Title
Structural and Biochemical Insights into Bacterial Cell Wall Glycopolymer Display

Author
Martinez, Orlando Edward

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Structural and Biochemical Insights into Bacterial Cell Wall Glycopolymer Display

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy in Biochemistry, Molecular, and Structural Biology

by

Orlando Edward Martinez

2022

ABSTRACT OF THE DISSERTATION

Structural and Biochemical Insights into Bacterial Cell Wall Glycopolymer Display

by

Orlando Edward Martinez

Doctor of Philosophy in Biochemistry, Molecular, and Structural Biology
University of California, Los Angeles, 2022
Professor Robert Thompson Clubb, Chair

Pathogenic Gram-positive bacteria have acquired resistance to many commonly used antibiotics. One potential strategy to combat this problem is to develop small molecule therapeutics that target bacterial virulence mechanisms which enable pathogens to cause disease. The bacterial cell surface contains many virulence factors that have key roles in host-pathogen interactions, but the mechanisms by which they are synthesized and displayed remains incompletely understood. The work described in this dissertation is concentrated on understanding how bacteria synthesize wall teichoic acid (WTA) glycopolymers. WTAs are peptidoglycan-anchored alditol-phosphate polymers that have important functions in host immune system evasion, biofilm formation, and antibiotic resistance, among other roles. This research is specifically focused on the TagA glycosyltransferase, which catalyzes the first committed enzymatic step in WTA biosynthesis. Small molecule TagA inhibitors might function as novel antivirulence therapeutics because cells that lack the enzyme in methicillin-resistant *Staphylococcus aureus* (MRSA) maintain viability but become re-sensitized to β-lactam antibiotics.

The research described in this dissertation advances our understanding of the enzymatic mechanism of TagA and its role in controlling the architecture of the bacterial cell wall. Chapter 1 surveys the biosynthetic pathways that are used to produce peptidoglycan and WTA polymers in *Bacillus subtilis* and *S. aureus*. Chapter 2 details the determination of the first crystal structure of the soluble portion of TagA from *Thermoanaerobacter italicus*, which is shown to adopt a novel glycosyltransferase structural fold. This study identified two conserved residues in TagA that are important for catalysis and demonstrated that C-terminal tail residues in the enzyme are essential for enzymatic activity *in vitro* and membrane association in cells.

Chapter 3 describes the construction of a solubility-enhanced TagA variant and crystal structures of this enzyme bound to its native substrate, UDP-ManNAc, and an epimer of the substrate, UDP-GlcNAc. These structures provide insight into stereospecific protein-ligand contacts that confer substrate specificity. Molecular dynamics simulations of full-length TagA models with and without its bound substrates are also presented which demonstrate that UDP-ManNAc stabilizes construction of the enzyme's active site through interactions with key catalytic residues in the C-terminal tail. Collectively, these data suggest a model of enzyme function in which membrane association via the C-terminal tail triggers a conformational change in TagA that is further stabilized by stereospecific contacts with its UDP-ManNAc substrate. Lastly, Chapter 4 details ongoing progress in capturing the active, monomeric form of the protein for structural investigations and studying the influence of TagA activity on cell morphology in *B. subtilis* using transmission electron microscopy. This work contributes to our knowledge of WTA biogenesis in bacteria and lays a foundation for the structure-guided development of TagA-specific inhibitors that could function as antivirulence agents.

The dissertation of Orlando Edward Martinez is approved.

Joseph Ambrose Loo

Jose Alfonso Rodriguez

Jeffery Floyd Miller

Robert Thompson Clubb, Committee Chair

University of California, Los Angeles

2022

I dedicate this dissertation to four beloved people who mean the world to me.

First is my father, who gives me invaluable counsel and life advice that has built me to be the man I am today. I hope to live by your virtuous example.

Second is my mother, who has shown me unconditional love and compassion from the day I was born. I hope to show just as much affection to myself and my family as you have to me.

Third is my sister, who is one of my closest friends and someone I can always rely on for support and reason. I hope to support your greatest ambitions, as you have supported me in mine.

Last, and certainly not least, is my wife, who encourages my scientific passions and every aspiration I've had in life. Emily, you are everything to me. I only hope to be everything to you in turn.

"We are always more afraid than we wish to be,

but we can always be braver than we expect."

—Robert Jordan, *Lord of Chaos*

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1. Introduction: The architecture and biosynthesis of bacterial secondary cell wall glycopolymers

Chapter 2. The structural and mechanistic role of the TagA glycosyltransferase core domain in wall teichoic acid synthesis

Chapter 3. Insight into the molecular basis of substrate recognition by the TagA glycosyltransferase

Chapter 4. Progress towards understanding how monomeric TagA controls WTA display and cell morphology

LIST OF TABLES

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Professor Robert T. Clubb, for his exemplary mentorship, encouragement, and support which contributed greatly to my growth as a scientist. I would also like to thank my collaborators, Professors Joseph Loo, Michael Jung, Eric Brown, Xi Chen, and members of their labs for their advice, resources, direction, and data that led to the success of my projects. I thank the members of my dissertation committee, Professors Joseph Loo, Jeff F. Miller, and Jose Rodriguez for their insightful feedback, consistent advocacy, and career development support throughout my graduate career, including their review of this dissertation. Additionally, I thank Drs. Duilio Cascio and Robert Peterson for sharing their vast knowledge of structural biology and protein structure determination techniques. Special thanks to Dr. Martin Phillips for mentoring me in numerous analytical protein biochemistry skills.

