., 2019)
pCfa:: D8B36_07150_R62A	Expression vector for $PurF^{EC\ R62A}$	(Wang et al., 2019)
pCfa:: PA14_23290	Expression vector for $PurF^{PA}$	This study
pSW196:: *lacZ*	*lacZ* in miniCTX1 plasmid	(Whitney et al., 2014)

Methods

Bacterial strains and growth conditions

P. aeruginosa strains generated in this study were derived from the sequenced strains PAO1 and PA14 (Table 3.5) (Lee et al., 2006a; Stover et al., 2000). For co-culture experiments, growth curves and secretion assays, *P. aeruginosa* strains were grown at 37°C in LB medium (10 g/L NaCl, 10 g/L tryptone, and 5 g/L yeast extract). Solid media contained 1.5% or 3% agar. For analysis of cellular extracts and preparation of samples for metabolomics, *P. aeruginosa* strains were grown at 30°C overnight and sub-inoculated at 37°C, in LB medium. Media were supplemented with gentamicin (30 µg/mL) and IPTG (500 µM) as appropriate. *E. coli* strains XL-1 Blue, SM10, BL21 (DE3) CodonPlus, and MG1655 were used for plasmid maintenance, conjugative transfer, gene expression, growth curves and nucleotide extraction experiments, respectively (Table 3.5). *E. coli* strains were grown 37°C in LB medium with the exception of the $PurF^{PA}$ expression and nucleotide extraction experiments shown in Figure 3.7. For $PurF^{PA}$, the expression strain was grown in M9 medium (14 g/L $Na_2HPO_4{\cdot}7H_2O$, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 1 g/L NH_4Cl, 1 mM $MgSO_4$, and 30 µM $CaCl_2$) supplemented with 0.4% glucose and 25 µM $Fe(SO_4)$-EDTA chelate. For nucleotide extraction experiments, cells were grown in M9 medium supplemented with 0.1% glucose, 0.25% each of L-serine and L-threonine, 0.0375% each of L-asparagine and L-glutamine, 0.015% each of all 16 other natural amino acids, and 1× Kao & Michayluk Vitamin Solution (abbreviated as M9GAV). Where appropriate, media

were supplemented with 150 µg/mL carbenicillin, 50 µg/mL kanamycin, 200 µg/mL trimethoprim, 15 µg/mL gentamicin, 500 µM IPTG, 0.1% (w/v) rhamnose or 40 µg/mL X-gal.

DNA manipulation and plasmid construction

All DNA manipulation procedures followed standard molecular biology protocols. Primers were synthesized and purified by Integrated DNA Technologies (IDT). Phusion polymerase, restriction enzymes and T4 DNA ligase were obtained from New England Biolabs (NEB). DNA sequencing was performed by Genewiz Incorporated.

In-frame chromosomal deletion mutants in *P. aeruginosa* were generated using the pEXG2 suicide plasmid as described previously (Rietsch et al., 2005). Briefly, ~500bp upstream and downstream of target gene were amplified by standard PCR and spliced together by overlap-extension PCR. The resulting DNA fragment was ligated into pEXG2 using standard cloning procedures (see Table 3.6 for plasmid details). Deletion constructs were introduced into *P. aeruginosa* via conjugal transfer and *sacB*-based allelic exchange was carried out as described previously (Hmelo et al., 2015). All deletions were confirmed by PCR.

Bioinformatics

Bioinformatic analysis of tse6 and tas1 distribution among P. aeruginosa strains

Complete or draft assembled genome sequences for 326 *P. aeruginosa* isolates representing a broad sampling of *P. aeruginosa* diversity were downloaded from the *Pseudomonas* Genome DB (Winsor et al., 2016). Open reading frames were predicted for each isolate using Prodigal v2.6.1 and the resulting putative proteomes compared to the Tse6 and Tas1 sequences using BLASTP v2.8.1, with automated and manual inspection of the results to identify all homologs and sequence variants within each genome (Camacho et al., 2009; Hyatt et al., 2010). Phylogenetic relationships of the isolates were reconstructed using whole-genome SNP analysis; homologous sites in the genomes containing nucleotide variation among isolates, but not involved in horizontal gene transfer or recombination, were identified using PARSNP v1.2 with PhiPack filtering (Treangen et al., 2014). The resulting SNP matrix was converted to PHYLIP format and the phylogenetic history of the isolates reconstructed using maximum likelihood as implemented in the RAxML-HPC BlackBox v8.2.10 hosted on the CIPRES Science Gateway server (Stamatakis, 2014). RAxML analysis included automatically generated bootstrapping and estimated proportion of invariable sites (GTRGAMMA+I). The resulting Tse6 and Tas1 homologs were mapped onto the isolate phylogenetic tree using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Bioinformatic identification of Tas1 homologs

We used *jackhmmer* to generate a sequence alignment hidden Markov model (HMM) for Tas1 using an iterative search procedure that queried the UniProtKB database which identified 2,468 homologs of Tas1 (Johnson et al., 2010). To build a phylogenetic tree, we aligned sequences using MAFFT v.7.455 `—auto` (Katoh and Standley, 2013) and trimmed the alignment using trimAl v1.4 `-gt 0.8 -st 0.001 -cons 60` (Capella-Gutiérrez et al., 2009). The phylogenetic tree was built off FastTree v2.1.10 `-gamma` (Price et al., 2009). To cluster genomes in the phylogenetic tree, we used MMseqs v11.e1a1c with `--min-seq-id 0.05 -c 0.05` parameters (Hauser et al., 2016). The clusters were then further subdivided by performing a pairwise sequence alignment with the first two sequences at the root of the tree. The sequences form a group if they are > 24.5% similar to the previous sequence. The group is discontinued when there are two sequences that are < 24.5% similar. The iterative grouping process continues until the last sequence of the tree.

Toxicity experiments

For assays shown in Figures 3.1G, 3.3B and 3.4E, *E. coli* XL-1 Blue cells were co-transformed with either pSCrhaB2-CV or pSCrhaB2-CV expressing either wild-type or active site mutants of $Tas1_{tox}$ and pPSV39-CV or pPSV39-CV expressing Tis1, $Tsi6^{PA14}$, $Tsi6^{PAO1}$. For Figure 3.15B, *E. coli* XL-1 Blue cells were transformed with either pSCrhaB2-CV or pSCrhaB2-CV expressing either $Tas1_{tox}$

or wild-type or catalytically inactive Apk2-4 (see Table 3.6 for more details about each construct). For Figure 3.16B-C and E, *E. coli* XL-1 Blue cells were co-transformed with either pSCrhaB2-CV or pSCrhaB2-CV expressing $Tas1_{tox}$ or Apk2 and pPSV39-CV or pPSV39-CV expressing *Bc*SAH, *Bc*Tis1 or Tis1. Overnight cultures of the above strains were diluted to 10^6 in 10-fold increments and each dilution was spotted onto LB agar plates containing 0.1% (w/v) L-rhamnose or 0.1 (w/v) L-rhamnose and 250 µM IPTG and the appropriate antibiotics. Photographs were taken after overnight growth at 37°C.

To compare Tas1 toxicity in the presence and absence of *Pa*SAH (Figure 3.13B-C), *E. coli* XL-1 Blue strains were co-transformed with pSCrhaB2-CV or pSCrhaB2-CV expressing wild-type $Tas1_{tox}$ and pPSV39-CV or pPSV39-CV containing $PaSAH^{PAO1}$ and $PaSAH_{E74A/D75A}{}^{PAO1}$. The above strain co-expressing $Tas1_{tox}$ and Tis1 was used as an additional control. Overnight cultures of these *E. coli* strains grown in LB medium (37°C, 200 rpm) were diluted 1:100 in 200 µL of fresh LB medium containing selection in a 96-well plate. The plate was placed into a Synergy 4 (BioTek) plate reader set at 37°C with medium shaking and the A_{600nm} of the cultures was measured every 30 min for the duration of the assay. Inducers were added after 2.5 hr once the strain containing empty vectors reached mid-log optical density (A_{600nm} of 0.2). For CFU enumeration, the above strains were grown in LB medium (37°C, 200 rpm) overnight. Stationary-phase overnight cultures were used to inoculate 50 mL LB and allowed to grow to an A_{600nm} of 0.4. Samples were withdrawn from these cultures, diluted to 10^{-6} using 10-fold serial dilutions, and

plated on noninducer containing agar to enumerate the CFUs before induction. Inducer was added to each culture followed by further incubation for 1 hr, after which samples for counting of postinduction CFUs were withdrawn, again diluted and plated on noninducer containing LB agar plates. CFUs were counted following 24 hr incubation on LB agar plates at 37°C.

To compare (p)ppApp versus (p)ppGpp toxicity or (p)ppApp synthetase toxicity (shown in Figures 3.8G-I and 3.15C-D, respectively), *E. coli* MG1655 was transformed with plasmids expressing either $Tas1_{tox}$, RelA' or Apk2-4 (see Table 3.6 for construct details) and grown in LB at 30°C (Figure 3.8G-I) or 37°C (Figure 3.15C-D) overnight with appropriate antibiotic selection. Stationary-phase overnight cultures were diluted to OD 0.01 in fresh LB medium and grown at 37°C with shaking. At OD 0.3, 1mL of culture was retrieved and chilled on ice water and the remaining culture was treated with 500 μM IPTG (RelA' construct) or 0.1% (w/v) rhamnose (all pSCrhaB2-CV constructs). At indicated time points post-induction, 1 mL of culture was withdrawn for OD measurements and 1 mL of the sample was chilled on ice water for 2 minutes, pelleted at 20,000 *g* at 4°C and re-suspended in ice-cold fresh LB without inducer. These samples were diluted to 10^6 in 10-fold increments and 10 μL from each dilution was spotted onto LB agar plates containing appropriate antibiotics. Nucleotides levels in these strains were examined by anion exchange chromatography as described below (see Metabolite extraction and quantification – *nucleotide quantification using anion-exchange chromatography*).

Secretion assays

Stationary-phase overnight cultures of *P. aeruginosa* strains were inoculated in 2 mL LB at a ratio of 1:500. Cultures were grown at 37 °C with shaking to mid-log phase, and cell and supernatant fractions were prepared as described previously (Hood et al., 2010).

Quantification of $Tas1_{tox}$ copy number

E. coli MG1655 harboring an anhydrotetracycline (aTC)-inducible His_6-$Tas1_{tox}$-VSV-G expression plasmid was grown to OD_{600} = 0.3 and either untreated or induced with 2 or 3 ng/mL aTC for 15 minutes. Viability of each culture was assessed by enumerating CFUs. Cells from 100 mL cultures (5 x 10^9 CFU) were collected for Ni-NTA enrichment of His_6-$Tas1_{tox}$-VSV-G. Briefly, cells from each culture were lysed in 1 mL lysis buffer (20 mM HEPES-Na 7.4, 150 mM NaCl) and applied to a column of 0.5 mL Ni-NTA resin. Each column was washed with 3 mL lysis buffer containing 20 mM imidazole followed by 3 mL 20 mM HEPES-Na 7.4, 500 mM NaCl, 20 mM imidazole containing 8 M Urea. Bound protein was eluted with a buffer containing 300 mM imidazole and 8 M urea, and each eluate was concentrated to 60 µL. 15 µL of concentrated eluate (25% of total protein) underwent anti-VSV-G immunoblotting for quantification, using 50 and 15 fmol of purified, recombinant His_6-$Tas1_{tox}$-VSV-G as internal standards. Assuming 100% recovery of His_6-$Tas1_{tox}$-VSV-G by Ni-NTA enrichment, cells induced by 2 or 3

ng/mL aTC contain 24 and 44 fmol Tas1, which, provided a cell count of 5 x 10^9, correspond to 3 and 5 copies of His_6-$Tas1_{tox}$-VSV-G per cell, respectively.

Western blot analyses

Western blot analyses of protein samples were performed as described previously using rabbit anti-Tse1 (diluted 1:5000; Genscript), rabbit anti-VSV-G (diluted 1:5,000; Sigma), rabbit anti-Hcp (diluted 1:5,000) and detected with anti-rabbit horseradish peroxidase-conjugated secondary antibodies (diluted 1:5,000; Sigma) (Whitney et al., 2015). Western blots were developed using chemiluminescent substrate (Clarity Max, Bio-Rad or SuperSignal™ West Femto Maximum Sensitivity Substrate, Thermofisher) and imaged with the ChemiDoc Imaging System (Bio-Rad).

Competition assays

A *lacZ* cassette was inserted into a neutral phage attachment site (*attB*) of recipient *P. aeruginosa* strains to differentiate these strains from unlabeled donors. For inter-strain competitions, *P. aeruginosa* PA14 donor and PAO1 recipient strains were grown in 2 ml LB medium overnight. One milliliter of each stationary-phase culture was pelleted at 7,600g for 3 min and washed twice with 1 ml of fresh LB medium. The resuspended cells were diluted to an A_{600nm} of 0.4 (approximately 3×10^5 CFU) and mixed at a 4:3 (v/v) (PAO1:PA14) ratio, with the except of Figure 3.13E where the ratios are indicated. For intraspecific competitions with *P.*

aeruginosa PA14 donor and recipient strains, stationary-phase cultures were directly mixed in a 1:1 (v/v) ratio except for the viability assays shown in Figure 3.13H, where PA14 donor and recipient strains (approximately 3×10^5 CFU) were mixed in a 10:1 (donor:recipient) ratio.

Starting ratios of donor and recipient were enumerated by plating on LB agar containing 40 µg/ml X-gal. Ten microlitres of each competition mixture was then spotted in triplicate on a 0.45 µm nitrocellulose membrane overlaid on a 3% LB agar plate and incubated face up at 37 °C for 20-24 h. Competitions were then harvested by resuspending cells in LB and enumerating colony forming units by plating on LB agar containing 40 µg/ml X-gal. The final ratio of donor/recipient colony forming units were normalized to the starting ratio of donor and recipient strains.

To monitor (p)ppApp production in recipients by anion exchange, 600µL of donor and recipient mixtures were plated on a 25-mm, 0.45 µM PVDF membranes using vacuum filtration. The membrane was overlaid onto 3% LB agar and incubated face up for 2-7 h. Following incubation, each membrane was immersed in 2 mL ice-cold lysis solvent, a methanol-acetonitrile-water mixture in a volume ratio of 40:40:20. After brief sonication to detach cells from the PVDF membrane, the membrane was removed, and the entire suspension was diluted into 6mL of 20mM Tris-HCl buffer (pH 8.0). The diluted mixture was spun at 10,000 *g* to pellet insoluble debris. After passage of a 0.22-µm syringe filter, 4 mL of the supernatant was analyzed using anion-exchange chromatography as described below (see

Metabolite extraction and quantification – *nucleotide quantification using anion-exchange chromatography*).

Tis1 depletion system

A C-terminal *ssrA*-like DAS+4 degradation tag (peptide sequence of tag: AANDENYSENYADAS) was fused to the 3' end of the native *tis1* locus in a *P. aeruginosa* strain bearing deletions in *retS* and *sspB* (McGinness et al., 2006). An IPTG-inducible plasmid containing *sspB* was used to stimulate controlled degradation of Tis1-DAS+4 (Tis1-D4). The SspB protein recognizes DAS+4 tagged proteins and delivers them to the ClpXP protease for degradation. Strains harbouring this plasmid were streaked on LB agar supplemented with 30 µg/mL gentamicin and 500 µM IPTG.

Protein expression and purification

$Tas1_{tox}$-Tis1 complex, BcSAH and Apk2-4

$Tas1_{tox}$ or $Tas1_{tox}^{E382A}$ were coexpressed with Tis1 from pETDuet-1, *Bc*SAH was expressed from pET29b, the Apk2 (*Bc* homolog, BACCAC_01148 locus tag) toxin domain (D339-CT, $Apk2_{tox}$) and *Bc*Tis1 were co-expressed from pETDuet-1, full-length Apk3 (*Sa* homolog, Salbus254_1145 locus tag) and Apk4 (*Vp* homolog, VP1295 locus tag) were each expressed from pET29b (see Table 3.6 for plasmid details). Expression for the above constructs was carried out using *E. coli* BL21 (DE3) CodonPlus cells. For Tas1-Tis1 and *Bc*SAH, 40 mL overnight cultures of

expression strains were diluted into 2 L of LB broth and grown to mid-log phase (OD_{600} = 0.6) in a shaking incubator at 37°C. For Apk2-*Bc*Tis1, Apk3 and Apk4, 1 mL of stationary-phase overnight culture grown in LB was diluted into 100 mL of LB broth and grown to mid-log phase at 37°C. Protein expression Tas1-Tis1, Apk3 was induced by the addition of 1mM IPTG and cells were further incubated for 3.5 h at 37°C, whereas cultures expressing Apk2 and Apk4 were cooled to 18°C prior to adding inducer and incubated at this temperature overnight prior to harvest. Cells were harvested by centrifugation at 9,800 *g* for 10 min and resuspended in 25 mL (for cultures greater than 1 L) or 3.5 mL (for 50-100 mL cultures) of lysis buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM imidazole) prior to rupture by sonication (6 x 30 second pulses, amplitude 30%). Cell lysates were cleared by centrifugation at 39,000 *g* for 60 min and the soluble fraction was loaded onto a gravity flow Ni-NTA column that had been equilibrated in lysis buffer.

To obtain $Tas1_{tox}$–Tis1 complex or *Bc*SAH for crystallization, Ni-NTA bound complex was washed twice with 25 mL of lysis buffer followed by elution in 10 mL of lysis buffer supplemented with 400 mM imidazole. The Ni-NTA purified complex was further purified by gel filtration using a HiLoad 16/600 Superdex 200 column equilibrated in 20 mM Tris-HCl pH 8.0 150 mM NaCl. Fractions with the highest purity were used for subsequent crystallization screening.

To obtain Apk3 and Apk4 for enzyme assays, a similar procedure was followed, except only 200 µL of Ni-NTA agarose was used for 3.5 mL of lysate,

washes were completed with only 1 mL of lysis buffer and final elution volume was 400 µL.

To obtain $Tas1_{tox}$, $Tas1_{tox}^{E382A}$ or $Apk2_{tox}$ for enzyme assays, Tis1 or *Bc*Tis1 was removed from Ni-NTA immobilized $Tas1_{tox}$, $Tas1_{tox}^{E382A}$ or $Apk2_{tox}$ by washing the column twice with 25 mL of lysis buffer supplemented with 8 M urea. On-column refolding was achieved by washing twice with 25 mL of lysis buffer followed by elution of the renatured proteins using lysis buffer supplemented with 400 mM imidazole. Refolded $Tas1_{tox}$, $Tas1_{tox}^{E382A}$ and $Apk2_{tox}$ were further purified by gel-filtration as described above except that the running buffer was comprised of 20 mM HEPES pH 7.4, 150 mM NaCl. Purified proteins were then flash frozen until needed.

$SpoT_{1-387}$

The SpoT 1-387 fragment from *P. aeruginosa* was expressed from pETDuet-1 in *E. coli* BL21 (DE3) CodonPlus cells. The same expression protocol was followed as described for the *$Tas1_{tox}$-Tis1 complex*. To obtain $SpoT_{1-387}$ for enzyme assays, cleared cell lysates containing $SpoT_{1-387}$ were loaded onto a gravity flow Ni-NTA column that had been equilibrated in lysis buffer. The Ni-NTA bound $SpoT_{1-387}$ was washed twice with 25 mL of lysis buffer followed by elution in 10 mL of lysis buffer supplemented with 400 mM imidazole. The Ni-NTA purified complex was further purified by gel-filtration as described above except that the

running buffer was comprised of 20 mM HEPES pH 7.4, 150 mM NaCl. Purified proteins were then flash frozen until needed.