I thank the past and present members of the Clubb lab for their experimental assistance, brainstorming, and general support inside and outside of the lab over the years, including Dr. Brendan Mahoney, Dr. Kat Ellis-Guardiola, Christopher Sue, Andrew Goring, Jess Soule, Allen Takayesu, Jordan Ford, Jason Gosschalk, Dr. Brendan Amer, Dr. Michele Kattke, Dr. Grace Huang, Dr. Ramsay Macdonald, and Dr. Scott McConnell. I would also like to thank the Chemistry-Biology Interface (CBI) Training Grant, the Audree Fowler Fellowship in Protein Science, and the Roberts A. Smith Research Award for their financial support.

Most importantly, I would like to thank the friends and family that supported me throughout my graduate training. Nicholas Bernier, Kristen Rothdeutsch, Brendan Mahoney, Christopher Sue, Jay McDaniel, and Michael Walters are the best friends and support network I could have asked for. My parents, Orlando G. and Sara, and my sister, Carmen, always have my back through thick and thin and never fail to pick me up when I feel down. My cat, Marble, is

the best listener and critic whenever I need an audience for presentation practice. Finally, my amazing wife, Emily, never stopped loving and supporting me from the day we met. She captured my heart and I strive to be the best person I can be because of her. I could not have made it here without any of you.

Chapter Two of this dissertation is a reformatted version of a published manuscript: Kattke MD*, Gosschalk JE*, Martinez OE, Kumar G, Gale RT, Cascio D, Sawaya MR, Phillips ML, Brown ED, Clubb RT. (2019) Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria. PLOS Pathog**.** *15*(4): e1007723. doi:https://doi.org/10.1371/journal.ppat.1007723. It is reproduced with permission under the PLOS Creative Commons Attribution (CC BY) license. Asterisks indicate that these authors contributed to the work equally. We express appreciation to Prof. Brown and members of his laboratory for their collaborative assistance.

Chapter Three of this dissertation is a reformatted version of a published manuscript: Martinez OE, Mahoney BJ, Goring AK, Yi SW, Tran DP, Cascio D, Phillips ML, Muthana MM, Chen X, Jung ME, Loo JA, Clubb RT. (2022) Insight into the molecular basis of substrate recognition by the wall teichoic acid glycosyltransferase TagA. J. Biol. Chem. 298(2): 101464. doi:https://doi.org/10.1016/j.jbc.2021.101464. It is reproduced with permission under the Elsevier terms of the Creative Commons Attribution Non-Commercial No-Derivatives (CC BY NC ND) License. We express appreciation to Profs. Loo, Jung, and Chen and members of their laboratories for their collaborative assistance.

Chapter Four details ongoing progress towards capturing the active, monomeric form of TagA for structure determination and characterizing the effects of TagA activity on *Bacillus subtilis* cell morphology to guide the development of future TagA-specific inhibitors.

VITA

Orlando Edward Martinez received primary and secondary education in Highlands Ranch, Colorado, completing high school with Honors in 2012. Orlando enrolled at the Colorado School of Mines in the fall of 2012. In May 2016, he earned a Bachelor of Science degree in Chemistry with a Biochemistry specialty and graduated with *Cum laude* honors. At the Colorado School of Mines, he received several awards including the Presidential Scholarship and the Department of Chemistry Outstanding Senior Award. In September 2016, Orlando entered UCLA through the Biochemistry, Molecular, and Structural Biology graduate program. During his graduate studies, he received financial support from several fellowships, including the Chemistry-Biology Interface Training Fellowship, Roberts A. Smith Research Award, and the Audree Fowler Fellowship in Protein Sciences. Orlando is an author on two publications and presented multiple posters at conferences across the country.

Publications resulting from this Dissertation Research

Kattke MD, Gosschalk JE, **Martinez OE**, Kumar G, Gale RT, Cascio D, Sawaya MR, Phillips ML, Brown ED, Clubb RT. (2019) Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria. **PLOS Pathog.** *15*(4): e1007723.

Martinez OE, Mahoney BJ, Goring AK, Yi SW, Tran DP, Cascio D, Phillips ML, Muthana MM, Chen X, Jung ME, Loo JA, Clubb RT. (2022) Insight into the molecular basis of substrate recognition by the wall teichoic acid glycosyltransferase TagA. **J. Biol. Chem**. 298(2): 101464.

Chapter 1

Introduction: The architecture and biosynthesis of bacterial cell wall glycopolymers

1.1 Overview

Antibiotic resistance in pathogenic Gram-positive bacteria is a growing global health crisis (1-4). Over 2.8 million infections and 35,000 deaths are caused by antibiotic resistant bacteria in the United States annually, with novel resistance mechanisms evolving soon after the clinical introduction of new antibiotics (4-7). Most antibiotic development to date has focused on the inhibition of cellular processes that are essential for bacterial survival. Another potentially attractive approach to battle drug-resistant bacteria instead focuses on targeting virulence factors—pathways whose function is necessary for bacterial pathogenesis but not cell viability (8-11). The use of these "antivirulence" compounds may be advantageous as compared to conventional antibiotics because they presumably exert reduced selective pressures that drive the evolution of drug resistance mechanisms in bacteria (12-14). Thus, a deeper molecular understanding of virulence mechanisms could lead to the discovery of new therapeutics.

The bacterial surface plays a key role in host-pathogen interactions and can vary greatly between species. Early approaches to characterize bacteria employed a staining procedure developed by Christian Gram in 1884, which categorized all bacteria into two groups based on distinct cell surface structures that either retained a crystal violet Gram stain (Gram-positive) or did not (Gram-negative) (15). Gram-positive bacteria were subsequently further categorized based on the content of guanine and cytosine (G+C) bases in their genome. High G+C bacteria (51-70%) generally originate from the *Actinobacteria* phylum and include species from the *Corynebacterium*, *Mycobacterium*, and *Streptomyces* genera (16). In contrast, low G+C content (<50%) bacterial species are primarily found in the *Firmicutes* phylum and include several pathogens of clinical interest from the *Clostridium*, *Streptococcus*, *Staphylococcus*, and *Bacillus* genera (17). Unfortunately, these basic operational definitions do not account for the wide diversity

of cell wall architectures that encapsulate Gram-positive bacteria. For example, despite sharing a common monoderm structure that is surrounded by a peptidoglycan (PG) sacculus, different species within the *Firmicutes* phylum decorate their surfaces with an array of chemically distinct macromolecules, including teichoic acids, surface layer (S-layer) proteins, polysaccharide capsules, mycolic acids, or combinations of these elements (18,19). Many of these macromolecules have important functions that contribute to virulence in pathogenic bacteria, such as mediating host immune system evasion, cell adhesion, biofilm formation, and antibiotic resistance (18-25). Thus, a comprehensive knowledge of the structure, function, and regulation of these virulence factors may lead to the discovery of novel antivirulence compounds.

In this chapter, I review what is known about the structures and assembly mechanisms of PG and wall teichoic acid surface glycopolymers in *Staphylococcus aureus* and *Bacillus subtilis*. These two model bacteria from the *Firmicutes* phylum have the most well-studied cell wall glycopolymers and are the focus of my thesis research. *S. aureus* is also a high priority member of the ESKAPE pathogens that pose a severe clinical threat due to their high virulence and resistance to antimicrobial treatments (26). Thus, the enzymes that *S. aureus* uses to synthesize and display its glycopolymers are potential drug targets. The biosynthetic pathways of lipoteichoic acids, polysaccharide capsules, S-layers, and mycolic acids are outside the scope of the work described in this thesis and are reviewed elsewhere (27-30).

1.2 The cell wall structure in *S. aureus* and *B. subtilis*

S. aureus and *B. subtilis* cell walls are composed of a thick and densely functionalized PG layer. The structure of PG is generally well-conserved and composed of polymerized disaccharide repeat units that are crosslinked by covalent bridges between stem peptides of adjacent chains (31,32). PG serves as a protective barrier against toxic small molecules and prevents cell lysis from internal turgor pressure (33,34). Additionally, it serves as a scaffold for anchoring extracellular proteins and glycopolymers that can vary between species.

Secondary cell wall polysaccharides (SCWPs) are also major components of the *S. aureus* and *B. subtilis* cell surface and can constitute up to 60% of their cell wall mass (23,30,31,35). Bacterial SCWPs are chemically diverse and are important for nutrient acquisition, host immune system evasion, cell wall architecture, and extracellular protein display (23,30,35). Broadly, most Gram-positive bacteria decorate their cell wall with two types of polymers, a lipid-linked SCWP and a PG-linked SCWP. *B. subtilis* is an exception, as it has four different types of SCWPs that are displayed depending on phosphate availability and extracellular stressors (18). The biosynthetic pathways that produce these polymers in pathogens are exciting targets for the development of antivirulence drugs due to their essential role in bacterial pathogenesis (20,36).

The SCWPs that are associated with the bacterial cell surface are classified by the composition of their main chain polymers. Classical SCWPs, also known as teichoic acids, contain main chains that are formed by alditol-based sugars and linked together by phosphodiester-bridges. The polymers are anchored to the cell wall by an oligosaccharide linkage unit that can vary based on the type of teichoic acid that is displayed (18,23). There are two types of teichoic acids which differ in how they are anchored to the cell. Wall teichoic acids (WTAs) are covalently attached by a phosphodiester bond to PG whereas lipoteichoic acids (LTAs) are attached to a glycolipid

embedded in the cell membrane (30,37). Non-classical SCWPs are chemically diverse and encompass all other lipid- or PG-anchored glycopolymers that do not contain phosphodiester linkages between their repeat units (38,39). Both *S. aureus* and *B. subtilis* produce WTAs, but their main chains are formed from chemically distinct repeat units (23).

WTA polymers are formed by either poly-glycerol-phosphate (Gro-P) or poly-ribitol-phosphate (Rbo-P) subunits that are joined via phosphodiester linkages. The main chain is then connected to the cell wall by a conserved disaccharide linkage unit (Fig. 1.1A) (23). The proteins responsible for producing Gro-P WTAs are called Tag (Teichoic acid glycerol) enzymes, while Rbo-P WTA-producing proteins are called Tar (Teichoic acid ribitol) enzymes. All known WTA polymers consist of a conserved linkage unit composed of Gro-P-N-acetylmannosamine (ManNAc)-β1,4-N-acetylglucosamine(GlcNAc)-PP-undecaprenyl(Und), a product of three early-stage enzymes called TagO, TagA, and TagB (Fig. 1.1) (23). As the functions of TarOAB and TagOAB are very similar, the term Tag is used to denote these enzymes hereafter. Following the production of the linkage unit, the main chain is then polymerized by enzymes from the TagF-like family of proteins. The polymer's main chain is further glycosylated by the TarM and TarS enzymes in *S. aureus* or the TagE enzyme in *B. subtilis*. Then the mature WTA is exported across the cell membrane by the ABC transporter, TarGH or TagGH. Finally, LCP anchoring enzymes covalently link WTAs to the C6 hydroxyl of MurNAc in PG (Fig. 1.1).