$PurF^{EC}$ and $PurF^{PA}$

$PurF^{EC}$ used for crystallization was expressed without an affinity tag (Table 3.6). $PurF^{EC}$ and $PurF^{PA}$ used in biochemical experiments was expressed as a fusion protein with a C-terminal self-cleaving Cfa-His_6 tag(Stevens et al., 2016). Cultures of expression strains were grown to mid-log phase, cooled to 22°C and induced with 200 µM IPTG for 20 h. For $PurF^{PA}$, M9 minimal medium was used for expression.

For untagged $PurF^{EC}$, cell pellets (~10 g wet weight) were resuspended in 40 mL lysis buffer containing 50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10 mM $MgCl_2$, 5 mM DTT, 20 µg/mL lysozyme and 1 mM PMSF. Cells were lysed by sonication and centrifuged at 15,000 *g* for 10 min. Cleared lysates were treated with protamine sulfate (8 mg per gram of cell pellet), vortexed. Precipitate was pelleted at 30, 000 *g* for 1 h and cleared lysate was fractionated using a DEAE Sepharose column equilibrated in buffer A (50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$ and 5 mM DTT). The column was washed with 50 mL 5% buffer B (buffer A + 1 M NaCl) and bound protein was eluted using a linear gradient with buffer B concentration increasing from 5% to 55% over 200 mL. Peak fractions were combined and saturated ammonium sulfate was added to samples at 4°C. Precipitated protein collected between 40% and 47.5% saturation was redissolved in gel-filtration buffer (20 mM

HEPES-Na, pH 7.4, 150mM NaCl, 2mM $MgCl_2$ and 1mM TCEP) and run over a Superdex-200 increase (10/300) column.

PurF-Cfa-His_6 was purified using a Ni-NTA affinity column, as previously described (Wang et al., 2019). To cleave the Cfa-His_6 tag, eluate was treated with 100 mM sodium 2-mercaptoethanesulfonate (MESNa), 100 mM L-cysteine, and 20 mM TCEP, pH 7.0 at room temperature overnight. The cleavage mixture was dialyzed against gel filtration buffer and then subjected to a reverse Ni-NTA process. Collected protein was run over a Superdex-200 column equilibrated with gel-filtration buffer.

PaSAH

Native and catalytically inactive *Pa*SAH proteins were expressed in E. coli BL21 (DE3) (NEB). Cells were grown in lysogeny broth (LB) medium supplemented with 50 µg/ml kanamycin and 12.5 g/L D(+)-lactose-monohydrate for 20 hr at 30°C and harvested by centrifugation (3,500g, 20 min, 4°C). The cells were resuspended in lysis buffer (20 mM of HEPES-Na pH 7.5, 20 mM KCl, 250 mM NaCl, and 40 mM imidazole) and lysed by French Press (SLM Aminco) at 1,000 psi pressure. After removal of cell debris by centrifugation (47,850g, 20 min, 4°C), the supernatant was loaded on a 1-ml HisTrap column (GE Healthcare) equilibrated with 10 column volumes (CV) lysis buffer. After washing with 10 CV of lysis buffer, PaSAH was eluted with 5 CV elution buffer (lysis buffer containing 500 mM imidazole). The protein was concentrated (Amicon Ultracel-10K (Millipore)) and

applied to size-exclusion chromatography (SEC) on a HiLoad 26/600 Superdex 200 pg column (GE Healthcare) equilibrated with SEC buffer (20 mM of HEPES-Na, pH 7.5, 20 mM KCl, and 200 mM NaCl). Fractions containing *Pa*SAH were pooled, concentrated (Amicon Ultracel-10K (Millipore)), deep-frozen in liquid nitrogen and stored at −80°C. Protein concentration was determined by a spectrophotometer (NanoDrop Lite, Thermo Scientific).

Crystallization and structure determination

Tas1$_{tox}$-Tis1 complex and BcSAH

Selenomethionine-incorporated Tas1$_{tox}$-Tis1 complex or *Bc*SAH was concentrated to 28 mg/mL (Tas1$_{tox}$-Tis1 complex) or 15 mg/mL (*Bc*SAH) by spin filtration (10 kDa MWCO, MilliporeSigma) and screened for crystallization conditions at 23°C using commercially available sparse matrix screens (MCSG1-4, Anatrace) and the hanging drop vapour diffusion technique. Diffraction quality crystals appeared after approximately one week in conditions containing sodium acetate pH 4.5, 0.8 M sodium phosphate monobasic, 1.2 M potassium phosphate dibasic (for Tas1$_{tox}$-Tis1) and 0.8M LiCl, 0.1M Tris-HCl, 32% PEG4000. Tas1$_{tox}$-Tis1 crystals were cryoprotected in crystallization buffer containing 20% (v/v) ethylene glycol and flash frozen in liquid nitrogen, whereas *Bc*SAH crystals were directly flash frozen prior to data collection. Synchrotron diffraction data were collected at 100K on a Pilatus3 X 6M detector using beamline 19-ID of the Structural Biology Center at the Advanced Photon Source. Diffraction data were processed using HKL3000 (Minor et al., 2006) and the structure was solved by

single wavelength anomalous dispersion (SAD) phasing using Phenix.autosol (Adams et al., 2010). Initial model building was performed using Phenix.autobuild followed by manual adjustment using Coot (Emsley et al., 2010). Refinement was carried out using Phenix.refine with TLS parameterization.

$PurF^{EC}$-ppApp complex

Crystals were grown by hanging-drop vapor diffusion with drops containing 2 µL of protein (25 mg/mL $PurF^{EC}$ in 20 mM HEPES-Na, pH 7.4, 150 mM NaCl, 1 mM TCEP, 5 mM ppApp, and 10 mM $MgCl_2$) mixed with 2 µL of well solution (0.1 M HEPES-Na, pH 7.4, 24% PEG 3350 and 4% iPrOH) at 18°C. After 1 week, crystals were flash frozen in liquid nitrogen without added cryoprotectant. Diffraction data were collected at the APS, with the NE-CAT beamline 25-IDC on a Pilatus 6M detector. Diffraction data were indexed, integrated and scaled using XDS/XSCALE (Kabsch, 2010) and refined with Phenix (Adams et al., 2010). The structure was solved by molecular replacement using Phaser(Bunkoczi et al., 2013) with chain A of PDB entry 6CZF as the search model. The asymmetric unit of the $C222_1$ cell contains two PurF chains forming a symmetric dimer. The D2 symmetry of PurF tetramer is generated by the crystallographic centering operation. As in the 6CZF crystal, each PurF tetramer has four ligand binding sites but can only bind two ligands because pairs of binding sites overlap each other across a twofold symmetry axis of the tetramer. Consequently, each PurF chain in

the PurF/ppApp crystal is modeled with a single ppApp (and its associated Mg^{2+} ion) at 0.5 occupancy.

PaSAH

For crystallization trials, *Pa*SAH was purified in SEC buffer supplemented with 20 mM $MgCl_2$. Crystallization was conducted at 20°C by sitting drop vapor diffusion method in SWISSCI MRC 2-well plates (Jena Bioscience) with a reservoir volume of 50 µl. 0.5 µl of *Pa*SAH concentrated to 1 mM was mixed with 0.5 µl of a solution containing 2.0 M sodium formate and 0.1 M sodium acetate pH 4.6. After three days, 0.5 µl of a cryo-protecting solution containing mother liquor supplemented with 20% (v/v) glycerol was added to the drop, crystals harvested and flash-frozen in liquid nitrogen. Diffraction data were collected at the European Synchrotron Radiation Facility (ESRF) Grenoble, France, at beamline ID29 under laminar nitrogen flow at 100K (Oxford Cryostream 700 Series) with a Pilatus 6M detector at 1.85 Å. Data were processed with XDS (Kabsch, 2010) and the structure determined with Phenix-implemented Phaser (Adams et al., 2010) using the hydrolase domain of RelSeq (PDB: 1VJ7 chain B, amino acids 1–200, (Hogg et al., 2004)) as a search model for molecular replacement. The structure was manually built in Coot (Emsley and Cowtan, 2004) and refined with Phenix refine (Adams et al., 2010). Figures were prepared with PyMOL (www.pymol.org).

Biochemical analyses of $Tas1_{tox}$, Apk2-4

Analysis by coupled enzyme assay

Each reaction (100 μL) contained 50 mM HEPES 7.4, 150 mM NaCl, 20 mM KCl, 10 mM $MgCl_2$, 1 mM TCEP-Na, 5 mM ATP, 1 mM GTP (if indicated) and $Tas1_{tox}$ or $Tas1_{tox}^{E382A}$ at the indicated concentrations. To couple production of AMP in the pyrophosphokinase reaction to the consumption of NADH, the reaction also contained 3.75 M phosphoenolpyruvate (PEP), 0.5 mM NADH, 10 U/mL myokinase (Adenylate kinase, ADK), 20 U/mL pyruvate kinase (PK) and 20 U/mL lactate dehydrogenase (LDH).

$$AMP + 2PEP + 2NADH + 2H^+ = ATP + 2Lactate + 2\ NAD^+$$

Reactions were assembled in 96-well plates, with Tas1 added at $t = 0$. The reactions were monitored at 25°C in a Spectramax M5 plate reader (Molecular Devices) and absorbance at 340 nm (A_{340}) was measured every 15 seconds.

Analysis by anion-exchange chromatography

Each reaction (100 μL total volume) contained 20 mM HEPES-Na 7.4, 300 mM NaCl, 10 mM $MgCl_2$, and substrates at above concentrations. $Tas1_{tox}$, $Apk2_{tox}$, Apk3 and Apk4 were diluted to 1 nM (Tas1) or 50 nM added last. Reactions were incubated at 37°C ($Tas1_{tox}$ turnover experiment in Figure 3.6A, Apk2-4 (p)ppApp synthesis in Figure 3.15E) or 25°C (all other reactions). At the indicated time points, each 50 μL reaction was diluted in 1 mL ice-cold water and then applied to a MonoQ 5/50 column (GE Healthcare). Bound nucleotides were eluted at 4°C using a linear gradient of buffer A (5 mM Tris-HCl pH 8.0) and buffer B (5 mM Tris-HCl

pH 8.0, 1M NaCl), with the percentage of buffer B increasing from 0 to 40% over 20 mL.

Tas1 turnover measurement

Each reaction (200 μL total volume) contained 20 mM HEPES-Na pH 7.4, 150 mM NaCl, 15 mM $MgCl_2$, 10 mM ATP, 25 mM PEP-K, 10U/mL each ADK and PK, and 1 nM $Tas1_{tox}$ or 1 μM $Tas1_{tox}^{E382A}$. Reactions were incubated at 37°C and 20μL was diluted in 1 mL ice-cold water at the indicated time points and analyzed by anion-exchange chromatography as described above. The 3′-pyrophosphorylated product, pppApp, was quantified based on the integration of the A_{254} trace.

Additional biochemical analyses by enzyme-coupling readout

Reactions were assembled in 96-well plates from 10X stocks of individual components, with enzymes added at t = 0. Then reactions were monitored at 30 ^{o}C in a Spectramax M5 plate reader (Molecular Devices) for the absorbance at 340 nm (A_{340}) every 30 seconds.

PurF glutamine amidophosphoribosyltransferase assay

Each reaction (100 μL) contained 50nM $PurF^{EC}$ or 100nM $PurF^{PA}$ in 50 mM HEPES-Na pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM TCEP, 5 mM glutamine, 1 mM pRpp-Mg and indicated concentrations of ppGpp-Mg, ppApp-Mg or pppApp-Mg. To couple production of 5′-phosphoribosylamine (PRA) by PurF to the consumption of NADH, the reaction also contains 5 mM ATP, 5 mM glycine and 1

μM *E. coli* PurD (these components ligate glycine to 5′-PRA to form glycinamide 5′-ribonucleotide (GAR) and generate ADP), as well as 3.75M PEP, 0.5 mM NADH, 20U/mL PK and 20U/mL LDH.

glutamine + pRpp + glycine + H_2O + PEP + NADH + H^+ = glutamate + GAR + PPi + Pi + lactate + NAD^+

P. aeruginosa $SpoT_{1-387}$ hydrolase assay

Each reaction (100 μL) contained 40 mM HEPES 7.4, 150 mM KCl, 5 mM $MnCl_2$, 2 mM $MgCl_2$, 1 mM TCEP-Na, 1 mM ATP, 10μM $SpoT_{1-387}$, and 1 mM ppGpp or ppApp. To couple production of ADP or GDP by $SpoT_{1-387}$ to the consumption of NADH, the reaction also contains 3.75M PEP, 0.5 mM NADH, 1 μM *E. coli* nucleoside disphosphate kinase, 20U/mL PK and 20U/mL LDH.

ppGpp + PEP + NADH + H^+ = GTP + PPi + Lactate + NAD^+

ppApp + PEP + NADH + H^+ = ATP + PPi + Lactate + NAD^+

Scale-up preparation of (p)ppApp

pApp

Because Tas1 can slowly convert pppApp and AMP to pApp, quantitative conversion of ATP to pApp is achieved after prolonged incubation with Tas1. Excess AMP was included to ensure the complete consumption of pppApp. We incubated 175 μmol AMP and 75 μmol ATP with 50 nmol Tas1 in 50 mL reaction

containing 20 mM HEPES 7.4, 150 mM NaCl, 5 mM $MgCl_2$ and 1 mM TCEP-Na. The reaction reached completion after 15 minutes at RT and then diluted to 200 mL with ice water. pApp was purified using anion-exchange chromatography.

ppApp

When ADP is present in excess to ATP, Tas1 activity preferentially produces ppApp (Figure 3.5C). Thus, we first incubated 50 μmol ADP with 50 pmol $Tas1_{tox}$ in 5 mL 20 mM HEPES 7.4, 150 mM NaCl, 20 mM $MgCl_2$ and 1 mM TCEP-Na at 37°C. Then, with vigorous stirring, we added 45 μmol ATP in 9 portions over 10 minutes. After another 5 minutes of incubation at 37°C, the reaction was complete and $Tas1_{tox}$ was inactivated with 2 mL chloroform. The aqueous phase was isolated, diluted to 25 mL with water, and ppApp purified using anion-exchange chromatography.

pppApp

After synthesizing pppApp, $Tas1_{tox}$ further converts pppApp and AMP into pApp (Figure 3.6E). To maximize the yield of pppApp, we included ADK, PK and PEP to regenerate ATP from AMP. The synthesis was thus carried out with 50 μmol ATP, 125 μmol PEP, 25 pmol $Tas1_{tox}$ and 200U/mL each PK and ADK in 5mL in the presence of 20 mM HEPES 7.4, 150 mM NaCl, 15mM $MgCl_2$ and 1mM TCEP-Na. After incubation at 37°C for 30 min, 250 nmol $Tas1_{tox}$ was added and the mixture was incubated for another 30 min. $Tas1_{tox}$ was then inactivated with 2

mL chloroform. The aqueous phase was isolated, diluted to 25 mL with water, and pppApp purified using anion-exchange chromatography.

Preparative anion-exchange chromatography

To purify (p)ppApp, a MonoQ 10/100 column (GE Healthcare) was operated at 5 mL/min at room temperature. (p)ppApp synthesis reactions were diluted with water and applied to the column. Bound nucleotides were eluted using a linear gradient of buffer A (5 mM Tris-HCl pH 8.0) and buffer B (5 mM Tris-HCl pH 8.0, 1M NaCl), with the percentage of buffer B increasing from 15 to 40% within 5 column volumes (~ 40 mL). Preparations of pppApp were purified in two runs, while preparations of ppApp were purified in 4 runs. Fractions containing the purified product purity were combined, and LiCl was added to the combined fractions to 1M final concentration. Then, 4x volumes of ethanol was added to precipitate the nucleotide. After incubation in an ice-water bath for 30 min, the nucleotide was collected by centrifugation at 8000 *g* for 10min and the mother liquor decanted. The product was washed with 10 mL 95% ethanol, then dissolved in water and dried on a lyophilizer. The powder was reconstituted in water and concentration determined by absorbance at 260nm (ε = 15,400 $M^{-1}cm^{-1}$).

Metabolite extraction and quantification

Culture and induction conditions

Prior to each experiment, strains were grown overnight at 30°C or 37°C stationary phase as indicated above. The starter culture was diluted to OD_{600} = 0.005 in fresh medium and grown at 37°C. Inducer was added after OD_{600} reached 0.10 for *P. aeruginosa* or 0.25 for *E. coli*. Untreated control samples were harvested 1 minute prior to induction.

Expression of $Tas1_{tox}$, $Tas1_{tox}^{E382A}$, $Apk2_{tox}$, $Apk3_{\Delta SS}$ and $Apk4_{\Delta SS}$ in *E. coli* MG1655: cells were grown in LB or M9GAV containing 250 μg/mL trimethoprim (TMP) and induced using 0.1% rhamnose.

Expression of RelA′ in *E. coli* MG1655: cells were grown in LB containing 100 μg/mL carbenicillin and induced using 500 μM IPTG.

Depletion of Tis1 in *P. aeruginosa* PA14: Tis1 inducible-degradation strain Δ*retS* Δ*sspB* PA14_01130-DAS+4 pPSV9-CV::*sspB* and its parental strain Δ*retS* Δ*sspB* PA14_01130-DAS+4 pPSV9-CV::*sspB* were grown in LB containing 50 μg/mL gentamycin and Tis1 depletion was induced using 500 μM IPTG.

Metabolite extraction from E. coli

E. coli cells (2.5~3.5 OD) were collected on a 0.22-μm hydrophilic PVDF membrane by vacuum filtration and washed briefly with 160 mM NaCl. At the same time, the culture was sampled for OD_{600} measurements. Cells on the membrane were subsequently immersed in ice-cold lysis solvent, a methanol-acetonitrile-water mixture in a volume ratio of 40:40:20. Lysates were briefly sonicated and,

after removal of PVDF membranes, diluted by the lysis solvent for a uniform cell density, typically 1.0 OD_{600} cells per mL solvent.

Metabolite extraction from P. aeruginosa

P. aeruginosa cells (1.25~1.75 OD) were collected on a 0.45-μm hydrophilic PVDF membrane by vacuum filtration and washed briefly with 160 mM NaCl. At the same time, the culture was sampled for OD_{600} measurements. Cells on the membrane were subsequently immersed in ice-cold lysis solvent, a methanol-acetonitrile-water mixture in a volume ratio of 40:40:20 containing 0.02% (*v*/*v*) Metabolomics Amino Acid Mix Standard solution (Cambridge Isotope Laboratories, MSK-A2-1.2) as the internal standard (ISTD). After brief sonication to detach cells from the PDVF membrane, the membrane was removed, and the suspension was diluted using the lysis solvent to 0.625 OD_{600} cells per mL solvent.

Nucleotide quantification using anion-exchange chromatography

(p)ppApp was quantified using anion-exchange chromatography: cell suspension in lysis solvent equivalent to 1.0 OD_{600} cells were diluted with aqueous solution of 10 mM Tris-HCl pH 8.0 until the content of organic solvent less than 20%. Insoluble material was pelleted at 10,000 *g*, and the supernatant was applied to a Mono Q 5/50 column (GE Healthcare) after passing through a 0.22-μm syringe filter. Bound metabolites were eluted at 4°C using a linear gradient of buffer A (5 mM Tris-HCl pH 8.0) and buffer B (5 mM Tris-HCl pH 8.0, 1M NaCl), with the

percentage of buffer B increasing from 0 to 35% over 17.5 mL. External standards containing equimolar of AMP, ADP, ATP, pApp, ppApp and pppApp was analyzed under the same condition to locate their peaks. Nucleotides were quantified according to their peak areas on the 254-nm chromatogram.