Membrane-associated LTA polymers exhibit much more chemical diversity than PG-attached WTA polymers. LTA polymers differ in both the structures of their polymerized subunits and the oligosaccharide linkages that connect the glycopolymer to the membrane-embedded lipid anchor. At present, five classes of LTAs have been characterized (Type I-V LTAs) (Fig. 1.2) (30). Type I LTAs are found in a wide range of bacteria in the *Firmicutes* phylum (*S. aureus, B. subtilis*,

and *Listeria monocytogenes*, among others) and contain a main chain of unbranched poly-Gro-P groups (40). Types II-V LTAs are more complex and less well-characterized, but incorporate cyclic hexose sugars in their main chain polymers (Fig. 1.2) (30,41). They are not discussed further here, but several reviews describe what is currently known about their structures and synthetic pathways (30,40-42).

1.3 Peptidoglycan structure and biosynthesis

The most-conserved feature of *S. aureus* and *B. subtilis* cell walls is the presence of PG, a mesh of polymeric β1,4-linked GlcNAc and N-acetylmuramic acid (MurNAc) disaccharide chains that are crosslinked by pentapeptide bridges. Its biosynthesis is completed by the action of approximately 20 internal and external enzymes and begins with the cytoplasmic biosynthesis of the disaccharide-pentapeptide repeat unit (Fig. 1.3) (32). The enzymes MurA and MurB are responsible for the conversion of the abundant nucleotide diphosphate sugar UDP-GlcNAc into UDP-MurNAc by appending and reducing a pyruvate group on the C3 hydroxyl of GlcNAc, respectively (Fig. 1.3) (32,43,44). Then the sequential action of four amino acid ligases, MurC-F, generate the pentapeptide stem (L-Ala-D-Glu-L-Lys-D-Ala-D-Ala in *S. aureus* and *B. subtilis*) on UDP-MurNAc (Fig. 1.3) (32,43).

The next steps of PG biosynthesis occur on the inner leaflet of the cell membrane and utilize Und-P, a polyprenyl lipid that is used as a molecular scaffold in many other biosynthetic pathways such as teichoic acid, non-classical SCWP, and O-antigen biosynthesis (32,45). The integral membrane protein MraY transfers the mature phospho-MurNAc-pentapeptide from its nucleotide phosphate carrier to Und-P, creating an intermediate known as lipid I (Und-PP-MurNAc-pentapeptide) (Fig. 1.3) (46,47). The first structure of MraY from *Aquifex aeolicus*

revealed ten transmembrane helices with three invariant Asp and one invariant His residue in a cytoplasmic loop that are essential for enzyme function adjacent to a putative hydrophobic pocket that presumably binds the Und-P substrate (47). From here, lipid I is passed to the glycosyltransferase (GT) MurG which catalyzes the transfer of GlcNAc from a UDP-GlcNAc sugar donor to lipid I and completes lipid II (Fig. 1.3) (48). MurG adopts a canonical GT-B structural fold (discussed in detail later) where two Rossmann-like domains form a wide cleft and separately bind to the donor (UDP-GlcNAc) and acceptor (Und-PP-MurNAc-pentapeptide) substrates (49). The *S. aureus* PG biosynthesis pathway contains an additional series of enzymatic steps following the production of lipid II that *B. subtilis* lacks. Three additional enzymes, FemX, FemA, and FemB, append a pentaglycine peptide on the ε-amino group of L-Lys in the stem peptide (50). This branched stem peptide serves as both a site for PG crosslinking that differs from *B. subtilis* and as a site for the assembly of pili, long multi-protein fibrous structures that are critical for the colonization of a host during infection (51). The final cytoplasmic step of PG biosynthesis involves the exporting of lipid II across the cell membrane by the flippases MurJ, RodA, or FtsW—the quantity and identity of lipid II flippases in bacteria remain contested after more than a decade of study (52-54).

Once lipid II is extracellular, peptidoglycan is assembled in two stages by monotopic membrane proteins called penicillin-binding proteins (PBPs): first, the polymerization of lipid II disaccharides is catalyzed by a GT domain; second, the crosslinking of stem peptides on adjacent PG chains is catalyzed by a transpepidase (TP) domain (Fig. 1.3). PBPs are often bifunctional, combining N-terminal GT domains with C-terminal TP domains that function in a concerted effort to maturate PG (Fig. 1.3) (32). The GT domains are unique and contain an α-helical 'head' catalytic subdomain and a hydrophobic 'jaw' subdomain that localizes the enzyme to the cell membrane

and binds to lipid II (55,56). The original crystal structures of PBP2 from *S. aureus* and PBP1a from *A. aeolicus* were perplexing because both structures lacked the canonical nucleotide-binding Rossmann-like fold that most GTs adopt (55,57). Instead, the functional head subdomain has seven α-helical elements and a small β-strand that form both a catalytic pocket at the interface with the jaw subdomain and a channel that leads towards the TP subdomain where the acceptor PG chain binds (55). As with other inverting GTs, PG polymerization occurs via an S_N2-like displacement reaction that is initiated when a conserved Glu in the head subdomain deprotonates the C4 hydroxyl of the GlcNAc acceptor on the growing PG chain. The nucleophile then attacks the anomeric C1 of MurNAc from a lipid II donor and forms a β1,4 glycosidic bond while releasing Und-P for recycling (31). After polymerization, the TP domain of the PBP crosslinks the nascent PG to an adjacent chain through a two-step enzymatic process. First, it forms an acyl-enzyme intermediate through serine-mediated acetylation of the fourth residue of the stem peptide (generally D-Ala) or the terminal Gly in *S. aureus*. Then, the third residue from the stem peptide on an adjacent PG chain deacylates the intermediate via a nucleophilic attack by a side chain nucleophile or terminal amino group, thereby forming covalently crosslinked PG strands (31).

Enzymes that build PG are common antibiotic targets. One of the first discovered antibiotics, penicillin, is the founding member of a large class of β-lactam drugs that mimic the donor peptide substrate and inhibit the TP activity of PBPs. Since the discovery of penicillin, several classes of PBP inhibitors have been developed that either impede transpeptidase activity or glycosyltransferase activity of the bifunctional enzyme (58,59). Additionally, nearly every enzyme involved in PG biosynthesis, polymerization, and crosslinking has a known inhibitor with marked antibacterial properties (32). However, bacteria have evolved mechanisms to acquire resistance to these and other antibiotics. Three primary mechanisms are used to impart this

resistance: the degradation of antibiotics, the alteration of the antibiotic target site, and the export of the antibiotics (60). The acquisition of these resistance mechanisms in clinical pathogens occurs at an alarmingly fast rate (4,6). However, some modern efforts aim to develop inhibitors against alternative enzyme targets that may slow or stop the evolution of new antibacterial resistance mechanisms.

1.4 Teichoic acid diversity and their roles in *S. aureus* and *B. subtilis*

S. aureus and *B. subtilis* elaborate their cell surfaces with LTAs and WTAs, which have distinct cellular functions despite their similar chemical structures. For example, LTAs are produced at the division septum and their genetic deletion leads to defects in cell separation (61). On the other hand, WTA production is localized to where nascent PG insertion occurs in elongating cells (62). WTA-devoid cells exhibit significant morphological changes (rod-shaped bacteria become spherical) and clump together, suggesting that these polymers are required for proper localization of cell elongation machinery and the spatiotemporal regulation of PG maturation (62-64).

The major classes of teichoic acids in *S. aureus* and *B. subtilis* have complex roles in maintaining cell viability. Separate studies showed that LTA display was essential for viability in *S. aureus* but dispensable in *B. subtilis* (61,65). In contrast, the elimination of WTA biosynthesis by the removal of early-stage enzymes only reduces the growth of *S. aureus* and *B. subtilis* but is not lethal (63,66). However, the deletion of enzymes after TagA in the pathway are lethal unless either the *tagO* or *tagA* genes are also eliminated (66,67). This conditional essentiality is assumed to be due to the buildup of late-stage intermediates in the pathway that deplete Und-P, a key molecular scaffold that is necessary for other biochemical pathways, such as PG biosynthesis (23).

Below I discuss what is known about the WTA biosynthetic pathway, which has promising targets for novel antivirulence therapeutic development and is the focus of this dissertation.

1.4.1 Wall teichoic acid structure and biosynthesis

The development of therapeutics that target virulence factors in bacteria have gained significant interest because they reduce the pathogenicity of infectious microbes without killing them and reduce the selective pressure to evolve resistance mechanisms (12-14). WTA production is a particularly promising pathway for the discovery of novel antivirulence drugs because the depletion of WTAs in cells renders *S. aureus* avirulent and re-sensitizes methicillin-resistant *S. aureus* (MRSA) to β-lactam antibiotics (22,63). Additionally, the conditional essentiality of WTA synthetic enzymes is well-characterized and uniform across model Gram-positive bacteria, unlike the LTA biosynthesis pathway. Thus, a deep understanding of the structure and mechanism of WTA biosynthetic steps may enable a new era of treatments for Gram-positive bacterial infections.

The WTA synthesis pathway is most studied in the model bacteria *S. aureus, B. subtilis* 168, and *B. subtilis* W23. Several enzymes work in sequential action on the cytoplasmic face of the membrane to produce anionic WTAs before their extracellular export and anchoring to PG (Fig. 1.1). Initially, the integral membrane glycosyltransferase TagO begins building the WTA linkage unit by transferring a GlcNAc-1-P from a UDP-GlcNAc sugar carrier to Und-P, thereby creating lipid-α (GlcNAc-PP-Und) (68). TagO shares significant sequence and functional similarity to MraY in PG biosynthesis and belongs to the same polyprenyl-phosphate N-acetyl hexosamine 1-phosphate transferase (PNPT) superfamily of enzymes, whose only structures to date have been determined for MraY from *A. aeolicus* (Fig. 1.4A) (47). Ten transmembrane helices compose MraY, which dimerizes in the cell membrane to complete a hydrophobic Und-P binding

surface that leads to a conserved active site pocket (47). At this point, the lipid-α intermediate can be utilized for numerous SCWP synthetic pathways, such as for O-antigens, teichoic acids, and teichuronic acids (68).