MS profiling of P. aeruginosa metabolites

Cell suspension in lysis solvent (0.625 OD/mL) was extracted with 1.5x volumes of water and cell debris removed by centrifugation. 330 μL cleared extract was mixed with 770 μL 50% methanol in acetonitrile (*v*/*v*), and the mixture was frozen at -40°C for 1hr. Any insoluble material was spun down at 4°C, 20000 *g* for 10 min. 1 mL supernatant (0.075 OD) was transferred to a fresh tube and solvent evaporated using a speedvac followed by a lyophilizer. The residual was reconstituted with 37.5μL water, and 4 uL was injected into a ZIC-pHILIC 150 x 2.1 mm (5 μm particle size) column (EMD Millipore). Analysis was conducted on a QExactive benchtop orbitrap mass spectrometer equipped with an Ion Max source and a HESI II probe, which was coupled to a Dionex UltiMate 3000 UPLC system (Thermo Fisher Scientific, San Jose, CA). External mass calibration was performed using the standard calibration mixture every 7 days. Chromatographic separation was achieved using the following conditions: Buffer A was 20 mM ammonium carbonate, 0.1% ammonium hydroxide; buffer B was acetonitrile. The column oven and autosampler tray were held at 25°C and 4°C, respectively. The chromatographic gradient was run at a flow rate of 0.150 ml/min as follows: 0–20

min.: linear gradient from 80% to 20% B; 20–20.5 min.: linear gradient from 20% to 80% B; 20.5–28 min.: hold at 80% B. The mass spectrometer was operated in full-scan, polarity switching mode with the spray voltage set to 3.0 kV, the heated capillary held at 275°C, and the HESI probe held at 350°C. The sheath gas flow was set to 40 units, the auxiliary gas flow was set to 15 units, and the sweep gas flow was set to 1 unit. The MS data acquisition was performed in a range of 70–1000 m/z, with the resolution set at 70,000, the AGC target at 10e6, and the maximum injection time at 20 msec. Relative quantitation of polar metabolites was performed with XCalibur QuanBrowser 2.2 (Thermo Fisher Scientific) using a 5 ppm mass tolerance and referencing an in-house library of chemical standards.

For relative quantifications, a peak area of $1.0x10^4$ was arbitrarily assigned to undetected metabolites. Then, peak area of each metabolite was normalized to the ISTD amino acid with the closest retention time and ionized by the same charge. Fold change of metabolite levels between conditions were then calculated based on normalized peak areas.

For absolute quantifications, standard samples containing AMP, ADP, ATP, GMP, GDP, GTP, pApp, ppApp, pppApp, IMP, UTP, dATP, dGTP, dCTP, dTTP and UDP-GlcNAc at a series of known concentrations were prepared in water containing 0.064% (v/v) ISTD and 4 μL was analyzed under the same conditions. Note that this ISTD concentration was identical to that in metabolome samples. Peak areas of all 16 purine nucleotides were therefore normalized to that of

$^{13}C_5{}^{15}N$-glutamate, and standard curves were generated. Absolute levels of the above nucleotides in unknown samples were then derived through interpolation.

Isothermal titration calorimetry (ITC)

All ITC experiments were performed in a VP-ITC (Malvern) instrument thermo-equilibrated at 25°C with water in the reference cell. Ligand solution is 1 mM (p)ppApp and 1 mM $MgCl_2$ in a buffer containing 20 mM HEPES-Na pH 7.4 and 150 mM NaCl, 2 mM $MgCl_2$ and 1 mM TCEP. Sample cell contains 100 μM $PurF^{EC}$ in the same buffer. ppApp-Mg was injected in 27 injections at 10 nmol/injection. Blank titrations were performed with protein-free gel filtration buffer in the sample cell. The blank-subtracted data were analyzed using the Origin software package (version 5.0, MicroCal, Inc.) and fit using a single-site binding model.

Microscopy

Phase contrast and propidium iodide (PI)-fluorescence images were taken on a Zeiss Observer Z1 microscope using a 100x/1.4 oil immersion objective and an LED-based Colibri illumination system using software Metamorph (Universal Imaging, PA). Cells were first washed with inducer-free medium and concentrated to OD_{600} = 0.5. 1 μL sample was spotted onto 1.5% agarose LB pads containing 2.5 μg/mL PI and incubated at 30°C. Time-lapse images were taken every 10 minutes over a 6-hour period.

Data Availability

All data supporting the findings of this study are available within the manuscript and associated appendix. X-ray crystallographic coordinates and structure factor files for $Tas1_{tox}$-Tis1, $PurF^{EC}$-ppApp and *Pa*SAH are available from the PDB: 6OX6, 6OTT, 6YVC, respectively. Structural data for *Bc*SAH and list of Tas1 homolog sequences are available upon request from Dr. John Whitney. Maximum likelihood estimates of *P. aeruginosa* strain relationships used for tree construction are provided in Newick format in Appendix Dataset A3.1.

Links to publisher websites for additional information:

https://www.nature.com/articles/s41586-019-1735-9#Sec45.

https://onlinelibrary.wiley.com/doi/10.1111/mmi.14684

CHAPTER IV – Conclusions and future directions

Overview

The T6SS is a sophisticated molecular machine that mediates interbacterial competition by directly injecting antibacterial effectors from a donor cell into a recipient cell. Effector proteins that transit this system target essential molecules and have diverse functions that can often act on a broad range of hosts (Russell et al., 2014a). The aim of this thesis was to characterize a widespread family of effectors that possess a PAAR domain fused to N- and C-terminal extensions of unknown function. Overall, my work highlights that the N-terminal regions possess domains that mediate effector trafficking between cells, while the C-terminus possesses a toxin domain that can inhibit target cell growth through a unique mechanism of action.

The second chapter of my thesis characterized domains present in the N-terminal extensions of these effectors. I show that in addition to the conserved PAAR domain, many of these effectors share a unique motif in their N-terminus, now referred to as prePAAR, that exhibits strict co-association with transmembrane domains (TMDs) and molecular chaperones called Eag proteins. prePAAR-containing effectors were categorized as containing either contain 1 TMD (class I) or 2 TMDs (class II). I used a combination of structural biology and biochemistry to demonstrate that Eag chaperones bind the TMDs of their cognate class I or class II prePAAR effectors, which I found is necessary for their stability, recruitment to the T6SS apparatus and transport into target cells. With the help of collaborators, I also showed that the TMDs of a specific effector permit insertion into liposomes

and allow for self-translocation of the C-terminal toxin domain into the liposome lumen.

In the third chapter of this thesis, I characterize a novel C-terminal toxin domain found within a prePAAR-containing effector from *P. aeruginosa* called Tas1. We find that this toxin domain resembles enzymes belonging to the ubiquitous RelA-SpoT homolog (RSH) family of proteins, which are known to synthesize the pro-survival signaling molecules (p)ppGpp during nutrient starvation. However, Tas1 does not synthesize (p)ppGpp and instead is a potent toxin that synthesizes (p)ppApp upon delivery into target cells. My work on Tas1 represents the first in-depth characterization of a (p)ppApp synthetase and demonstrates a physiological role for these nucleotides in bacteria. I have since extended my analysis of Tas1 by bioinformatically identifying hundreds of putative (p)ppApp synthetases that appear in genomic contexts outside of the T6SS.

In this chapter, I will be discussing outstanding questions that emerge from the work presented in the preceding chapters and providing initial experimental strategies based on the literature to address these questions in the future.

Chapter II summary and discussion

Why do different families of effectors require different families of molecular chaperones?

prePAAR effectors constitute a new family of T6SS effectors that are defined by the existence of a prePAAR motif, N-terminal TMDs, a PAAR domain and a C-terminal toxin domain. Most notably, we show that this group of effectors

strictly co-occur with Eag chaperones and that chaperone interaction with prePAAR effector TMDs is a conserved property of this protein family. While previous work has relied on genetic context to identify the cognate effector of an Eag chaperone (Alcoforado Diniz and Coulthurst, 2015; Whitney et al., 2015), our use of the prePAAR motif as an effector discovery tool enables the identification of these effectors in any genomic context. Other families of chaperones, such as the DUF4123 or DUF2169 protein families, have also been shown to affect the stability and/or export of their cognate effectors (Bondage et al., 2016; Burkinshaw et al., 2018; Pei et al., 2020). However, little is known about the specificity of these chaperones for their effector targets, which do not contain predicted N-terminal TMDs. DUF4123 chaperones are encoded next to effectors with diverse domain architectures and studies on several members of this family have shown chaperone interactions occur with domains of effectors possessing no apparent shared sequence properties (Liang et al., 2015). A lack of structural information for these and DUF2169 chaperones has hindered an understanding of why certain T6SS effectors require members of these chaperone families for export from the cell.

What is the structural basis for the role of prePAAR in binding PAAR?

Our structural characterization of Eag-TMD complexes provides important information about Eag specificity for their cognate effectors. However, this data also suggests that prePAAR is not a key component of the Eag binding site. Our

prePAAR mutagenesis experiments coupled with structural models suggest that the PAAR domain of prePAAR-containing effectors is unstable and that prePAAR likely facilitates proper folding of PAAR. However, we did not directly show that these two regions interact to form a properly folded PAAR domain. Future biophysical and structural studies will be necessary to quantify and visualize the binding of prePAAR to PAAR. Below, I propose several technical considerations that could be taken into account for these studies.

First, full-length effectors co-expressed with their cognate Eag proteins could be screened for crystallization. However, considering that class I effectors typically possess a large Rhs repeat-containing cage (> 100 kDa) and large proteins are often recalcitrant to crystal formation, this strategy may only be possible for class II effectors, which are significantly smaller in size. An alternative strategy would be to instead only use the prePAAR-TMD-PAAR region of an effector and express this region with its cognate Eag. A structure of this complex will not only reveal the molecular basis of PAAR and prePAAR binding but will also indicate the position of the Eag proteins in relation to the PAAR domain. This approach could be employed for both class I and class II effectors. A potential shortcoming of this strategy is that the PAAR domain of prePAAR effectors appears to exhibit instability when heterologously expressed. While the N-terminal extensions of PAAR will likely be stable in an Eag:prePAAR-TMD-PAAR complex, the C-terminal boundaries for PAAR will need to be estimated. Thus, many truncations across multiple prePAAR-containing effectors will need to be tested to

determine a suitable candidate for crystallography. Many effectors also possess multiple protein partners, which may hinder crystallization efforts. Therefore, another alternative strategy could be to construct artificial fusions of prePAAR and PAAR, which would circumvent problems associated with co-expressing other protein partners such as Eag chaperones. Regardless, structure determination will likely provide the most in-depth information regarding prePAAR-PAAR interactions and provide significant insight into the structural diversity of the PAAR domain, which has not been previously explored.

How do TMDs mediate toxin domain translocation?

Our work on the *P. aeruginosa* Tse6-VgrG1a-EagT6$_2$-EF-Tu complex revealed that the TMDs of the class II effector Tse6 spontaneously insert into liposomes, which permits self-translocation of the effector toxin domain into the lumen of the liposome. Based on this, we speculate that the TMDs of Tse6 and other related effectors likely insert into the inner membranes of recipient cells to permit toxin domain translocation into the cytoplasm. This mechanism may be similar to the mechanism previously described for AB toxins which involve 'threading' of an unfolded toxin domain through a transmembrane channel on the endosome used to enter the cytosol (Murphy, 2011).

Effectors associated with other pathways such as type I-V secretion systems typically bind and hijack a receptor on the recipient membrane to enter the recipient cytoplasm (Costa et al., 2015). However, to our knowledge, no

receptor has been identified to mediate T6SS effector entry into recipient cells. Our work indicates that effector TMDs are highly unstable in the absence of Eag chaperones and that these chaperones are not secreted with the effector, further supporting the proposed model where TMDs insert into the inner membrane of recipient cells following a T6SS firing event. However, this work is currently limited and further analyses should be conducted to test the validity of this model.

First, the membrane inserting properties of other prePAAR-containing effectors should be explored. Our work is based solely on the class II effector Tse6, which possesses a 17 kDa toxin domain. By contrast, class I effectors are significantly larger in size (> 100 kDa), possess a β-barrel cage that encapsulates the toxin and contain only a single TMD in their N-terminus. The release of the toxin domain from the cage is suspected to be triggered by an autoproteolysis event, however, this has not been shown in the context of a T6SS competition (Busby et al., 2013; Donato et al., 2020; Pei et al., 2020). If a toxin is released after entering the periplasm, it is possible it traverses the small pore made by the single TMD of the class I effectors, in a mechanism similar to that proposed for Tc toxins (Gatsogiannis et al., 2016). Structures of full-length effectors or their TMD-containing N-terminus in their membrane-inserted state will also be useful in providing insights into the organization of a toxin-translocating pore. The prePAAR-TMD-PAAR constructs described above may be useful for this structural work.

It is also important to explore the possibility that effectors located at the tip of VgrG proteins are delivered directly into the target cell cytoplasm during a T6SS

firing event. Microscopic methods to track effectors in transit have been with limited success in the field, likely owing to the single copy of each PAAR effector injected per firing event. However, previous studies have effectively tracked T6SS assembly using fluorescent imaging of high-copy structural components like the TssB/C sheath proteins (Borgeaud et al., 2015; Lin et al., 2019). Using this approach, a previous study suggested that the T6SS of *Vibrio cholera* cells directly inject structural components like Hcp and VgrG into the cytoplasm of recipient *V. cholera* cells lacking these components. This, in turn, enables the reuse of these structural components by recipients to assemble their own T6SSs (Vettiger and Basler, 2016). Other work showed that expressing a *V. cholera* peptidoglycan-targeting effector, VgrG3, in the cytoplasm of *E. coli* in the absence of its immunity protein, was toxic to cells (Dong et al., 2013). A follow-up study demonstrated that replacing the peptidoglycan-targeting toxin domain of VgrG3 with a DNA-targeting toxin domain that is active in the cytoplasm, did not impact the ability of the effector to inhibit target cell growth (Ho et al., 2017). All three of these studies suggest that *V. cholera* T6SSs secrete structural components and effectors directly into the recipient cytoplasm. Of note, our informatics work did not identify any prePAAR effectors in *Vibrio* species. Additionally, many studies have shown that cytoplasmic expression of T6SS effectors that are normally active in the periplasm is not toxic to cells (Russell et al., 2011; Whitney et al., 2013; Wood et al., 2019a). Thus, similar assays to those conducted in the above two studies may be explored for future work with prePAAR effectors in *Pseudomonas* and other genera identified in our

analyses. For example, a feasible experimental setup in *P. aeruginosa* would involve replacement of the toxin domain of an effector active in the periplasm, such as the recently identified peptidoglycan-targeting VgrG2b (Wood et al., 2019a), with the cytoplasmic toxin domains typically found in prePAAR effectors, such as the (p)ppApp synthetase domain described in this work (ex. VgrG2b-Tas1 fusion). These engineered effectors could be tested in competition assays to determine their effect on the growth of recipient cells and to infer effector localization following T6SS-dependent delivery. If these mutants are not toxic to recipient cells, the same competition assays may be conducted with donor strains expressing prePAAR effectors lacking their toxin domain, which would provide the effector TMDs into a recipient cell. This may improve the toxicity of the proposed VgrG2b-Tas1 effector or other periplasmic effectors with artificially fused toxin domains that are active in the cytoplasm, which would further support our hypotheses surrounding the role of the TMDs in type VI secretion.

Do all prePAAR-containing proteins contain C-terminal toxin domains that act in the cytoplasm?

Studies conducted in several different bacteria suggest that many T6SSs export multiple effectors during a single firing event (Cianfanelli et al., 2016a; Hood et al., 2010; Silverman et al., 2013). The precise subcellular location for effector delivery in recipient cells is not well understood, however, it is noteworthy that many effectors that interact with Hcp or C-terminal extensions of VgrG target periplasmic structures such as peptidoglycan or membranes (Brooks et al., 2013;

Flaugnatti et al., 2016; LaCourse et al., 2018; Silverman et al., 2013). By contrast, all characterized prePAAR proteins act on cytoplasmic targets by mechanisms that include the hydrolysis of NAD^+ and $NADP^+$, ADP-ribosylation of FtsZ, deamination of cytidine bases in double-stranded DNA and as shown in this work, pyrophosphorylation of ADP and ATP (Chapter III) (Mok et al., 2020; Ting et al., 2018; Whitney et al., 2015). This observation supports the proposal that the TMDs in prePAAR effectors function to promote toxin entry into the cytoplasm of target cells. It should be noted that PAAR effectors with nuclease activity that lack N-terminal TMDs have been identified, suggesting that other cell entry mechanisms likely exist, and future work may address whether these proteins have important motifs or domains that permit an alternative translocation mechanism into recipient cells (Jana et al., 2019; Pissaridou et al., 2018). Further characterization of other prePAAR-containing effectors will likely lead to the identification of other toxins with novel activities.

Chapter III summary and discussion

Does Tas1 have other interacting protein partners?

A key finding we discuss in chapter three is that Tas1 exhibits significant homology to previously characterized effector Tse6, which is found in a closely related *P. aeruginosa* strain to Tas1. We showed that both effectors are homologous at their N-terminus (prePAAR-TMD-PAAR region) but differ at their C-terminal toxin domains. The N-terminal homology suggests that both proteins bind

cognate Eag proteins and a cognate VgrG protein, while we showed that the C-terminal differences lead to distinct toxin functions and account for the ability of the two effectors to bind two different immunity proteins. However, it was previously shown that Tse6 also possesses a binding site for the essential housekeeping GTPase EF-Tu, which is required for interbacterial growth inhibition of recipient cells (Whitney et al., 2015). While we did study full-length Tas1 for metabolite profiling experiments, we did not determine whether this effector also has other protein partners beyond its cognate immunity. Our work shows that the EF-Tu binding site found in Tse6 is not conserved in Tas1, suggesting that these two proteins may use different protein partners to inhibit target cell growth. Pull-down experiments in *P. aeruginosa* using full-length Tas1 as bait will likely provide important insights into other protein interacting partners and their role in Tas1-mediated intoxication.

What is the molecular basis for adenosine specificity in Tas1 and other adenosine pyrophosphokinases?

Our characterization of Tas1 indicates that this toxin kills cells by rapidly synthesizing (p)ppApp. The rate of (p)ppApp synthesis by Tas1 is over 100-fold higher than any known (p)ppGpp synthetase and as shown in our metabolite profiling experiments, Tas1 specifically synthesizes (p)ppApp, not (p)ppGpp, in cells. Our preliminary characterization of other (p)ppApp synthetases suggests that these enzymes also exhibit very high rates of (p)ppApp synthesis. By comparing the sequences of Tas1 and known (p)ppApp synthetases to the conserved

(p)ppGpp synthetase RelA, we identified a number of key residues that may mediate the specificity for adenosine acceptor nucleotides over guanosine nucleotides. Future work should focus on two areas to understand more about adenosine specificity in the identified (p)ppApp synthetases: 1) structural characterization of substrate- (ATP/ADP) or product- (ppApp/pppApp) bound (p)ppApp synthetases and 2) in vitro and in vivo mutagenesis experiments of Tas1 and other identified (p)ppApp synthetases.

Several ligand-bound structures of (p)ppGpp synthetases have been solved and provide valuable insights into the mechanism of acceptor and donor nucleotide binding (Hogg et al., 2004; Manav et al., 2018; Steinchen et al., 2015). Some of these structures have been solved using non-hydrolysable ATP analog, AMPCPP, which may be similarly useful for structural characterization of (p)ppApp synthetases. The high rates of synthesis by Tas1 indicate that this enzyme may not be ideal for such experiments and using mutants with attenuated activity or testing the activity of other less potent synthetases in vitro will be a necessary first step prior to conducting crystallographic studies.