The first committed step in WTA production is catalyzed by the N-acetylmannosaminetransferase TagA (Fig. 1.4B) (63). This inverting glycosyltransferase uses a UDP-ManNAc sugar donor—generated by the conversion of UDP-GlcNAc to UDP-ManNAc by the MnaA epimerase—to affix ManNAc to lipid-α by a β1,4 glycosyl linkage to GlcNAc (63,69). TagA functions by a concerted effort between a soluble N-terminal Rossmann-like fold and a dynamic and amphipathic C-terminal appendage (Fig. 1.4B) (70,71). The resulting lipid disaccharide product is called Lipid-β (ManNAc-β1,4-GlcNAc-PP-Und) and resembles the repeat unit of some non-classical SCWPs (38,72). Then, the TagD cytidylyltransferase combines L-α-Gro-3-P with CTP to produce CDP-Gro, which is in turn used as a substrate by TagB to append a Gro-P unit to the terminal ManNAc of Lipid-β (73,74). To validate the enzymes responsible for linkage unit construction, Gro-3-P-ManNAc-β1,4-GlcNAc-PP-Und biosynthesis was reconstituted *in vitro* using purified *B. subtilis* enzymes and authentic semisynthetic substrates (75,76). These studies laid the foundation for the mechanistic determination of linkage unit biosynthetic enzymes. The TagO, TagA, and TagB enzymes are well-conserved in WTA-producing bacterial strains and their study may help advance the development of novel antivirulence compounds.

Following linkage unit synthesis, the WTA synthetic pathways diverge based on the composition of the main chain polymer. The polymerization of the WTA main chain occurs by the action of one or more GTs from the TagF-like family of proteins (Fig. 1.4C). For poly-Gro-P WTAs, such as those in *B. subtilis* 168 and *Staphylococcus epidermidis* cells, a single enzyme

(TagF) utilizes the primed terminal Gro-P from the linkage unit and polymerizes further Gro-P subunits from a CDP-Gro sugar donor (77,78). Poly-Rbo-P-producing organisms primarily differ in the sugars that are used for the main chain polymer. Two methods of Rbo-P polymerization have been characterized so far where (1) the linkage unit is extended by an additional Gro-P before Rbo-P polymerization or (2) Rbo-P is polymerized directly from the conserved linkage unit. Scenario 1, which occurs in *S. aureus*, requires the addition of (most commonly) one Gro-P to the linkage unit by the essential enzyme TarF (79,80). Then, the enzyme TarL acts as both a primase and polymerase, adding over 40 Rbo-P units from CDP-Rbo donors that are synthesized beforehand by a cytidylyltransferase, TarI, and alcohol dehydrogenase, TarJ (81-83). Scenario 2 occurs in *B. subtilis* W23 and eschews the extension of the linkage unit. Instead, the primase TarK adds a single Rbo-P unit to the linkage unit which is then used for polymerization by TarL (79,84). Interestingly, the *B. subtilis* W23 genome contains a TarF gene akin to *S. aureus* and *B. subtilis* 168, however it is dispensable and not transcribed when the cells are cultured in laboratory conditions (79).

The TagF, TarF, and TarL primases and polymerases all fall under a family of peripheral membrane proteins where only one protein has been structurally characterized: TagF from *S. epidermidis* (Fig. 1.4C) (85). The catalytic domain (residues 312-721) forms a canonical GT-B structural fold with two Rossmann-like domains that separately bind the nucleotide (C-terminal) and polyol (N-terminal) and a C-terminal kinked helix that bridges the domains (Fig. 1.4C) (85). This binding pattern is reminiscent of several GT-B superfamily members, especially in the glycogen phosphorylase/GT family of enzymes that includes TagF. Additionally, the apo- and substrate-bound TagF structures contain an N-terminal amphipathic two-helix domain that was shown to be responsible for peripheral membrane association by burying hydrophobic side chains

into one leaflet of the cell membrane (85,86). This structure is particularly unique among GT-B members due to its localization to the cell membrane by a membrane-binding motif that is separated from the Rossmann-like folds and its utilization of a lipid-bound substrate (Fig. 1.4C). TagF is structurally distinct from other known GT-B enzymes, but a search for structurally homologous proteins revealed that the MnaA UDP-GlcNAc 2-epimerase—also used in WTA biosynthesis—is the most structurally similar to TagF. The most appreciable functional difference between TagF and other GT-B enzymes is that TagF acts on an open-ring sugar and hydrolyzes a pyrophosphate bond instead of an anomeric C-O bond typically seen in GT reactions (85).

Once the WTA polymer is built, the polyol chain is glycosylated before it is exported to the cell surface. WTA glycosylation has been shown to modulate the microbe's ability to evade the host's immune system, avoid the uptake of antibiotics, and promote gene transfer (87,88). In *S. aureus*, WTAs are glycosylated by GlcNAc sugars bound to the C4 hydroxyls of the Rbo repeat unit by either α- or β-linkages (87,89). α-GlcNAc glycosylation is catalyzed by TarM, a retaining GT-B enzyme that contains an N-terminal oligomerization domain which mediates homotrimerization (Fig. 1.4D) (88,90). One Rossmann-like fold in the GT-B domain forms an electropositive surface with the oligomerization domain of a neighboring protomer to facilitate WTA polymer acceptor binding (Fig. 1.4D). The opposite Rossmann-like fold of the GT-B domain binds UDP-GlcNAc such that a front-facing S_Ni-like retaining GT reaction mechanism can produce an α-GlcNAc glycosylated Rbo. β-GlcNAc glycosylation occurs separately by the action of an inverting GT-A enzyme, TarS, which contains a C-terminal homotrimerization domain that is distinct from the TarM trimerization domain (Fig. 1.4E) (91). Two dynamic substrate binding loops in the GT-A domain of TarS are essential for catalytic activity and suggest donor and

acceptor positioning to utilize the S_N2-like direct displacement mechanism assigned to many inverting GTs.