Our evolutionary analysis of (p)ppApp synthetases may also provide a strong basis for initial mutagenesis experiments. Conversion of (p)ppApp-specific residues in Tas1 to residues found in RelA may lead to an alteration of Tas1 specificity such that it is able to accommodate guanosine nucleotide acceptors. It is also possible that there exist more residues than the four predicted amino acids identified herein that mediate binding to adenosine acceptor nucleotides, thus,

high-throughput methods such as error-prone PCR may provide an alternative and unbiased strategy for the analysis of Tas1 specificity. This method may also allow for rapid in vivo screening of mutants. It was previously shown that minimal medium supplemented with serine, methionine and glycine (SMG) is a growth-permissive condition for (p)ppGpp producing cells and growth-inhibitory for non-(p)ppGpp producers (Δ*relA*) (Nanamiya et al., 2008; Uzan and Danchin, 1976). Thus, a library of Tas1 variants generated through error-prone PCR could be transformed into *E. coli* Δ*relA* strains and tested for growth on SMG plates, which may provide strong negative selection for Tas1 mutants that synthesize (p)ppGpp at high enough concentrations to activate the stringent response.

Are there other targets for (p)ppApp in cells?

Our metabolite profiling experiments revealed that the rapid synthesis of (p)ppApp by Tas1 has multiple consequences for cellular metabolism. First, accumulation of (p)ppApp leads to a depletion of ATP and ADP and subsequently, dysregulation of several metabolic pathways. While ATP depletion affects ATP-dependent processes in central metabolism, we also show that (p)ppApp directly inhibits PurF, an essential enzyme involved in purine biosynthesis. PurF converts phosphoribosylpyrophosphate to 5'-phosphoribosylamine, which is a generic building block for purines. Notably, the PurF enzyme is a target that is shared between (p)ppGpp and (p)ppApp and our data suggest that (p)ppApp inhibits PurF to a similar extent and in a similar manner to (p)ppGpp. This work could be followed

by mutagenesis experiments where mutated '(p)ppApp-blind' PurF-expressing cells could be tested in competition assays to determine if this target is of relevance during bacterial competition.

During starvation, the concentration of (p)ppGpp rapidly increases from micromolar to millimolar levels, leading to the inhibition of over 50 protein targets in diverse cellular pathways including DNA replication, transcription, translation and nucleotide metabolism (Wang et al., 2019). Based on our findings surrounding (p)ppApp-dependent inhibition of purine biosynthesis, it is possible that (p)ppApp also inhibits other (p)ppGpp targets in vivo. Indeed, (p)ppApp was shown, like (p)ppGpp, to bind RNA polymerase in *E. coli*, however the consequences of this binding on transcription are not known (Bruhn-Olszewska et al., 2018).

In the case of Tas1, (p)ppApp accumulation coincides with a significant drop in ATP, however, our preliminary assays suggest that the other (p)ppApp synthetases identified do not have this effect on the cellular ATP pool. Metabolite profiling and transcriptomics experiments in cells expressing these enzymes will likely provide further insights into the effects of (p)ppApp on different pathways and its protein targets. Previously, biologically relevant protein targets were identified using affinity-based capture of protein targets using (p)ppGpp as bait (Wang et al., 2019). Similar assays could also be used for (p)ppApp to identify the range of targets for these nucleotides in different bacteria.

Are there role(s) for (p)ppApp synthetases beyond interbacterial competition?

Our work identified hundreds of putative (p)ppApp synthetases, many of which are not found to be associated with T6SSs or other pathways that mediate interbacterial antagonism. Three Tas1 homologs used for our experiments are found in a MuF domain-containing prophage (MuF phage), *Streptomyces* species and *Vibrio* species. Our preliminary work on these proteins suggests that all three proteins are toxic to *E. coli* and like Tas1, produce (p)ppApp in vitro.

Several MuF domain phage-associated toxins were previously identified using bioinformatics (Jamet et al., 2017). This informatics study along with other studies assessing the biochemistry and genetics of MuF domain-containing proteins suggest that these proteins may be involved in any one general functions: 1) phage infection, 2) interbacterial competition or 3) defense against phage infection.

MuF domain-containing proteins are minor phage capsid proteins that have elusive roles in phage biology (Jamet et al., 2017). The most well-studied MuF protein is gp7 from *B. subtilis* phage SPP1, which lacks a C-terminal toxin domain (Stiege et al., 2003; Vinga et al., 2006). Deletion of gp7 from the SPP1 phage does not affect phage assembly but does cause a significant reduction in phage infectivity (Vinga et al., 2006). Biochemical analyses showed that gp7 binds the portal protein, gp6, which facilitates recruitment of 2-3 copies of gp7 to the phage head (Dröge et al., 2000; Stiege et al., 2003). In addition to interacting with gp6, the gp7 protein binds DNA and has been proposed to be delivered with the phage

DNA into the cytoplasm of target cells during phage infection (Stiege et al., 2003). The evidence above strongly suggests a role for these proteins in phage infection, however as mentioned above, gp7 does not have a toxin domain that is often fused to MuF domain-containing proteins (Jamet et al., 2017). If the MuF domain is important for phage infection, what are the role(s) of the toxin domain? And by extension, what are the role for the associated immunity proteins? It is possible that the MuF-fused toxins are injected with phage DNA into the host cell to facilitate infection, which is analogous to protein toxin injection previously described for phage T4 (Bossi et al., 2003; Moak and Molineux, 2004). Future work studying these toxins should start by assessing the impact of toxin deletions or catalytically inactive mutants of the toxin on phage infection. The SPP1 phage lifestyle, its receptor and host are well-established and this phage may be a good heterologous system used to express and isolate SPP1 containing MuF toxins from other phage (Sao-Jose et al., 2006). However, it is possible that other interacting partners, such as a cognate portal protein described above, may also need to be included in this heterologous system to ensure proper phage assembly (Stiege et al., 2003). Using this system may also provide a way to study the role of immunity proteins associated with MuF toxins. For example, it should be determined if a packaged phage requires immunity proteins to be present in the capsid with the toxin prior to host infection or if these proteins are only necessary for phage assembly in the phage-encoding bacterium. The effects of MuF phage immunity proteins could be

studied by expressing these proteins in a host cell and assaying the efficiency of phage infection by toxin-containing phage.

Like T6SS effectors, MuF domain-containing toxins are diverse in function and are predicted to target essential molecules in the cell. In all cases, including the (p)ppApp synthetase described in our work, MuF toxins are encoded adjacent to predicted immunity genes (Jamet et al., 2017). Based on this and the fact that some bacteria activate prophage release in the presence of competing cells (Bossi et al., 2003), Jamet and colleagues hypothesized that MuF-associated toxins may be involved in interbacterial competition. In this case, lysogenic populations may produce MuF phage that inhibit target population growth by injecting DNA and the MuF-associated toxins into host cells, inhibiting their growth and preventing lysogeny of the phage into competitors. Other (p)ppApp synthetases have also been identified in toxin-antitoxin (TA) systems and within other prophages (Dedrick et al., 2017; Jimmy et al., 2020b). Recent work also shows that some TA systems provide defense against phage (Guegler and Laub, 2021). Thus, it is also possible that (p)ppApp synthetases associated with MuF phages, other prophages or TA systems play a role in preventing phage infection. Understanding the basis of MuF-domain containing phage expression, whether the toxins are carried in the phage and injected into recipients and whether the presence of phage provides a competitive advantage to the bacteria that harbor them should be explored in the future.

The Tas1 homologs identified in *Vibrio* and *Streptomyces* both possess N-terminal Sec signal sequences, which likely direct these proteins to the periplasm or extracellular milieu, respectively. As described in chapter three, the *Vibrio* homolog has been previously annotated as a (p)ppGpp synthetase called RelV (Das et al., 2009; Dasgupta et al., 2014). However, our work demonstrates that this toxin readily synthesizes (p)ppApp, not (p)ppGpp, in vitro. The T2SS of *Vibrio* species is known to secrete many periplasmic proteins important for the lifestyle of these bacteria into the extracellular milieu, including cholera toxin (Rivera-Chávez and Mekalanos, 2019). Future work should consider examining whether the identified *Vibrio* (p)ppApp synthetases are secreted in a T2SS-dependent manner.

What is the role of the *Vibrio* (p)ppApp synthetases? Because of their lack of an apparent immunity protein and because these synthetases are found in several pathogenic species of *Vibrio* in our dataset, we hypothesized that they may be associated with virulence. Several studies have characterized the transcriptional profiles of *Vibrio* species under various laboratory conditions and in the context of a host to better understand the genetic basis for their pathogenicity. However, no study for either *V. cholera* or *V. parahaemolyticus* showed differential expression of (p)ppApp synthetase genes when comparing expression in rich versus minimal medium (Mandlik et al., 2011), during infection (Das et al., 2000; García et al., 2017; Livny et al., 2014; Mandlik et al., 2011; Rivera-Chávez and Mekalanos, 2019) or have associated them with the regulon for biofilm formation (Papenfort et al., 2015) or the T6SS (Dong and Mekalanos, 2012). Together this

suggests that these enzymes are not likely to function as virulence factors, or (p)ppGpp synthetases, which are active under minimal medium (Irving et al., 2021), or T6SS effectors.

The above studies do suggest that the (p)ppApp synthetases from *V. cholera* and *V. parahaemolyticus* are constitutively expressed under standard laboratory conditions (Livny et al., 2014; Mandlik et al., 2011). Therefore, in addition to studying their secretion, early studies may also benefit from determining whether these proteins have additional protein interaction partners, which can be assayed through chromosomal tagging of the synthetase-encoding gene and conducting pull-down experiments. Isogenic mutants should also be compared to wild-type to determine if deletion of this gene affects growth or other aspects of *Vibrio* species lifestyles, such as interbacterial competition (Fridman et al., 2020). Linking this gene to a phenotype may be challenging, however, the general assays described above may provide a starting point for more sophisticated assays in the future.

Previous work showed that culture supernatants isolated from various *Streptomyces* species possessed (p)ppApp-synthesizing activity (Nishino and Murao, 1975; Oki et al., 1976). Many of these species were identified in our informatics search and possess predicted Sec signal sequence-containing (p)ppApp synthetases, providing some evidence that these enzymes are secreted and are active in the culture supernatants. *Streptomyces* species have sophisticated developmental lifestyles and are known to secrete many proteins and

signaling factors that regulate this process (Flärdh and Buttner, 2009; Zhou et al., 2021). During the transition from vegetative growth to the production of aerial mycelium, *Streptomyces* undergo periods of cell lysis, which may release ATP into the environment (McCormick and Flärdh, 2012). Therefore, extracellular (p)ppApp synthetases may produce (p)ppApp during specific growth stages to regulate *Streptomyces* growth. Previous work suggests that (p)ppApp specifically inhibits spore germination in these bacteria, further supporting the idea that these nucleotides are involved in regulating *Streptomyces* differentiation (Hamagishi et al., 1980). Our analyses identified multiple genetically tractable *Streptomyces* species that could be subject to straightforward assays to determine the role of these (p)ppApp synthetases in *Streptomyces* physiology. Important first steps to characterizing these proteins should validate previous findings by assessing the impact of (p)ppApp on bacterial differentiation through microscopy, determine if these proteins are secreted and measuring (p)ppApp synthesis in culture supernatants. In addition to their complex lifestyle, *Streptomyces* have been a historical source of antibiotics and often require signals for the secretion of these molecules that are often linked to specific growth phases of these organisms (Chater et al., 2010; Procópio et al., 2012; Zhang et al., 2020). The potential role of (p)ppApp in affecting *Streptomyces* differentiation suggest that it could also impact antibiotic secretion. This could be explored using metabolomic and secretomic analyses of strains grown in the presence and absence of (p)ppApp or that differentially express a (p)ppApp synthetase (Xu et al., 2020). Another

potential way to explore the role of (p)ppApp on antibiotic secretion is through competition assays between *Streptomyces* and susceptible recipient bacteria such as *Bacillus subtilis* (Culp et al., 2019; Gehrke et al., 2019). Using *Streptomyces* strains expressing a wild-type and catalytically inactive (p)ppApp synthetase for these assays may provide insights into the impact of these enzymes on the fitness of these bacteria during interbacterial competition.

Our work collectively suggests that (p)ppApp synthetases are widespread in bacteria and thus are not strictly limited to functioning as T6SS effectors. Unlike (p)ppGpp synthetases, (p)ppApp synthetases appear to more often be secreted from cells, either by transport through a secretion system or through association with phages. Many questions remain about the role of these synthetases in bacteria, however, key experiments should broadly explore the roles of these proteins in interbacterial competition, understand their biochemical and structural diversity and characterize their regulon in their native contexts.

Concluding Remarks

This work has explored T6SS effector transport and toxicity for a broadly distributed family of effector proteins. However, several other effector families exist and likely use distinct mechanisms from those described herein to mediate trafficking and toxicity (Zhang et al., 2012). Exploring these mechanisms using in-depth structural, biochemical and genetic analyses promises to provide novel insights into the functioning of the T6SS. We also demonstrate an example of the

unique toxic functions associated with T6SS effectors with the discovery of a (p)ppApp synthetase. Indeed, many other toxins also possess novel functions (de Moraes et al., 2021; Tang et al., 2018; Ting et al., 2018). A complete understanding of the function of these proteins provides a platform for bioengineering efforts that could be used in the treatment of disease (Mok et al., 2020) or as shown with (p)ppApp, may help identify other families of non-T6SS proteins with similar functions that have previously undescribed roles in bacteria.

APPENDIX

Chapter II Appendix

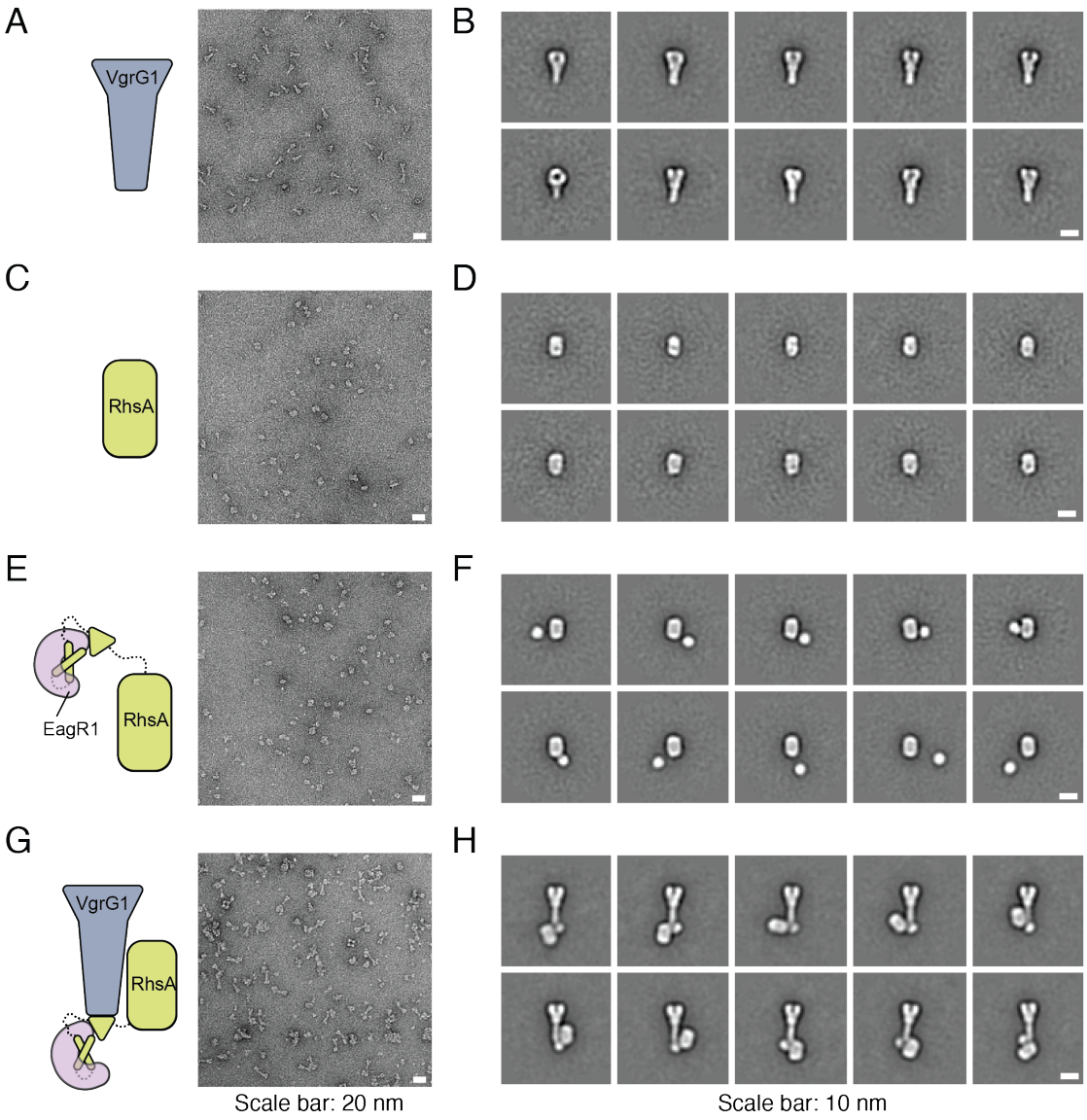


Figure A2.1. RhsA, EagR1 and VgrG1 form a ternary complex in vitro. Unprocessed micrographs **(A, C, E, G)** and representative 2-D class averages **(B, D, F, H)** of negatively stained VgrG1 **(A, B)**, $RhsA_{\Delta NT}$ **(C, D)**, EagR1-RhsA complex **(E, F)** and EagR1-RhsA-VgrG1 complex **(G, H)**. Scale bar represents 20 nm for unprocessed micrographs and 10 nm for class averages.

Chapter III Appendix

Dataset A3.1. Maximum likelihood estimate of *P. aeruginosa* strain relationships constructed using the RAxML-HPC Blackbox. Provided in Newick format below.