B. subtilis strains only contain a single type of sugar covalently attached to WTA polymers (92). In *B. subtilis* 168, the enzyme TagE (a predicted GT-B according to the Carbohydrate Active enZYmes database, CAZY) attaches an α-glucose to the free hydroxyl on the repeating Gro of the WTA main chain (93,94). *B. subtilis* W23 differs slightly, as TarQ (a predicted GT-A-like enzyme according to CAZY and UniProt) appends a β-glucose to the free hydroxyl on Rbo (87). However, there appears to be significant variability in WTA modifications in other microbes. For example, a recent sequence similarity search and teichoic acid glycosylation analysis suggested that *L. monocytogenes* WTAs are glycosylated extracellularly by a GT-C enzyme, YfhO, which has a homolog responsible for LTA glycosylation in *B. subtilis* (95).

After its synthesis and modification, the WTA is transported through the membrane by an ATP-binding cassette (ABC) transporter called TarGH or TagGH. A recent cryo-EM structure showed that TarH contains six transmembrane helices and a well-conserved N-terminal helix (the IF helix) that is primarily responsible for interfacing with TarG, the nucleotide binding domain (Fig. 1.5A) (96). These enzymes form a heterotetramer (a dimer of homodimers) where two L-shaped tunnels form with electropositive residues at the cytoplasmic side of the membrane leaflet and hydrophobic residues that span through the membrane channel (Fig. 1.5A). These channels complement the anionic diphosphates and hydrophobic lipid of Und-PP carrier, suggesting that the ABC transporter evolved to recognize Und-PP instead of specific WTA polymers produced by the cell. Indeed, WTA ABC transporters do not discriminate between poly-Rbo-P and poly-Gro-P main chains, as *B. subtilis* cell viability can be completely recovered when native TagGH is replaced by *S. aureus* TarGH (97). How the polymer threads through TarGH remains unknown,

but a comparison of the TarGH structure to related flippases bound and unbound to ATP implied that TarGH may function like a crankshaft conrod mechanism (96). The entrance to the tunnel in TarH is seemingly shifted between the inner and outer membrane leaflets by the translational movement of the IF helix, thereby pumping the WTA through the membrane.

Once the mature WTA is extracellular, it undergoes covalent attachment to PG and the release of Und-P for recycling. Three enzymes from the LytR-CpsA-Psr (LCP) family of phosphotransferases are responsible for the attachment of WTA to the C6 hydroxyl of MurNAc in polymerized, but not crosslinked, PG—TagT, TagU, and TagV in *B. subtilis* 168 and LcpA, LcpB, and LcpC in *S. aureus* (98-100). Interestingly, the structures of LcpA, TagT, TagU, and TagV are similar to other known LCP enzymes, like Cps2A from *S. penumoniae*, that have a central five to six-stranded β-sheet sandwiched between several helical elements (Fig. 1.5B) (99,101). The central sheets pack such that a hydrophobic cavity is formed to facilitate Und-P lipid binding (represented by Cps2A binding to octaprenyl-pyrophosphate) (Fig. 1.5B) (101). Additional highly conserved WTA and PG binding sites in LCP enzymes with similar electrostatic surfaces suggest a redundancy in WTA-anchoring LCP enzymes, which supports findings where WTAs in *S. aureus* could be anchored to PG by any of its three phosphotransferases (99,102). The seemingly redundant function of LCPs remains incompletely understood, but modern crystal structures continue to unravel the complexities of LCP enzymes and may aid the development of LCP inhibitors to be used as future antibiotics.

1.4.2 Extracellular teichoic acid modifications

Teichoic acids produced by low G+C bacteria (as in the *Firmicutes* phylum) are modified with D-alanine on the C2 hydroxyl of the polymeric repeat unit which imparts a positive charge

and enables zwitterionic chemical properties of the WTA (18). The mechanism of D-alanyl transfer is poorly understood, but in *S. aureus*, pulse-chase experiments using ^{13}C-labeled alanine indicate that the amino acid is first incorporated in LTAs before its subsequent transfer to WTAs (103).

The initial modification of mature LTAs is mediated by four protein products, DltA, DltB, DltC, and DltD (104,105). Intracellularly, DltA utilizes ATP to covalently modify the DltC peptidyl carrier with D-alanine (106). Then, DltB, a membrane-integrated O-acyltransferase homolog, presumably ligates D-alanine from DltC to a lipid carrier and transfers it across the cell membrane (106). Und-P was proposed to serve as the lipid scaffold for export, however this has yet to be experimentally determined (103). Finally, DltD transfers D-alanine to its acyl acceptor, a free hydroxyl on the LTA main chain, either directly or through an acyl-intermediate (106). The enzyme(s) involved in the transfer of D-alanine from LTA to WTA remains unknown (106,107). The D-alanylation of PG-bound WTA polymers marks the final step in its biosynthesis.