((AZPAE14700_2310_Tas1_1:0.066311,(((BL18_544_Tas1_1:1.89E-4,WH-SGI-V-07070_3342_Tas1_1:4.2E-5):0.062915,(AZPAE14393_2359_Tas1_4:1.32E-4,AZPAE14359_2351_Tas1_4:1.66E-4):0.064376):0.012115,((((AZPAE13850_2339_Tas1_1:2.78E-4,AZPAE14726_2315_Tas1_1:4.86E-4):0.067856,AZPAE15004_2231_Tse6_6:0.056785):0.010267,(AZPAE14395_2361_Tas1_1:0.065157,(ATCC_700888_679_Tse6_10:0.002942,279_PAER_3097_Tse6_10:0.002261):0.059863):0.010508):0.004075,((((AZPAE14404_2364_Tas1_1:0.057368,681_PAER_2876_Tse6_6:0.065767):0.012607,(((WH-SGI-V-07293_3581_Tas1_2:5.26E-4,AZPAE14905_2542_Tas1_2:3.35E-4):0.014407,MRSN_317_2627_Tas1_2:0.014801):0.044745,WH-SGI-V-07380_3391_Tas1_1:0.065655):0.008978):0.005045,((PS75_2063_Tas1_7:0.0582,(AZPAE14930_2275_Tas1_1:0.065813,(((PDR_2511_Tas1_10:0.064091,(AZPAE14957_2529_Tas1_3:0.039667,WH-SGI-V-07420_3533_Tas1_2:0.0401):0.045827):0.244298,(WH-SGI-V-07050_3348_Tse6_22:0.039299,CF_PA39_212_Tse6_15:0.030601):0.854349):0.05214,((((((AZPAE14919_2516_Tas1_1:0.063229,WH-SGI-V-07692_3503_Tse6_2:0.061981):0.008759,((((((((((WH-SGI-V-07053_3393_Tse6_2:0.048016,((PAG_2088_Tse6_2:1.96E-4,(MH19_2858_Tse6_2:0.013663,1152_PAER_2895_Tse6_2:0.001011):2.8E-5):0.007529,(AZPAE14715_2390_Tse6_2:0.002545,14651_3307_Tse6_2:8.07E-4):0.010209):0.034547):0.005194,((AZPAE14820_2329_Tse6_2:0.04513,AZPAE15028_2209_Tse6_4:0.042829):0.00651,(M9A_1_532_Tse6_13:0.043968,((AZPAE14828_2336_Tse6_17:0.036232,(_3246_Tse6_2:0.022891,((AZPAE12136_2383_Tse6_2:0.003206,138244_305_Tse6_2:0.009946):0.008828,(CF27_600_Tse6_2:0.027681,AZPAE15014_2241_Tse6_2:0.012411):0.007429):0.005881):0.009587):0.010098,(AZPAE15055_2251_Tse6_2:0.037334,WH-SGI-V-07710_3565_Tas1_1:0.041429):0.007447):0.004324):0.003038):0.004306):0.002981,((XMG_378_Tse6_2:0.0401,AZPAE14931_2271_Tse6_2:0.041751):0.00586,((WH-SGI-V-07695_3560_Tse6_2:0.028862,C7447m_494_Tse6_23:0.034104):0.009871,AZPAE15056_2250_Tse6_2:0.044662):0.009784):0.003618):0.002633,(((MRSN_321_2628_Tas1_1:0.044157,BWH056_2068_Tse6_2:0.042186):0.006449,((AZPAE14876_2130_Tse6_2:0.025419,14650_3305_Tse6_2:0.023023):0.025194,((AZPAE14910_2112_Tse6_2:0.013057,AZPAE14873_2127_Tse6_2:0.008485):0.024598,((AZPAE13872_2559_Tse6_2:0.018653,WH-SGI-V-07372_3385_Tse6_4:0.026242):0.007551,BL07_555_Tse6_11:0.02553):0.012343):0.008245):0.004817):0.001841,((AZPAE14943_2169_Tse6_2:0.030813,AZPAE14995_2222_Tse6_2:0.038335):0.016947,(((1079_PAER_2864_Tas1_3:0.039462,WH-SGI-V-07071_3340_Tse6_2:0.033575):0.009427,(WH-SGI-V-07488_3458_Tse6_14:0.036881,(AZPAE14891_2148_Tse6_2:0.001856,AZPAE12410_2404_Tse6_2:0.001857):0.028215):0.01089):0.003593,(1336_PAER_2896_Tse6_2:0.037086,BWHPSA008_582_Tse6_2:0.043988):0.009802):0.001964):0.002495):0.002566):0.00121,(((BWHPSA013_577_Tse6_8:0.045355,BWHPSA016_574_Tse6_21:0.044731):0.007014,(WH-SGI-V-07620_3541_Tse6_2:0.048579,3575_2059_Tas1_3:0.042791):0.009841):0.00283,

((((AZPAE14691_2303_Tse6_2:0.021275,(AZPAE12151_2397_Tse6_2:0.003272,AZP
AE14884_2140_Tse6_2:0.002736):0.031191):0.025541,AZPAE14918_2093_Tse6_3:
0.043567):0.007327,((BWHPSA042_2037_Tse6_2:0.043553,(AZPAE14443_2372_Tse
6_2:0.023925,(AZPAE14817_2326_Tse6_5:0.025415,((AZPAE15022_2248_Tse6_2:0
.028428,AZPAE15023_2249_Tse6_5:0.023133):0.003252,(AZPAE14731_2553_Tse6_
5:0.018431,(576_PAER_2874_Tse6_2:0.012923,AZPAE15057_2252_Tse6_5:0.01391
7):0.001423):0.001387):0.003663):0.009347):0.015012):0.006635,(BWH029_20
52_Tse6_2:0.011926,WH-SGI-V-
07300_3586_Tse6_2:0.01487):0.0298):0.004445):0.002444,(((AZPAE14917_2266
_Tse6_2:0.047217,WH-SGI-V-07375_3387_Tse6_2:0.043643):0.005885,WH-SGI-V-
07490_3460_Tse6_2:0.043806):0.002804,(AZPAE14913_2274_Tas1_1:0.046223,AZ
PAE14852_2106_Tse6_5:0.048929):0.006672):0.002233):0.002216):0.001134):0
.001337,(((AZPAE14697_2307_Tas1_1:0.038838,AZPAE12142_2387_Tse6_2:0.0443
48):0.009937,(C41_528_Tse6_2:0.03986,RB_670_Tse6_2:0.035884):0.013186):0
.003681,(((AZPAE12145_2391_Tse6_2:0.048562,M8A_4_533_Tse6_2:0.043127):0.
006386,(AZPAE14934_2277_Tse6_2:0.043704,WH-SGI-V-
07187_3365_Tse6_2:0.042529):0.006435):0.001809,((WH-SGI-V-
07623_3542_Tse6_2:0.035522,(AZPAE12156_2401_Tas1_1:0.036353,AZPAE14847_2
102_Tas1_12:0.029906):0.009526):0.013323,((3578_2056_Tse6_2:0.038011,AZP
AE14848_2104_Tse6_2:0.043653):0.007124,(LESlike4_669_Tse6_2:1.61E-
4,JD304_2012_Tse6_2:2.28E-
4):0.039411):0.004486):0.003124):0.001525):0.001519):0.002491,((((AZPAE1
4975_2202_Tse6_2:0.044074,WH-SGI-V-
07428_3452_Tse6_2:0.041366):0.004773,(H1l_2007_Tas1_13:0.036155,(BL10_55
2_Tse6_12:0.02892,(BWHPSA004_586_Tse6_2:0.003606,((991_rep1_PAER_2886_Ts
e6_2:4.25E-4,792_PAER_2884_Tse6_2:4.25E-4):1.0E-
6,BWHPSA012_578_Tse6_2:0.00389):0.001853):0.03252):0.008671):0.007136):0
.005147,((AZPAE14999_2226_Tas1_1:0.048992,(AZPAE14996_2223_Tas1_3:0.0379
16,WH-SGI-V-
07491_3461_Tse6_2:0.038631):0.008893):0.009452,(AZPAE14550_2296_Tse6_2:0
.035022,AZPAE14907_2268_Tse6_2:0.029943):0.017579):0.001998):0.002052,((
(ATCC_33349_3310_Tas1_2:0.046144,AZPAE14810_2320_Tse6_2:0.043485):0.0039
95,(PAK_2043_Tse6_2:0.038833,WH-SGI-V-
07226_3369_Tse6_2:0.039932):0.017011):0.002477,((WH-SGI-V-
07233_3412_Tse6_2:0.048802,AZPAE12419_2414_Tse6_2:0.038269):0.01153,((DR
1_2575_Tse6_2:0.059165,(M8A_2_535_Tas1_3:0.039495,(C48_527_Tse6_16:0.037
883,AZPAE14809_2319_Tse6_2:0.034111):0.007091):0.009107):0.005684,(CF5_5
99_Tse6_2:0.042792,AZPAE15065_2538_Tse6_2:0.042903):0.010647):0.003613):
0.001983):0.002121):0.002276):0.001364,(((((WH-SGI-V-
07317_3594_Tse6_2:0.047396,(AZPAE14855_2109_Tas1_2:0.029354,ATCC_33358_3
328_Tse6_2:0.029593):0.019331):0.004679,((AZPAE15048_2536_Tse6_3:0.03032
4,BTP034_3111_Tse6_2:0.030979):0.021549,(AZPAE14441_2370_Tas1_1:0.050112
,(AZPAE13864_2343_Tse6_2:0.095949,(AZPAE14535_2379_Tse6_3:7.04E-
4,(((AZPAE14453_2373_Tse6_3:5.53E-4,WH-SGI-V-07325_3598_Tse6_3:9.8E-
5):1.0E-6,(P1_London_28_IMP_1_04_05_3113_Tse6_3:1.62E-4,WH-SGI-V-
07259_3570_Tse6_3:9.8E-5):1.0E-6):1.0E-6,AZPAE14415_2366_Tse6_3:4.88E-
4):4.36E-
4):0.044973):0.016708):0.00386):0.00184):0.002213,(((AZPAE14533_2567_Tse
6_2:0.023034,(AZPAE14895_2153_Tse6_3:0.022762,_3249_Tse6_3:0.011823):0.0
10686):0.019924,AZPAE14977_2205_Tse6_3:0.040799):0.014532,(AZPAE15052_21
75_Tse6_3:0.051258,_3295_Tse6_2:0.046254):0.005513):0.002233):0.002584,(
(((ATCC_43732_3332_Tas1_5:0.030314,(ATCC_33363_3325_Tas1_5:3.99E-
4,VRFPA06_274_Tas1_5:6.15E-
4):0.030521):0.020151,(AZPAE13866_2558_Tas1_2:0.045083,(WH-SGI-V-

07310_3589_Tse6_2:0.032863,(BL06_556_Tas1_3:0.014778,((AZPAE14986_2214_T
as1_1:0.019134,(AZPAE14813_2323_Tse6_2:0.020009,(ATCC_43390_3327_Tse6_2:
0.004808,((WH-SGI-V-07635_3483_Tse6_2:9.7E-
5,AZPAE14923_2162_Tse6_2:6.5E-5):1.0E-6,WH-SGI-V-07484_3454_Tse6_2:2.6E-
4):0.004659):0.011221):0.001691):0.00174,AZPAE12154_2400_Tas1_2:0.016996
):0.004065):0.014646):0.018466):0.010543):0.003788,((660_PAER_2899_Tse6_
2:0.045621,AZPAE14838_2547_Tse6_4:0.043399):0.007811,((AZPAE14410_2365_T
as1_2:0.048865,(AZPAE15001_2228_Tse6_2:0.022862,AZPAE12422_2416_Tse6_2:0
.030964):0.017686):0.005971,((DQ8_381_Tse6_4:0.001419,WH-SGI-V-
07178_3359_Tse6_4:3.67E-
4):0.035757,(AZPAE15053_2138_Tse6_Perfect:0.006465,ATCC_33352_3314_Tse6_
Perfect:0.006947):0.032757):0.01203):0.003899):0.004845):0.002536,(((O40
0_2625_Tse6_2:3.49E-4,AZPAE14956_2528_Tse6_2:2.67E-
4):0.040685,(BL19_543_Tse6_2:0.029767,(WH-SGI-V-07389_3427_Tse6_6:2.03E-
4,(WH-SGI-V-07404_3440_Tse6_6:3.65E-4,(AZPAE14701_2550_Tse6_6:3.26E-
4,WH-SGI-V-07267_3528_Tse6_6:9.8E-5):2.6E-5):9.7E-
5):0.031778):0.014951):0.009937,(IGB83_2077_Tse6_8:0.045187,((859_PAER_2
880_Tse6_4:1.31E-4,957_PAER_2892_Tse6_4:9.7E-
5):0.039625,AZPAE14402_2362_Tas1_6:0.042911):0.007036):0.003689):0.00270
5):0.001945):0.002265,(((AZPAE12411_2405_Tse6_2:0.044923,BWHPSA018_572_T
se6_2:0.047708):0.003631,(WH-SGI-V-
07708_3602_Tas1_2:0.031802,(AZPAE14964_2190_Tse6_3:1.62E-
4,0C2E_3623_Tse6_3:2.58E-
4):0.033308):0.01477):0.005095,((((AZPAE14947_2283_Tse6_4:0.048533,PASS1
_3205_Tas1_3:0.048315):0.008119,((AZPAE14877_2132_Tse6_2:0.029349,SMC438
6_3621_Tse6_2:0.024159):0.022086,BWHPSA020_570_Tse6_2:0.046826):0.003936
):0.002784,(((WH-SGI-V-
07060_3333_Tse6_2:0.041289,MSH_10_2488_Tse6_2:0.049164):0.007901,(AZPAE1
4868_2123_Tse6_5:0.016691,WH-SGI-V-
07282_3530_Tse6_5:0.016284):0.026858):0.00364,(AZPAE14864_2119_Tse6_2:0.
046091,(ATCC_33356_3321_Tse6_2:0.04186,UM-
01_2646_Tse6_2:0.042022):0.00608):0.002763):0.004037):0.003216,(((AZPAE1
4916_2281_Tse6_2:0.038396,(WH-SGI-V-
07485_3455_Tse6_2:0.033928,AZPAE14954_2182_Tas1_1:0.029988):0.009751):0.
009947,(AZPAE15036_2097_Tse6_8:0.056311,(B3-
1811_2011_Tse6_4:0.001528,(BWH031_2050_Tse6_2:6.79E-
4,1268_PAER_2861_Tse6_2:3.61E-4):7.24E-
4):0.038757):0.012594):0.004142,((AZPAE14936_2264_Tse6_2:0.04379,_3247_T
se6_Perfect:0.046115):0.004804,(UDL_597_Tse6_2:0.040055,(AZPAE12417_2412
_Tse6_2:0.027579,PASS4_3204_Tse6_2:0.033899):0.014593):0.005616):0.00287
7):0.002264):0.001028):0.001935):8.84E-
4):0.001691,(((((AZPAE12140_2385_Tse6_2:0.012898,JD303_2497_Tse6_24:0.01
9209):0.022102,(WH-SGI-V-07279_3573_Tse6_2:0.034866,WH-SGI-V-
07373_3418_Tse6_2:0.025456):0.01938):0.028234,3577_2057_Tas1_2:0.048688)
:0.005472,ATCC_14886_678_Tse6_3:0.052947):0.003294,(BWHPSA047_2032_Tse6_
2:0.050333,(((3573_2061_Tse6_2:0.001757,545_PAER_2873_Tse6_2:0.004742):0
.00236,WH-SGI-V-
07067_3396_Tse6_2:0.004498):0.049059,AZPAE14946_2174_Tse6_2:0.044711):0.
004177):0.003637):0.003424):0.001895,(JD312_2014_Tse6_11:0.056393,(AZPAE
14992_2219_Tse6_2:0.049249,(AZPAE12152_2398_Tse6_3:0.046429,AZPAE14352_2
349_Tse6_5:0.062123):0.009069):0.003236):0.004393):0.001551,(((_3231_Tse
6_3:0.049614,(BL12_550_Tse6_20:0.030959,(AZPAE14878_2522_Tas1_2:0.084319
,_3238_Tse6_2:0.003561):0.036987):0.01747):0.005076,(((AZPAE14882_2137_T
se6_2:0.028199,(WH-SGI-V-

07228_3370_Tse6_2:0.011543,(AZPAE13877_2346_Tse6_2:0.015793,18A_661_Tse6
_2:0.018607):0.014917):0.004969):0.0157,AZPAE14526_2378_Tse6_2:0.023224)
:0.023191,AZPAE14939_2541_Tse6_2:0.054014):0.009867):0.004017,((((JD318_
2498_Tse6_18:0.059206,719_PAER_2877_Tse6_2:0.069517):0.009552,(((AZPAE14
862_2517_Tas1_3:0.002977,AZPAE14937_2543_Tas1_3:0.001596):0.048711,WH-
SGI-V-
07638_3548_Tas1_3:0.055183):0.014002,(AZPAE12146_2392_Tse6_9:0.059142,(A
ZPAE12416_2410_Tas1_2:0.018134,AZPAE14899_2158_Tas1_2:0.020236):0.039066
):0.005217):0.003863):0.003773,(AZPAE15062_2255_Tse6_4:0.030583,BL01_561
_Tse6_9:0.02499):0.03317):0.003619,((WH-SGI-V-
07422_3448_Tse6_2:0.045663,AZPAE14825_2334_Tse6_9:0.061551):0.009386,(AZ
PAE14538_2380_Tse6_2:0.047475,AZPAE14908_2272_Tse6_2:0.049925):0.010959)
:0.005763):0.00208):0.001893):0.002482):0.001729,(BWH052_2072_Tse6_2:0.0
19983,AZPAE14963_2189_Tse6_19:0.022152):0.037953):0.002922,((WH-SGI-V-
07410_3442_Tas1_1:0.007121,(WH-SGI-V-07501_3537_Tas1_3:0.010276,WH-SGI-
V-
07266_3574_Tas1_1:0.006669):0.004781):0.040191,AZPAE14840_2515_Tas1_6:0.
047748):0.012286):0.006097,AZPAE14856_2111_Tse6_2:0.055437):0.00671,(_32
39_Tse6_5:0.061358,(AZPAE14903_2525_Tas1_1:0.078892,AZPAE14967_2192_Tas1
_1:0.060991):0.017774):0.009614):0.005495):0.013387):0.041043):0.074929,
(((WH-SGI-V-07426_3451_Tas1_1:0.069237,((((WH-SGI-V-
07632_3547_Tse6_7:0.061342,BWHPSA043_2036_Tse6_7:0.059632):0.010053,(AZP
AE14373_2355_Tas1_1:0.062596,(BL04_558_Tse6_7:0.056725,WH-SGI-V-
07384_3422_Tas1_1:0.060377):0.009119):0.005689):0.006795,AZPAE14982_2210
_Tse6_7:0.069726):0.006765,(WH-SGI-V-
07252_3520_Tas1_1:0.071165,_3248_Tas1_1:0.078225):0.01028):0.022654):0.0
23773,WH-SGI-V-07314_3592_Tas1_1:0.071077):0.018669,WH-SGI-V-
07286_3578_Tas1_3:0.064031):0.00643):0.007587):0.0062,((((AZPAE14865_212
0_Tas1_4:3.66E-4,WH-SGI-V-07243_3379_Tas1_4:3.0E-5):0.056248,WH-SGI-V-
07707_3513_Tas1_11:0.063281):0.012164,(((BWH049_2075_Tas1_3:0.007422,AZP
AE14988_2216_Tas1_3:0.009497):0.054356,AZPAE14372_2561_Tas1_2:0.058092):
0.010311,(AZPAE14858_2114_Tas1_8:0.05507,(AZPAE14398_2565_Tse6_Perfect:0
.012639,AZPAE14911_2280_Tse6_Perfect:0.012091):0.038494):0.015612):0.005
674):0.005786,(((AZPAE14972_2199_Tse6_6:0.06778,(WH-SGI-V-
07239_3414_Tas1_1:0.003819,AZPAE14442_2371_Tas1_1:0.03089):0.057397):0.0
08687,((953_PAER_2885_Tse6_2:0.002387,(230_PAER_2866_Tse6_2:3.3E-
4,(927_PAER_2891_Tse6_2:3.96E-4,(462_PAER_2871_Tse6_2:1.65E-
4,PA103_2044_Tse6_2:3.3E-5):1.0E-6):1.65E-4):5.64E-
4):0.05548,BL15_547_Tas1_1:0.065388):0.012053):0.00562,(((AZPAE15027_216
4_Tse6_2:0.062239,AZPAE14920_2270_Tas1_1:0.05965):0.012497,BTP038_3108_T
as1_3:0.071375):0.005357,(ATCC_33364_3326_Tas1_1:0.061035,BWHPSA038_2041
_Tas1_9:0.059425):0.011736):0.003415):0.003835):0.004025):0.003813):0.00
5061):0.008322):0.063705,WH-SGI-V-
07698_3562_Tas1_1:0.006628,_3229_Tas1_1:0.00487);

REFERENCES

1. Adamian, L., and Liang, J. (2002). Interhelical hydrogen bonds and spatial motifs in membrane proteins: polar clamps and serine zippers. Proteins *47*, 209-218.
2. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica Section D, Biological crystallography *66*, 213-221.
3. Alcoforado Diniz, J., and Coulthurst, S.J. (2015). Intraspecies Competition in Serratia marcescens Is Mediated by Type VI-Secreted Rhs Effectors and a Conserved Effector-Associated Accessory Protein. J Bacteriol *197*, 2350-2360.
4. Allsopp, L.P., Wood, T.E., Howard, S.A., Maggiorelli, F., Nolan, L.M., Wettstadt, S., and Filloux, A. (2017). RsmA and AmrZ orchestrate the assembly of all three type VI secretion systems in Pseudomonas aeruginosa. Proc Natl Acad Sci U S A *114*, 7707-7712.
5. Aloulou, A., Ali, Y.B., Bezzine, S., Gargouri, Y., and Gelb, M.H. (2012). Phospholipases: an overview. Methods Mol Biol *861*, 63-85.
6. Aoki, S.K., Diner, E.J., de Roodenbeke, C.t.K., Burgess, B.R., Poole, S.J., Braaten, B.A., Jones, A.M., Webb, J.S., Hayes, C.S., Cotter, P.A.*, et al.* (2010). A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. Nature *468*, 439-442.
7. Aoki, S.K., Pamma, R., Hernday, A.D., Bickham, J.E., Braaten, B.A., and Low, D.A. (2005). Contact-Dependent Inhibition of Growth in Escherichia coli. Science *309*, 1245-1248.
8. Aravind, L., and Koonin, E.V. (1998). The HD domain defines a new superfamily of metal-dependent phosphohydrolases. Trends Biochem Sci *23*, 469-472.
9. Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res *44*, W344-350.
10. Atkinson, G.C., Tenson, T., and Hauryliuk, V. (2011). The RelA/SpoT homolog (RSH) superfamily: distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life. PLoS One *6*, e23479.
11. Bai, X.C., McMullan, G., and Scheres, S.H. (2015). How cryo-EM is revolutionizing structural biology. Trends Biochem Sci *40*, 49-57.
12. Ballister, E.R., Lai, A.H., Zuckermann, R.N., Cheng, Y., and Mougous, J.D. (2008). In vitro self-assembly of tailorable nanotubes from a simple protein building block. Proc Natl Acad Sci U S A *105*, 3733-3738.

13. Barker, M.M., Gaal, T., Josaitis, C.A., and Gourse, R.L. (2001). Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. J Mol Biol *305*, 673-688.
14. Basler, M., Ho, B.T., and Mekalanos, J.J. (2013). Tit-for-tat: type VI secretion system counterattack during bacterial cell-cell interactions. Cell *152*, 884-894.
15. Basler, M., Pilhofer, M., Henderson, G.P., Jensen, G.J., and Mekalanos, J.J. (2012). Type VI secretion requires a dynamic contractile phage tail-like structure. Nature *483*, 182-186.
16. Bayer-Santos, E., Cenens, W., Matsuyama, B.Y., Oka, G.U., Di Sessa, G., Mininel, I.D.V., Alves, T.L., and Farah, C.S. (2019). The opportunistic pathogen Stenotrophomonas maltophilia utilizes a type IV secretion system for interbacterial killing. PLoS Pathog *15*, e1007651.
17. Baynham, P.J., Ramsey, D.M., Gvozdyev, B.V., Cordonnier, E.M., and Wozniak, D.J. (2006). The Pseudomonas aeruginosa ribbon-helix-helix DNA-binding protein AlgZ (AmrZ) controls twitching motility and biogenesis of type IV pili. J Bacteriol *188*, 132-140.
18. Beljantseva, J., Kudrin, P., Andresen, L., Shingler, V., Atkinson, G.C., Tenson, T., and Hauryliuk, V. (2017). Negative allosteric regulation of Enterococcus faecalis small alarmone synthetase RelQ by single-stranded RNA. Proc Natl Acad Sci U S A *114*, 3726-3731.
19. Bi, E.F., and Lutkenhaus, J. (1991). FtsZ ring structure associated with division in Escherichia coli. Nature *354*, 161-164.
20. Bondage, D.D., Lin, J.S., Ma, L.S., Kuo, C.H., and Lai, E.M. (2016). VgrG C terminus confers the type VI effector transport specificity and is required for binding with PAAR and adaptor-effector complex. Proc Natl Acad Sci U S A *113*, E3931-3940.
21. Bonemann, G., Pietrosiuk, A., Diemand, A., Zentgraf, H., and Mogk, A. (2009). Remodelling of VipA/VipB tubules by ClpV-mediated threading is crucial for type VI protein secretion. EMBO J *28*, 315-325.
22. Borgeaud, S., Metzger, L.C., Scrignari, T., and Blokesch, M. (2015). The type VI secretion system of Vibrio cholerae fosters horizontal gene transfer. Science *347*, 63-67.
23. Bossi, L., Fuentes, J.A., Mora, G., and Figueroa-Bossi, N. (2003). Prophage contribution to bacterial population dynamics. J Bacteriol *185*, 6467-6471.
24. Brooks, T.M., Unterweger, D., Bachmann, V., Kostiuk, B., and Pukatzki, S. (2013). Lytic activity of the Vibrio cholerae type VI secretion toxin VgrG-3 is inhibited by the antitoxin TsaB. The Journal of biological chemistry *288*, 7618-7625.
25. Bruhn-Olszewska, B., Molodtsov, V., Sobala, M., Dylewski, M., Murakami, K.S., Cashel, M., and Potrykus, K. (2018). Structure-function comparisons of (p)ppApp vs (p)ppGpp for Escherichia coli RNA polymerase binding sites and for rrnB P1 promoter regulatory responses in vitro. Biochim Biophys Acta Gene Regul Mech *1861*, 731-742.

26. Bugg, T.D., Braddick, D., Dowson, C.G., and Roper, D.I. (2011). Bacterial cell wall assembly: still an attractive antibacterial target. Trends Biotechnol *29*, 167-173.
27. Bunduc, C.M., Bitter, W., and Houben, E.N.G. (2020). Structure and Function of the Mycobacterial Type VII Secretion Systems. Annual Review of Microbiology *74*, 315-335.
28. Bunkoczi, G., Echols, N., McCoy, A.J., Oeffner, R.D., Adams, P.D., and Read, R.J. (2013). Phaser.MRage: automated molecular replacement. Acta Crystallogr D Biol Crystallogr *69*, 2276-2286.
29. Burkinshaw, B.J., Liang, X., Wong, M., Le, A.N.H., Lam, L., and Dong, T.G. (2018). A type VI secretion system effector delivery mechanism dependent on PAAR and a chaperone-co-chaperone complex. Nat Microbiol *3*, 632-640.
30. Busby, J.N., Panjikar, S., Landsberg, M.J., Hurst, M.R., and Lott, J.S. (2013). The BC component of ABC toxins is an RHS-repeat-containing protein encapsulation device. Nature *501*, 547-550.
31. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics *10*, 421.
32. Cao, Z., Casabona, M.G., Kneuper, H., Chalmers, J.D., and Palmer, T. (2016). The type VII secretion system of Staphylococcus aureus secretes a nuclease toxin that targets competitor bacteria. Nat Microbiol *2*, 16183.
33. Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972-1973.
34. Cardona, S.T., and Valvano, M.A. (2005). An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in Burkholderia cenocepacia. Plasmid *54*, 219-228.
35. Castang, S., McManus, H.R., Turner, K.H., and Dove, S.L. (2008). H-NS family members function coordinately in an opportunistic pathogen. Proc Natl Acad Sci U S A *105*, 18947-18952.
36. Chater, K.F., Biró, S., Lee, K.J., Palmer, T., and Schrempf, H. (2010). The complex extracellular biology of Streptomyces. FEMS Microbiol Rev *34*, 171-198.
37. Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr *66*, 12-21.
38. Cherrak, Y., Rapisarda, C., Pellarin, R., Bouvier, G., Bardiaux, B., Allain, F., Malosse, C., Rey, M., Chamot-Rooke, J., Cascales, E.*, et al.* (2018). Biogenesis and structure of a type VI secretion baseplate. Nat Microbiol *3*, 1404-1416.
39. Chou, S., Bui, N.K., Russell, A.B., Lexa, K.W., Gardiner, T.E., LeRoux, M., Vollmer, W., and Mougous, J.D. (2012). Structure of a peptidoglycan amidase

effector targeted to Gram-negative bacteria by the type VI secretion system. Cell Rep *1*, 656-664.
40. Cianfanelli, F.R., Alcoforado Diniz, J., Guo, M., De Cesare, V., Trost, M., and Coulthurst, S.J. (2016a). VgrG and PAAR Proteins Define Distinct Versions of a Functional Type VI Secretion System. PLoS pathogens *12*, e1005735.
41. Cianfanelli, F.R., Monlezun, L., and Coulthurst, S.J. (2016b). Aim, Load, Fire: The Type VI Secretion System, a Bacterial Nanoweapon. Trends Microbiol *24*, 51-62.
42. Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., Doron, S., Amitai, G., and Sorek, R. (2019). Cyclic GMP-AMP signalling protects bacteria against viral infection. Nature *574*, 691-695.
43. Cornforth, D.M., and Foster, K.R. (2013). Competition sensing: the social side of bacterial stress responses. Nat Rev Microbiol *11*, 285-293.
44. Costa, T.R., Felisberto-Rodrigues, C., Meir, A., Prevost, M.S., Redzej, A., Trokter, M., and Waksman, G. (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. Nat Rev Microbiol *13*, 343-359.
45. Culp, E.J., Yim, G., Waglechner, N., Wang, W., Pawlowski, A.C., and Wright, G.D. (2019). Hidden antibiotics in actinomycetes can be identified by inactivation of gene clusters for common antibiotics. Nat Biotechnol *37*, 1149-1154.
46. Curran, A.R., and Engelman, D.M. (2003). Sequence motifs, polar interactions and conformational changes in helical membrane proteins. Current opinion in structural biology *13*, 412-417.
47. Das, B., Pal, R.R., Bag, S., and Bhadra, R.K. (2009). Stringent response in Vibrio cholerae: genetic analysis of spoT gene function and identification of a novel (p)ppGpp synthetase gene. Mol Microbiol *72*, 380-398.
48. Das, S., Chakrabortty, A., Banerjee, R., Roychoudhury, S., and Chaudhuri, K. (2000). Comparison of global transcription responses allows identification of Vibrio cholerae genes differentially expressed following infection. FEMS Microbiol Lett *190*, 87-91.
49. Das, S., and Chaudhuri, K. (2003). Identification of a unique IAHP (IcmF associated homologous proteins) cluster in Vibrio cholerae and other proteobacteria through in silico analysis. In Silico Biol *3*, 287-300.
50. Dasgupta, S., Basu, P., Pal, R.R., Bag, S., and Bhadra, R.K. (2014). Genetic and mutational characterization of the small alarmone synthetase gene relV of Vibrio cholerae. Microbiology (Reading) *160*, 1855-1866.
51. Dawson, J.P., Melnyk, R.A., Deber, C.M., and Engelman, D.M. (2003). Sequence context strongly modulates association of polar residues in transmembrane helices. J Mol Biol *331*, 255-262.
52. Dawson, J.P., Weinger, J.S., and Engelman, D.M. (2002). Motifs of serine and threonine can drive association of transmembrane helices. J Mol Biol *316*, 799-805.
53. De Kwaadsteniet, M., Fraser, T., Van Reenen, C.A., and Dicks, L.M. (2006). Bacteriocin T8, a novel class IIa sec-dependent bacteriocin produced by

Enterococcus faecium T8, isolated from vaginal secretions of children infected with human immunodeficiency virus. Appl Environ Microbiol *72*, 4761-4766.

54. de Moraes, M.H., Hsu, F., Huang, D., Bosch, D.E., Zeng, J., Radey, M.C., Simon, N., Ledvina, H.E., Frick, J.P., Wiggins, P.A.*, et al.* (2021). An interbacterial DNA deaminase toxin directly mutagenizes surviving target populations. Elife *10*.
55. Dedrick, R.M., Jacobs-Sera, D., Bustamante, C.A., Garlena, R.A., Mavrich, T.N., Pope, W.H., Reyes, J.C., Russell, D.A., Adair, T., Alvey, R.*, et al.* (2017). Prophage-mediated defence against viral attack and viral counter-defence. Nat Microbiol *2*, 16251.
56. Desvaux, M., Hébraud, M., Talon, R., and Henderson, I.R. (2009). Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. Trends Microbiol *17*, 139-145.
57. Dillard, J.P., and Seifert, H.S. (2001). A variable genetic island specific for Neisseria gonorrhoeae is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates. Mol Microbiol *41*, 263-277.
58. Donato, S.L., Beck, C.M., Garza-Sánchez, F., Jensen, S.J., Ruhe, Z.C., Cunningham, D.A., Singleton, I., Low, D.A., and Hayes, C.S. (2020). The β-encapsulation cage of rearrangement hotspot (Rhs) effectors is required for type VI secretion. Proc Natl Acad Sci U S A *117*, 33540-33548.
59. Dong, T.G., Ho, B.T., Yoder-Himes, D.R., and Mekalanos, J.J. (2013). Identification of T6SS-dependent effector and immunity proteins by Tn-seq in Vibrio cholerae. Proc Natl Acad Sci U S A *110*, 2623-2628.
60. Dong, T.G., and Mekalanos, J.J. (2012). Characterization of the RpoN regulon reveals differential regulation of T6SS and new flagellar operons in Vibrio cholerae O37 strain V52. Nucleic Acids Res *40*, 7766-7775.
61. Dröge, A., Santos, M.A., Stiege, A.C., Alonso, J.C., Lurz, R., Trautner, T.A., and Tavares, P. (2000). Shape and DNA packaging activity of bacteriophage SPP1 procapsid: protein components and interactions during assembly. J Mol Biol *296*, 117-132.
62. Dudley, E.G., Thomson, N.R., Parkhill, J., Morin, N.P., and Nataro, J.P. (2006). Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative Escherichia coli. Mol Microbiol *61*, 1267-1282.
63. Edwards, J.S., and Palsson, B.O. (2000). The Escherichia coli MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. Proc Natl Acad Sci U S A *97*, 5528-5533.
64. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.
65. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

66. Eshraghi, A., Kim, J., Walls, A.C., Ledvina, H.E., Miller, C.N., Ramsey, K.M., Whitney, J.C., Radey, M.C., Peterson, S.B., Ruhland, B.R.*, et al.* (2016). Secreted Effectors Encoded within and outside of the Francisella Pathogenicity Island Promote Intramacrophage Growth. Cell Host Microbe *20*, 573-583.
67. Flärdh, K., and Buttner, M.J. (2009). Streptomyces morphogenetics: dissecting differentiation in a filamentous bacterium. Nat Rev Microbiol *7*, 36-49.
68. Flaugnatti, N., Le, T.T., Canaan, S., Aschtgen, M.S., Nguyen, V.S., Blangy, S., Kellenberger, C., Roussel, A., Cambillau, C., Cascales, E.*, et al.* (2016). A phospholipase A1 antibacterial Type VI secretion effector interacts directly with the C-terminal domain of the VgrG spike protein for delivery. Mol Microbiol *99*, 1099-1118.
69. Flemming, H.-C., and Wuertz, S. (2019). Bacteria and archaea on Earth and their abundance in biofilms. Nature Reviews Microbiology *17*, 247-260.
70. Folkesson, A., Löfdahl, S., and Normark, S. (2002). The Salmonella enterica subspecies I specific centisome 7 genomic island encodes novel protein families present in bacteria living in close contact with eukaryotic cells. Res Microbiol *153*, 537-545.
71. Fridman, C.M., Keppel, K., Gerlic, M., Bosis, E., and Salomon, D. (2020). A comparative genomics methodology reveals a widespread family of membrane-disrupting T6SS effectors. Nat Commun *11*, 1085.
72. Fung, D.K., Yang, J., Stevenson, D.M., Amador-Noguez, D., and Wang, J.D. (2020). Small Alarmone Synthetase SasA Expression Leads to Concomitant Accumulation of pGpp, ppApp, and AppppA in Bacillus subtilis. Front Microbiol *11*, 2083.
73. Gaca, A.O., Kudrin, P., Colomer-Winter, C., Beljantseva, J., Liu, K., Anderson, B., Wang, J.D., Rejman, D., Potrykus, K., Cashel, M.*, et al.* (2015). From (p)ppGpp to (pp)pGpp: Characterization of Regulatory Effects of pGpp Synthesized by the Small Alarmone Synthetase of Enterococcus faecalis. J Bacteriol *197*, 2908-2919.
74. Galán, J.E., and Waksman, G. (2018). Protein-Injection Machines in Bacteria. Cell *172*, 1306-1318.
75. García, K., Yáñez, C., Plaza, N., Peña, F., Sepúlveda, P., Pérez-Reytor, D., and Espejo, R.T. (2017). Gene expression of Vibrio parahaemolyticus growing in laboratory isolation conditions compared to those common in its natural ocean environment. BMC Microbiol *17*, 118.
76. Garcia-Bayona, L., and Comstock, L.E. (2018). Bacterial antagonism in host-associated microbial communities. Science *361*.
77. Garcia-Bayona, L., Gozzi, K., and Laub, M.T. (2019). Mechanisms of Resistance to the Contact-Dependent Bacteriocin CdzC/D in Caulobacter crescentus. J Bacteriol *201*.

78. Garcia-Bayona, L., Guo, M.S., and Laub, M.T. (2017). Contact-dependent killing by Caulobacter crescentus via cell surface-associated, glycine zipper proteins. Elife *6*.
79. Gatsogiannis, C., Merino, F., Prumbaum, D., Roderer, D., Leidreiter, F., Meusch, D., and Raunser, S. (2016). Membrane insertion of a Tc toxin in near-atomic detail. Nature Structural & Molecular Biology *23*, 884-890.
80. Gehrke, E.J., Zhang, X., Pimentel-Elardo, S.M., Johnson, A.R., Rees, C.A., Jones, S.E., Hindra, Gehrke, S.S., Turvey, S., Boursalie, S.*, et al.* (2019). Silencing cryptic specialized metabolism in Streptomyces by the nucleoid-associated protein Lsr2. Elife *8*.
81. Gentry, D.R., and Cashel, M. (1996). Mutational analysis of the Escherichia coli spoT gene identifies distinct but overlapping regions involved in ppGpp synthesis and degradation. Mol Microbiol *19*, 1373-1384.
82. Goodman, A.L., Kulasekara, B., Rietsch, A., Boyd, D., Smith, R.S., and Lory, S. (2004). A signaling network reciprocally regulates genes associated with acute infection and chronic persistence in Pseudomonas aeruginosa. Dev Cell *7*, 745-754.
83. Granato, E.T., Meiller-Legrand, T.A., and Foster, K.R. (2019). The Evolution and Ecology of Bacterial Warfare. Current biology : CB *29*, R521-R537.
84. Guegler, C.K., and Laub, M.T. (2021). Shutoff of host transcription triggers a toxin-antitoxin system to cleave phage RNA and abort infection. Mol Cell.
85. Haas, T.M., Ebensperger, P., Eisenbeis, V.B., Nopper, C., Durr, T., Jork, N., Steck, N., Jessen-Trefzer, C., and Jessen, H.J. (2019). Magic spot nucleotides: tunable target-specific chemoenzymatic synthesis. Chem Commun (Camb) *55*, 5339-5342.
86. Hamagishi, Y., Tone, H., Oki, T., and Inui, T. (1980). Effect of adenosine-5'-triphosphate-3'-diphosphate and related nucleoside polyphosphates on the spore germination of Streptomyces galilaeus. Arch Microbiol *125*, 285-289.
87. Hauryliuk, V., Atkinson, G.C., Murakami, K.S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of (p)ppGpp in bacterial physiology. Nature reviews Microbiology *13*, 298-309.
88. Hauser, M., Steinegger, M., and Söding, J. (2016). MMseqs software suite for fast and deep clustering and searching of large protein sequence sets. Bioinformatics *32*, 1323-1330.
89. Hersch, S.J., Watanabe, N., Stietz, M.S., Manera, K., Kamal, F., Burkinshaw, B., Lam, L., Pun, A., Li, M., Savchenko, A.*, et al.* (2020). Envelope stress responses defend against type six secretion system attacks independently of immunity proteins. Nature Microbiology *5*, 706-714.
90. Hibbing, M.E., Fuqua, C., Parsek, M.R., and Peterson, S.B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. Nat Rev Microbiol *8*, 15-25.
91. Hmelo, L.R., Borlee, B.R., Almblad, H., Love, M.E., Randall, T.E., Tseng, B.S., Lin, C., Irie, Y., Storek, K.M., Yang, J.J.*, et al.* (2015). Precision-

engineering the Pseudomonas aeruginosa genome with two-step allelic exchange. Nature protocols *10*, 1820-1841.
92. Ho, B.T., Fu, Y., Dong, T.G., and Mekalanos, J.J. (2017). Vibrio cholerae type 6 secretion system effector trafficking in target bacterial cells. Proc Natl Acad Sci U S A *114*, 9427-9432.
93. Hofreuter, D., Odenbreit, S., and Haas, R. (2001). Natural transformation competence in Helicobacter pylori is mediated by the basic components of a type IV secretion system. Mol Microbiol *41*, 379-391.
94. Hogg, T., Mechold, U., Malke, H., Cashel, M., and Hilgenfeld, R. (2004). Conformational antagonism between opposing active sites in a bifunctional RelA/SpoT homolog modulates (p)ppGpp metabolism during the stringent response [corrected]. Cell *117*, 57-68.
95. Hood, R.D., Singh, P., Hsu, F., Guvener, T., Carl, M.A., Trinidad, R.R., Silverman, J.M., Ohlson, B.B., Hicks, K.G., Plemel, R.L.*, et al.* (2010). A type VI secretion system of Pseudomonas aeruginosa targets a toxin to bacteria. Cell Host Microbe *7*, 25-37.
96. Hu, H., Zhang, H., Gao, Z., Wang, D., Liu, G., Xu, J., Lan, K., and Dong, Y. (2014). Structure of the type VI secretion phospholipase effector Tle1 provides insight into its hydrolysis and membrane targeting. Acta Crystallogr D Biol Crystallogr *70*, 2175-2185.
97. Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics *11*, 119.
98. Irving, S.E., Choudhury, N.R., and Corrigan, R.M. (2021). The stringent response and physiological roles of (pp)pGpp in bacteria. Nat Rev Microbiol *19*, 256-271.
99. Jamet, A., Touchon, M., Ribeiro-Gonçalves, B., Carriço, J.A., Charbit, A., Nassif, X., Ramirez, M., and Rocha, E.P.C. (2017). A widespread family of polymorphic toxins encoded by temperate phages. BMC Biol *15*, 75.
100. Jana, B., Fridman, C.M., Bosis, E., and Salomon, D. (2019). A modular effector with a DNase domain and a marker for T6SS substrates. Nat Commun *10*, 3595.
101. Jimmy, S., Saha, C.K., Kurata, T., Stavropoulos, C., Oliveira, S.R.A., Koh, A., Cepauskas, A., Takada, H., Rejman, D., Tenson, T.*, et al.* (2020a). A widespread toxin-antitoxin system exploiting growth control via alarmone signaling. Proc Natl Acad Sci U S A.
102. Jimmy, S., Saha, C.K., Kurata, T., Stavropoulos, C., Oliveira, S.R.A., Koh, A., Cepauskas, A., Takada, H., Rejman, D., Tenson, T.*, et al.* (2020b). A widespread toxin-antitoxin system exploiting growth control via alarmone signaling. Proc Natl Acad Sci U S A *117*, 10500-10510.
103. Johnson, L.S., Eddy, S.R., and Portugaly, E. (2010). Hidden Markov model speed heuristic and iterative HMM search procedure. BMC bioinformatics *11*, 431.

104. Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G.*, et al.* (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics *30*, 1236-1240.
105. Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.
106. Käll, L., Krogh, A., and Sonnhammer, E.L. (2007). Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. Nucleic Acids Res *35*, W429-432.
107. Kanonenberg, K., Schwarz, C.K., and Schmitt, L. (2013). Type I secretion systems - a story of appendices. Res Microbiol *164*, 596-604.
108. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol *30*, 772-780.
109. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc *10*, 845-858.
110. Kim, S., Jeon, T.J., Oberai, A., Yang, D., Schmidt, J.J., and Bowie, J.U. (2005). Transmembrane glycine zippers: physiological and pathological roles in membrane proteins. Proc Natl Acad Sci U S A *102*, 14278-14283.
111. Klein, T.A., Ahmad, S., and Whitney, J.C. (2020). Contact-Dependent Interbacterial Antagonism Mediated by Protein Secretion Machines. Trends Microbiol *28*, 387-400.
112. Koboldt, D.C., Steinberg, K.M., Larson, D.E., Wilson, R.K., and Mardis, E.R. (2013). The next-generation sequencing revolution and its impact on genomics. Cell *155*, 27-38.
113. Koskiniemi, S., Lamoureux, J.G., Nikolakakis, K.C., t'Kint de Roodenbeke, C., Kaplan, M.D., Low, D.A., and Hayes, C.S. (2013). Rhs proteins from diverse bacteria mediate intercellular competition. Proceedings of the National Academy of Sciences of the United States of America *110*, 7032-7037.
114. Krampen, L., Malmsheimer, S., Grin, I., Trunk, T., Luhrmann, A., de Gier, J.W., and Wagner, S. (2018). Revealing the mechanisms of membrane protein export by virulence-associated bacterial secretion systems. Nature communications *9*, 3467.
115. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol *305*, 567-580.
116. LaCourse, K.D., Peterson, S.B., Kulasekara, H.D., Radey, M.C., Kim, J., and Mougous, J.D. (2018). Conditional toxicity and synergy drive diversity among antibacterial effectors. Nat Microbiol *3*, 440-446.
117. Le, N.-H., Peters, K., Espaillat, A., Sheldon, J.R., Gray, J., Di Venanzio, G., Lopez, J., Djahanschiri, B., Mueller, E.A., Hennon, S.W.*, et al.* (2020). Peptidoglycan editing provides immunity to Acinetobacter baumannii during bacterial warfare. Science Advances *6*, eabb5614.

118. Lee, D.G., Urbach, J.M., Wu, G., Liberati, N.T., Feinbaum, R.L., Miyata, S., Diggins, L.T., He, J., Saucier, M., Deziel, E.*, et al.* (2006a). Genomic analysis reveals that Pseudomonas aeruginosa virulence is combinatorial. Genome Biol *7*, R90.
119. Lee, P.A., Tullman-Ercek, D., and Georgiou, G. (2006b). The bacterial twin-arginine translocation pathway. Annu Rev Microbiol *60*, 373-395.
120. Leiman, P.G., Basler, M., Ramagopal, U.A., Bonanno, J.B., Sauder, J.M., Pukatzki, S., Burley, S.K., Almo, S.C., and 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*, 4154-4159.
121. Leo, J.C., Grin, I., and Linke, D. (2012). Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. Philos Trans R Soc Lond B Biol Sci *367*, 1088-1101.
122. LeRoux, M., Kirkpatrick, R.L., Montauti, E.I., Tran, B.Q., Peterson, S.B., Harding, B.N., Whitney, J.C., Russell, A.B., Traxler, B., Goo, Y.A.*, et al.* (2015). Kin cell lysis is a danger signal that activates antibacterial pathways of Pseudomonas aeruginosa. Elife *4*.
123. Lewis, K.N., Liao, R., Guinn, K.M., Hickey, M.J., Smith, S., Behr, M.A., and Sherman, D.R. (2003). Deletion of RD1 from Mycobacterium tuberculosis mimics bacille Calmette-Guerin attenuation. The Journal of infectious diseases *187*, 117-123.
124. Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics *22*, 1658-1659.
125. Liang, X., Moore, R., Wilton, M., Wong, M.J., Lam, L., and Dong, T.G. (2015). Identification of divergent type VI secretion effectors using a conserved chaperone domain. Proc Natl Acad Sci U S A *112*, 9106-9111.
126. Lin, L., Lezan, E., Schmidt, A., and Basler, M. (2019). Abundance of bacterial Type VI secretion system components measured by targeted proteomics. Nat Commun *10*, 2584.
127. Livny, J., Zhou, X., Mandlik, A., Hubbard, T., Davis, B.M., and Waldor, M.K. (2014). Comparative RNA-Seq based dissection of the regulatory networks and environmental stimuli underlying Vibrio parahaemolyticus gene expression during infection. Nucleic Acids Res *42*, 12212-12223.
128. Lu, D., Shang, G., Zhang, H., Yu, Q., Cong, X., Yuan, J., He, F., Zhu, C., Zhao, Y., Yin, K.*, et al.* (2014a). Structural insights into the T6SS effector protein Tse3 and the Tse3-Tsi3 complex from Pseudomonas aeruginosa reveal a calcium-dependent membrane-binding mechanism. Mol Microbiol *92*, 1092-1112.
129. Lu, D., Zheng, Y., Liao, N., Wei, L., Xu, B., Liu, X., and Liu, J. (2014b). The structural basis of the Tle4-Tli4 complex reveals the self-protection mechanism of H2-T6SS in Pseudomonas aeruginosa. Acta Crystallogr D Biol Crystallogr *70*, 3233-3243.

130. Ma, J., Pan, Z., Huang, J., Sun, M., Lu, C., and Yao, H. (2017). The Hcp proteins fused with diverse extended-toxin domains represent a novel pattern of antibacterial effectors in type VI secretion systems. Virulence *8*, 1189-1202.
131. Ma, L.S., Hachani, A., Lin, J.S., Filloux, A., and Lai, E.M. (2014). Agrobacterium tumefaciens deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta. Cell Host Microbe *16*, 94-104.
132. Magnusson, L.U., Farewell, A., and Nyström, T. (2005). ppGpp: a global regulator in Escherichia coli. Trends Microbiol *13*, 236-242.
133. Mahairas, G.G., Sabo, P.J., Hickey, M.J., Singh, D.C., and Stover, C.K. (1996). Molecular analysis of genetic differences between Mycobacterium bovis BCG and virulent M. bovis. J Bacteriol *178*, 1274-1282.
134. Manav, M.C., Beljantseva, J., Bojer, M.S., Tenson, T., Ingmer, H., Hauryliuk, V., and Brodersen, D.E. (2018). Structural basis for (p)ppGpp synthesis by the Staphylococcus aureus small alarmone synthetase RelP. J Biol Chem *293*, 3254-3264.
135. Mandlik, A., Livny, J., Robins, W.P., Ritchie, J.M., Mekalanos, J.J., and Waldor, M.K. (2011). RNA-Seq-based monitoring of infection-linked changes in Vibrio cholerae gene expression. Cell Host Microbe *10*, 165-174.
136. Marden, J.N., Diaz, M.R., Walton, W.G., Gode, C.J., Betts, L., Urbanowski, M.L., Redinbo, M.R., Yahr, T.L., and Wolfgang, M.C. (2013). An unusual CsrA family member operates in series with RsmA to amplify posttranscriptional responses in Pseudomonas aeruginosa. Proc Natl Acad Sci U S A *110*, 15055-15060.
137. Mariano, G., Trunk, K., Williams, D.J., Monlezun, L., Strahl, H., Pitt, S.J., and Coulthurst, S.J. (2019). A family of Type VI secretion system effector proteins that form ion-selective pores. Nat Commun *10*, 5484.
138. McCormick, J.R., and Flärdh, K. (2012). Signals and regulators that govern Streptomyces development. FEMS Microbiol Rev *36*, 206-231.
139. McGinness, K.E., Baker, T.A., and Sauer, R.T. (2006). Engineering controllable protein degradation. Mol Cell *22*, 701-707.
140. McKenna, R., and Nielsen, D.R. (2011). Styrene biosynthesis from glucose by engineered E. coli. Metab Eng *13*, 544-554.
141. Minor, W., Cymborowski, M., Otwinowski, Z., and 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*, 859-866.
142. Miyata, S.T., Unterweger, D., Rudko, S.P., and Pukatzki, S. (2013). Dual expression profile of type VI secretion system immunity genes protects pandemic Vibrio cholerae. PLoS Pathog *9*, e1003752.
143. Moak, M., and Molineux, I.J. (2004). Peptidoglycan hydrolytic activities associated with bacteriophage virions. Mol Microbiol *51*, 1169-1183.

144. Mok, B.Y., de Moraes, M.H., Zeng, J., Bosch, D.E., Kotrys, A.V., Raguram, A., Hsu, F., Radey, M.C., Peterson, S.B., Mootha, V.K.*, et al.* (2020). A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. Nature *583*, 631-637.
145. Moriya, T., Saur, M., Stabrin, M., Merino, F., Voicu, H., Huang, Z., Penczek, P.A., Raunser, S., and Gatsogiannis, C. (2017). High-resolution Single Particle Analysis from Electron Cryo-microscopy Images Using SPHIRE. J Vis Exp.
146. Mougous, J.D., Cuff, M.E., Raunser, S., Shen, A., Zhou, M., Gifford, C.A., Goodman, A.L., Joachimiak, G., Ordonez, C.L., Lory, S.*, et al.* (2006). A virulence locus of Pseudomonas aeruginosa encodes a protein secretion apparatus. Science *312*, 1526-1530.
147. Mougous, J.D., Gifford, C.A., Ramsdell, T.L., and Mekalanos, J.J. (2007). Threonine phosphorylation post-translationally regulates protein secretion in Pseudomonas aeruginosa. Nature cell biology *9*, 797-803.
148. Murphy, J.R. (2011). Mechanism of diphtheria toxin catalytic domain delivery to the eukaryotic cell cytosol and the cellular factors that directly participate in the process. Toxins (Basel) *3*, 294-308.
149. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta crystallographica Section D, Biological crystallography *53*, 240-255.
150. Nadell, C.D., Drescher, K., and Foster, K.R. (2016). Spatial structure, cooperation and competition in biofilms. Nature Reviews Microbiology *14*, 589-600.
151. Nanamiya, H., Kasai, K., Nozawa, A., Yun, C.S., Narisawa, T., Murakami, K., Natori, Y., Kawamura, F., and Tozawa, Y. (2008). Identification and functional analysis of novel (p)ppGpp synthetase genes in Bacillus subtilis. Mol Microbiol *67*, 291-304.
152. Nano, F.E., Zhang, N., Cowley, S.C., Klose, K.E., Cheung, K.K., Roberts, M.J., Ludu, J.S., Letendre, G.W., Meierovics, A.I., Stephens, G.*, et al.* (2004). A Francisella tularensis pathogenicity island required for intramacrophage growth. Journal of bacteriology *186*, 6430-6436.
153. Nguyen, V.S., Jobichen, C., Tan, K.W., Tan, Y.W., Chan, S.L., Ramesh, K., Yuan, Y., Hong, Y., Seetharaman, J., Leung, K.Y.*, et al.* (2015). Structure of AcrH-AopB Chaperone-Translocator Complex Reveals a Role for Membrane Hairpins in Type III Secretion System Translocon Assembly. Structure *23*, 2022-2031.
154. Nishino, T., and Murao, S. (1975). Characterization of Pyrophosphoryl Transfer Reaction of ATP: Nucleotide Pyrophosphotransferase. Agricultural and Biological Chemistry *39*, 1007-1014.
155. Noegel, A., Rdest, U., Springer, W., and Goebel, W. (1979). Plasmid cistrons controlling synthesis and excretion of the exotoxin alpha-haemolysin of Escherichia coli. Mol Gen Genet *175*, 343-350.

156. Oki, T., Yoshimoto, A., Ogasawara, T., Sato, S., and Takamatsu, A. (1976). Occurrence of pppApp-synthesizing activity in actinomycetes and isolation of purine nucleotide pyrophosphotransferase. Arch Microbiol *107*, 183-187.
157. Oki, T., Yoshimoto, A., Sato, S., and Takamatsu, A. (1975). Purine nucleotide pyrophosphotransferase from Streptomyces morookaensis, capable of synthesizing pppApp and pppGpp. Biochim Biophys Acta *410*, 262-272.
158. Papenfort, K., Förstner, K.U., Cong, J.P., Sharma, C.M., and Bassler, B.L. (2015). Differential RNA-seq of Vibrio cholerae identifies the VqmR small RNA as a regulator of biofilm formation. Proc Natl Acad Sci U S A *112*, E766-775.
159. Paulsen, I.T., Press, C.M., Ravel, J., Kobayashi, D.Y., Myers, G.S.A., Mavrodi, D.V., DeBoy, R.T., Seshadri, R., Ren, Q., Madupu, R.*, et al.* (2005). Complete genome sequence of the plant commensal Pseudomonas fluorescens Pf-5. Nature Biotechnology *23*, 873-878.
160. Pausch, P., Abdelshahid, M., Steinchen, W., Schäfer, H., Gratani, F.L., Freibert, S.A., Wolz, C., Turgay, K., Wilson, D.N., and Bange, G. (2020). Structural Basis for Regulation of the Opposing (p)ppGpp Synthetase and Hydrolase within the Stringent Response Orchestrator Rel. Cell Rep *32*, 108157.
161. Pei, T.T., Li, H., Liang, X., Wang, Z.H., Liu, G., Wu, L.L., Kim, H., Xie, Z., Yu, M., Lin, S.*, et al.* (2020). Intramolecular chaperone-mediated secretion of an Rhs effector toxin by a type VI secretion system. Nat Commun *11*, 1865.
162. Pell, L.G., Kanelis, V., Donaldson, L.W., Howell, P.L., and Davidson, A.R. (2009). The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system. Proc Natl Acad Sci U S A *106*, 4160-4165.
163. Peterson, S.B., Bertolli, S.K., and Mougous, J.D. (2020). The Central Role of Interbacterial Antagonism in Bacterial Life. Curr Biol *30*, R1203-r1214.
164. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.
165. Pissaridou, P., Allsopp, L.P., Wettstadt, S., Howard, S.A., Mavridou, D.A.I., and Filloux, A. (2018). The Pseudomonas aeruginosa T6SS-VgrG1b spike is topped by a PAAR protein eliciting DNA damage to bacterial competitors. Proceedings of the National Academy of Sciences of the United States of America *115*, 12519-12524.
166. Potrykus, K., Murphy, H., Philippe, N., and Cashel, M. (2011). ppGpp is the major source of growth rate control in E. coli. Environ Microbiol *13*, 563-575.
167. Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol *26*, 1641-1650.
168. Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. PloS one *5*, e9490.

169. Procópio, R.E., Silva, I.R., Martins, M.K., Azevedo, J.L., and Araújo, J.M. (2012). Antibiotics produced by Streptomyces. Braz J Infect Dis *16*, 466-471.
170. Pukatzki, S., Ma, A.T., Revel, A.T., Sturtevant, D., and Mekalanos, J.J. (2007). Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. Proc Natl Acad Sci U S A *104*, 15508-15513.
171. Pukatzki, S., Ma, A.T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W.C., Heidelberg, J.F., and Mekalanos, J.J. (2006). Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system. Proc Natl Acad Sci U S A *103*, 1528-1533.
172. Purcell, M., and Shuman, H.A. (1998). The Legionella pneumophila icmGCDJBF genes are required for killing of human macrophages. Infect Immun *66*, 2245-2255.
173. Pym, A.S., Brodin, P., Brosch, R., Huerre, M., and 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*, 709-717.
174. Quentin, D., Ahmad, S., Shanthamoorthy, P., Mougous, J.D., Whitney, J.C., and Raunser, S. (2018). Mechanism of loading and translocation of type VI secretion system effector Tse6. Nat Microbiol *3*, 1142-1152.
175. Raetz, C.R. (1978). Enzymology, genetics, and regulation of membrane phospholipid synthesis in Escherichia coli. Microbiol Rev *42*, 614-659.
176. Rao, P.S., Yamada, Y., Tan, Y.P., and Leung, K.Y. (2004). Use of proteomics to identify novel virulence determinants that are required for Edwardsiella tarda pathogenesis. Mol Microbiol *53*, 573-586.
177. Rapisarda, C., Cherrak, Y., Kooger, R., Schmidt, V., Pellarin, R., Logger, L., Cascales, E., Pilhofer, M., Durand, E., and Fronzes, R. (2019). In situ and high-resolution cryo-EM structure of a bacterial type VI secretion system membrane complex. EMBO J *38*.
178. Renault, M.G., Zamarreno Beas, J., Douzi, B., Chabalier, M., Zoued, A., Brunet, Y.R., Cambillau, C., Journet, L., and Cascales, E. (2018). The gp27-like Hub of VgrG Serves as Adaptor to Promote Hcp Tube Assembly. J Mol Biol *430*, 3143-3156.
179. Rhaese, H.J., and Groscurth, R. (1976). Control of development: role of regulatory nucleotides synthesized by membranes of Bacillus subtilis in initiation of sporulation. Proc Natl Acad Sci U S A *73*, 331-335.
180. Rhaese, H.J., Hoch, J.A., and Groscurth, R. (1977). Studies on the control of development: isolation of Bacillus subtilis mutants blocked early in sporulation and defective in synthesis of highly phosphorylated nucleotides. Proc Natl Acad Sci U S A *74*, 1125-1129.
181. Rietsch, A., Vallet-Gely, I., Dove, S.L., and Mekalanos, J.J. (2005). ExsE, a secreted regulator of type III secretion genes in Pseudomonas aeruginosa. Proc Natl Acad Sci U S A *102*, 8006-8011.

182. Rivera-Chávez, F., and Mekalanos, J.J. (2019). Cholera toxin promotes pathogen acquisition of host-derived nutrients. Nature *572*, 244-248.
183. Rosales-Reyes, R., Skeldon, A.M., Aubert, D.F., and Valvano, M.A. (2012). The Type VI secretion system of Burkholderia cenocepacia affects multiple Rho family GTPases disrupting the actin cytoskeleton and the assembly of NADPH oxidase complex in macrophages. Cell Microbiol *14*, 255-273.
184. Ross, B.D., Verster, A.J., Radey, M.C., Schmidtke, D.T., Pope, C.E., Hoffman, L.R., Hajjar, A.M., Peterson, S.B., Borenstein, E., and Mougous, J.D. (2019). Human gut bacteria contain acquired interbacterial defence systems. Nature *575*, 224-228.
185. Ruhe, Z.C., Subramanian, P., Song, K., Nguyen, J.Y., Stevens, T.A., Low, D.A., Jensen, G.J., and Hayes, C.S. (2018). Programmed Secretion Arrest and Receptor-Triggered Toxin Export during Antibacterial Contact-Dependent Growth Inhibition. Cell *175*, 921-933 e914.
186. Russell, A.B., Hood, R.D., Bui, N.K., LeRoux, M., Vollmer, W., and Mougous, J.D. (2011). Type VI secretion delivers bacteriolytic effectors to target cells. Nature *475*, 343-347.
187. Russell, A.B., LeRoux, M., Hathazi, K., Agnello, D.M., Ishikawa, T., Wiggins, P.A., Wai, S.N., and Mougous, J.D. (2013). Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. Nature *496*, 508-512.
188. Russell, A.B., Peterson, S.B., and Mougous, J.D. (2014a). Type VI secretion system effectors: poisons with a purpose. Nat Rev Microbiol *12*, 137-148.
189. Russell, A.B., Singh, P., Brittnacher, M., Bui, N.K., Hood, R.D., Carl, M.A., Agnello, D.M., Schwarz, S., Goodlett, D.R., Vollmer, W.*, et al.* (2012). A widespread bacterial type VI secretion effector superfamily identified using a heuristic approach. Cell Host Microbe *11*, 538-549.
190. Russell, A.B., Wexler, A.G., Harding, B.N., Whitney, J.C., Bohn, A.J., Goo, Y.A., Tran, B.Q., Barry, N.A., Zheng, H., Peterson, S.B.*, et al.* (2014b). A type VI secretion-related pathway in Bacteroidetes mediates interbacterial antagonism. Cell Host Microbe *16*, 227-236.
191. Ruwe, M., Rückert, C., Kalinowski, J., and Persicke, M. (2018). Functional Characterization of a Small Alarmone Hydrolase in Corynebacterium glutamicum. Front Microbiol *9*, 916.
192. Samsonov, V.V., Samsonov, V.V., and Sineoky, S.P. (2002). DcrA and dcrB Escherichia coli genes can control DNA injection by phages specific for BtuB and FhuA receptors. Res Microbiol *153*, 639-646.
193. Santoriello, F.J., Michel, L., Unterweger, D., and Pukatzki, S. (2020). Pandemic Vibrio cholerae shuts down site-specific recombination to retain an interbacterial defence mechanism. Nat Commun *11*, 6246.
194. Sao-Jose, C., Lhuillier, S., Lurz, R., Melki, R., Lepault, J., Santos, M.A., and Tavares, P. (2006). The ectodomain of the viral receptor YueB forms a fiber that triggers ejection of bacteriophage SPP1 DNA. J Biol Chem *281*, 11464-11470.

195. Sarubbi, E., Rudd, K.E., Xiao, H., Ikehara, K., Kalman, M., and Cashel, M. (1989). Characterization of the spoT gene of Escherichia coli. J Biol Chem *264*, 15074-15082.
196. Schell, M.A., Ulrich, R.L., Ribot, W.J., Brueggemann, E.E., Hines, H.B., Chen, D., Lipscomb, L., Kim, H.S., Mrazek, J., Nierman, W.C.*, et al.* (2007). Type VI secretion is a major virulence determinant in Burkholderia mallei. Mol Microbiol *64*, 1466-1485.
197. Schreiber, G., Metzger, S., Aizenman, E., Roza, S., Cashel, M., and Glaser, G. (1991). Overexpression of the relA gene in Escherichia coli. J Biol Chem *266*, 3760-3767.
198. Schwarz, S., Hood, R.D., and Mougous, J.D. (2010a). What is type VI secretion doing in all those bugs? Trends Microbiol *18*, 531-537.
199. Schwarz, S., West, T.E., Boyer, F., Chiang, W.C., Carl, M.A., Hood, R.D., Rohmer, L., Tolker-Nielsen, T., Skerrett, S.J., and Mougous, J.D. (2010b). Burkholderia type VI secretion systems have distinct roles in eukaryotic and bacterial cell interactions. PLoS Pathog *6*, e1001068.
200. Seubert, A., Hiestand, R., de la Cruz, F., and Dehio, C. (2003). A bacterial conjugation machinery recruited for pathogenesis. Mol Microbiol *49*, 1253-1266.
201. Shneider, M.M., Buth, S.A., Ho, B.T., Basler, M., Mekalanos, J.J., and Leiman, P.G. (2013). PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. Nature *500*, 350-353.
202. Silverman, J.M., Agnello, D.M., Zheng, H., Andrews, B.T., Li, M., Catalano, C.E., Gonen, T., and Mougous, J.D. (2013). Haemolysin coregulated protein is an exported receptor and chaperone of type VI secretion substrates. Mol Cell *51*, 584-593.
203. Smith, W.P.J., Vettiger, A., Winter, J., Ryser, T., Comstock, L.E., Basler, M., and Foster, K.R. (2020). The evolution of the type VI secretion system as a disintegration weapon. PLoS Biol *18*, e3000720.
204. Souza, D.P., Oka, G.U., Alvarez-Martinez, C.E., Bisson-Filho, A.W., Dunger, G., Hobeika, L., Cavalcante, N.S., Alegria, M.C., Barbosa, L.R., Salinas, R.K.*, et al.* (2015). Bacterial killing via a type IV secretion system. Nat Commun *6*, 6453.
205. Spínola-Amilibia, M., Davó-Siguero, I., Ruiz, F.M., Santillana, E., Medrano, F.J., and Romero, A. (2016). The structure of VgrG1 from Pseudomonas aeruginosa, the needle tip of the bacterial type VI secretion system. Acta Crystallogr D Struct Biol *72*, 22-33.
206. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312-1313.
207. Steinchen, W., and Bange, G. (2016). The magic dance of the alarmones (p)ppGpp. Mol Microbiol *101*, 531-544.
208. Steinchen, W., Schuhmacher, J.S., Altegoer, F., Fage, C.D., Srinivasan, V., Linne, U., Marahiel, M.A., and Bange, G. (2015). Catalytic mechanism and

allosteric regulation of an oligomeric (p)ppGpp synthetase by an alarmone. Proc Natl Acad Sci U S A *112*, 13348-13353.
209. Stevens, A.J., Brown, Z.Z., Shah, N.H., Sekar, G., Cowburn, D., and Muir, T.W. (2016). Design of a Split Intein with Exceptional Protein Splicing Activity. J Am Chem Soc *138*, 2162-2165.
210. Stiege, A.C., Isidro, A., Dröge, A., and Tavares, P. (2003). Specific targeting of a DNA-binding protein to the SPP1 procapsid by interaction with the portal oligomer. Mol Microbiol *49*, 1201-1212.
211. Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M.*, et al.* (2000). Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. Nature *406*, 959-964.
212. Sun, D., Lee, G., Lee, J.H., Kim, H.Y., Rhee, H.W., Park, S.Y., Kim, K.J., Kim, Y., Kim, B.Y., Hong, J.I.*, et al.* (2010). A metazoan ortholog of SpoT hydrolyzes ppGpp and functions in starvation responses. Nat Struct Mol Biol *17*, 1188-1194.
213. Tamman, H., Van Nerom, K., Takada, H., Vandenberk, N., Scholl, D., Polikanov, Y., Hofkens, J., Talavera, A., Hauryliuk, V., Hendrix, J.*, et al.* (2020). A nucleotide-switch mechanism mediates opposing catalytic activities of Rel enzymes. Nat Chem Biol *16*, 834-840.
214. Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol *157*, 38-46.
215. Tang, J.Y., Bullen, N.P., Ahmad, S., and Whitney, J.C. (2018). Diverse NADase effector families mediate interbacterial antagonism via the type VI secretion system. J Biol Chem *293*, 1504-1514.
216. Tareen, A., and Kinney, J.B. (2020). Logomaker: beautiful sequence logos in Python. Bioinformatics *36*, 2272-2274.
217. Taylor, J.C., Gao, X., Xu, J., Holder, M., Petrosino, J., Kumar, R., Liu, W., Höök, M., Mackenzie, C., Hillhouse, A.*, et al.* (2021). A type VII secretion system of Streptococcus gallolyticus subsp. gallolyticus contributes to gut colonization and the development of colon tumors. PLoS Pathog *17*, e1009182.
218. Ting, S.Y., Bosch, D.E., Mangiameli, S.M., Radey, M.C., Huang, S., Park, Y.J., Kelly, K.A., Filip, S.K., Goo, Y.A., Eng, J.K.*, et al.* (2018). Bifunctional Immunity Proteins Protect Bacteria against FtsZ-Targeting ADP-Ribosylating Toxins. Cell *175*, 1380-1392 e1314.
219. Ting, S.Y., Martinez-Garcia, E., Huang, S., Bertolli, S.K., Kelly, K.A., Cutler, K.J., Su, E.D., Zhi, H., Tang, Q., Radey, M.C.*, et al.* (2020). Targeted Depletion of Bacteria from Mixed Populations by Programmable Adhesion with Antagonistic Competitor Cells. Cell host & microbe.
220. 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. Mol Microbiol *115*, 478-489.

221. Treangen, T.J., Ondov, B.D., Koren, S., and Phillippy, A.M. (2014). The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol *15*, 524.
222. Tsirigotaki, A., De Geyter, J., Šoštaric, N., Economou, A., and Karamanou, S. (2017). Protein export through the bacterial Sec pathway. Nat Rev Microbiol *15*, 21-36.
223. Typas, A., Banzhaf, M., Gross, C.A., and Vollmer, W. (2011). From the regulation of peptidoglycan synthesis to bacterial growth and morphology. Nat Rev Microbiol *10*, 123-136.
224. Unterweger, D., Kostiuk, B., and Pukatzki, S. (2017). Adaptor Proteins of Type VI Secretion System Effectors. Trends Microbiol *25*, 8-10.
225. Uzan, M., and Danchin, A. (1976). A rapid test for the rel A mutation in E. coli. Biochem Biophys Res Commun *69*, 751-758.
226. van der Heul, H.U., Bilyk, B.L., McDowall, K.J., Seipke, R.F., and van Wezel, G.P. (2018). Regulation of antibiotic production in Actinobacteria: new perspectives from the post-genomic era. Nat Prod Rep *35*, 575-604.
227. Vergunst, A.C., Schrammeijer, B., den Dulk-Ras, A., de Vlaam, C.M., Regensburg-Tuink, T.J., and Hooykaas, P.J. (2000). VirB/D4-dependent protein translocation from Agrobacterium into plant cells. Science *290*, 979-982.
228. Vettiger, A., and Basler, M. (2016). Type VI Secretion System Substrates Are Transferred and Reused among Sister Cells. Cell *167*, 99-110 e112.
229. Vinga, I., Dröge, A., Stiege, A.C., Lurz, R., Santos, M.A., Daugelavicius, R., and Tavares, P. (2006). The minor capsid protein gp7 of bacteriophage SPP1 is required for efficient infection of Bacillus subtilis. Mol Microbiol *61*, 1609-1621.
230. Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D.*, et al.* (2019). SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. Commun Biol *2*, 218.
231. Wang, B., Dai, P., Ding, D., Del Rosario, A., Grant, R.A., Pentelute, B.L., and Laub, M.T. (2019). Affinity-based capture and identification of protein effectors of the growth regulator ppGpp. Nat Chem Biol *15*, 141-150.
232. Wang, R.Y., Song, Y., Barad, B.A., Cheng, Y., Fraser, J.S., and DiMaio, F. (2016). Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. Elife *5*.
233. Wexler, A.G., Bao, Y., Whitney, J.C., Bobay, L.M., Xavier, J.B., Schofield, W.B., Barry, N.A., Russell, A.B., Tran, B.Q., Goo, Y.A.*, et al.* (2016). Human symbionts inject and neutralize antibacterial toxins to persist in the gut. Proc Natl Acad Sci U S A *113*, 3639-3644.
234. Whitney, J.C., Beck, C.M., Goo, Y.A., Russell, A.B., Harding, B.N., De Leon, J.A., Cunningham, D.A., Tran, B.Q., Low, D.A., Goodlett, D.R.*, et al.* (2014). Genetically distinct pathways guide effector export through the type VI secretion system. Mol Microbiol *92*, 529-542.

235. Whitney, J.C., Chou, S., Russell, A.B., Biboy, J., Gardiner, T.E., Ferrin, M.A., Brittnacher, M., Vollmer, W., and Mougous, J.D. (2013). Identification, structure, and function of a novel type VI secretion peptidoglycan glycoside hydrolase effector-immunity pair. J Biol Chem *288*, 26616-26624.
236. Whitney, J.C., Peterson, S.B., Kim, J., Pazos, M., Verster, A.J., Radey, M.C., Kulasekara, H.D., Ching, M.Q., Bullen, N.P., Bryant, D.*, et al.* (2017). A broadly distributed toxin family mediates contact-dependent antagonism between gram-positive bacteria. Elife *6*.
237. Whitney, J.C., Quentin, D., Sawai, S., LeRoux, M., Harding, B.N., Ledvina, H.E., Tran, B.Q., Robinson, H., Goo, Y.A., Goodlett, D.R.*, et al.* (2015). An interbacterial NAD(P)(+) glycohydrolase toxin requires elongation factor Tu for delivery to target cells. Cell *163*, 607-619.
238. Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta crystallographica Section D, Biological crystallography *67*, 235-242.
239. Winn, M.D., Isupov, M.N., and Murshudov, G.N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta crystallographica Section D, Biological crystallography *57*, 122-133.
240. Winsor, G.L., Griffiths, E.J., Lo, R., Dhillon, B.K., Shay, J.A., and Brinkman, F.S. (2016). Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. Nucleic Acids Res *44*, D646-653.
241. Wood, T.E., Howard, S.A., Forster, A., Nolan, L.M., Manoli, E., Bullen, N.P., Yau, H.C.L., Hachani, A., Hayward, R.D., Whitney, J.C.*, et al.* (2019a). The Pseudomonas aeruginosa T6SS Delivers a Periplasmic Toxin that Disrupts Bacterial Cell Morphology. Cell Rep *29*, 187-201 e187.
242. Wood, T.E., Howard, S.A., Wettstadt, S., and Filloux, A. (2019b). PAAR proteins act as the 'sorting hat' of the type VI secretion system. Microbiology *165*, 1203-1218.
243. Xu, M., Wang, W., Waglechner, N., Culp, E.J., Guitor, A.K., and Wright, G.D. (2020). GPAHex-A synthetic biology platform for Type IV-V glycopeptide antibiotic production and discovery. Nat Commun *11*, 5232.
244. Yang, X.Y., Li, Z.Q., Gao, Z.Q., Wang, W.J., Geng, Z., Xu, J.H., She, Z., and Dong, Y.H. (2017). Structural and SAXS analysis of Tle5-Tli5 complex reveals a novel inhibition mechanism of H2-T6SS in Pseudomonas aeruginosa. Protein Sci *26*, 2083-2091.
245. Yang, Z., Fang, J., Chittuluru, J., Asturias, F.J., and Penczek, P.A. (2012). Iterative stable alignment and clustering of 2D transmission electron microscope images. Structure *20*, 237-247.
246. Yu, G. (2020). Using ggtree to Visualize Data on Tree-Like Structures. Curr Protoc Bioinformatics *69*, e96.

247. Zhang, D., de Souza, R.F., Anantharaman, V., Iyer, L.M., and 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.
248. Zhang, D., Iyer, L.M., and Aravind, L. (2011). A novel immunity system for bacterial nucleic acid degrading toxins and its recruitment in various eukaryotic and DNA viral systems. Nucleic acids research *39*, 4532-4552.
249. Zhang, Z., Du, C., de Barsy, F., Liem, M., Liakopoulos, A., van Wezel, G.P., Choi, Y.H., Claessen, D., and Rozen, D.E. (2020). Antibiotic production in Streptomyces is organized by a division of labor through terminal genomic differentiation. Sci Adv *6*, eaay5781.
250. Zheng, J., and Leung, K.Y. (2007). Dissection of a type VI secretion system in Edwardsiella tarda. Mol Microbiol *66*, 1192-1206.
251. Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods *14*, 331-332.
252. Zhou, S., Bhukya, H., Malet, N., Harrison, P.J., Rea, D., Belousoff, M.J., Venugopal, H., Sydor, P.K., Styles, K.M., Song, L.*, et al.* (2021). Molecular basis for control of antibiotic production by a bacterial hormone. Nature *590*, 463-467.