Modification of teichoic acid polymers with D-alanine imparts a positive charge and is widely believed to be a bacterial defense mechanism to evade host recognition systems and cationic antibiotics that localize to the anionic cell membrane of bacteria. This is consistent with findings that have shown that the concerted function of the DltA, DltB, DltC, and DltD enzymes confers resistance to the inhibitor growth effects of host antimicrobial peptides and cationic antibiotics (105,108,109). The diversity of tailoring modifications on WTA polymers is not limited to glycosylation and D-alanylation; other microbes, such as *Streptococcus pneumoniae*, decorate their teichoic acids with phosphocholine which anchors choline binding proteins that have important functions in tissue adherence and virulence (110-112).

1.5 Glycosyltransferase enzymes play a key role in cell wall construction

The Gram-positive bacterial cell wall is a complex and multilayered structure that is primarily composed of polysaccharides that serve a key role in maintaining the structural integrity of the cell. In addition, the glycopolymers impart distinct physiochemical properties that mediate the proper construction of the cell surface, which has factors that mediate cell shape determination, cell division, and other fundamental aspects of the cell's physiology. While the structures of cell wall glycopolymers can drastically vary between species, one class of enzymes is essential for the incorporation of sugars into cell wall glycopolymers: glycosyltransferases (GTs). GTs are defined as enzymes which transfer a donor sugar (commonly from a substituted nucleotide-phosphate leaving group) to an acceptor. The acceptor is often another sugar, but can also be a lipid, protein, or other small molecule (113). GT enzymes were originally classified into three distinct structural folds (GT-A, GT-B, and GT-C) derived from sequence similarities in more than 110 GT families (114). This superfamily categorization has since considerably expanded, with two new GT folds (GT-D and GT-E) structurally defined in the last decade (70,115).

GTs produce new glycosidic bonds through a process that involves either the inversion or retention of the donor sugar's anomeric carbon stereochemistry in the product molecule (Fig. 1.6) (113,116). Most GTs facilitate the acceptor molecule's nucleophilic attack on the anomeric carbon of the donor sugar either through an S_N2-like (inverting) or S_Ni-like (retaining) mechanism. For inverting GTs, a simultaneous removal of the leaving phosphate group of the donor occurs with the formation of the glycosidic bond via a short-lived oxocarbenium ion-like transition state (Fig. 1.6A) (113). For retaining GTs, the nucleophile attacks the anomeric carbon of the donor sugar on the same face as the leaving phosphate group and the reaction proceeds through a covalent glycosyl-enzyme intermediate involving a catalytic acidic residue in the active site (Fig. 1.6B)

(113). A large proportion of GTs have been shown to adopt either GT-A or GT-B folds, and both superfamilies have members that function as either inverting or retaining enzymes (113). The remaining GTs that are functionally characterized as GT-C, GT-D, and GT-E enzymes are all inverting GTs (70,115,117).

From a structural perspective, all nucleotide-utilizing GTs contain at least one Rossmann-like fold consisting of alternating β/α/β secondary structure elements that form a central β-sheet. The GT superfamilies are distinguished by several features such as the quantity and topology of Rossmann-like folds, the conservation of catalytic motifs, or the inclusion of additional domains (113,116,117). GT-A enzymes have two closely abutting Rossmann-like folds that form a continuous β-sheet and have clear donor and acceptor binding sites in the two folds (113). Additionally, a vast majority of the GT-A superfamily contains a canonical DXD or DXE amino acid motif (where X can be any amino acid) whose carboxylate groups coordinate a divalent metal cation that is essential for catalysis because it stabilizes the leaving group's negative charge (113,114). Some GT enzymes can function without metal cofactors and instead utilize cationic residues, helix dipole interactions, and imidazole groups from nearby histidine residues to stabilize the negative charges in inverting GT reactions (118-120).

Similarly to GT-A enzymes, most GT-B enzymes also possess DXD/DXE motifs and two Rossmann-like folds, but the β/α/β domains are not continuous and instead form a catalytic cleft (113). The cleft contains highly conserved residues in both Rossmann-like folds that are responsible for binding the donor and acceptor substrates as well as additional cofactors. The canonical GT-A and GT-B folds are frequently found in enzymes that produce glycopolymers displayed on the cell surface. Some of these enzymes include: MurG, the GT involved in cytoplasmic lipid II production for PG biosynthesis; TagF and the TagF-family of WTA

polymerases; and glycosylation enzymes (TarM and TarS in *S. aureus*, TagE in *B. subtilis* 168, and TarQ in *B. subtilis* W23) in WTA biosynthesis. Additionally, several nonclassical SCWPs incorporate pyruvyl groups on their oligosaccharide repeat units or only on their terminal unit using a pyruvyltransferase called CsaB, whose structurally characterized homologs adopt a GT-B fold (38). These few examples demonstrate that the GT-A and GT-B families are heavily involved in the construction of the Gram-positive bacterial cell wall.

The GT-C family enzymes are not as widespread amongst bacteria, as a majority of the characterized members of this group have been shown to perform protein and lipid glycosylation in eukaryotes (117). Enzymes with the GT-C structural fold are distinct because they have 8 to 13 transmembrane helices appended to a single Rossmann-like fold. To date, they have only been shown to catalyze inverting GT reactions and utilize lipid-linked phosphate sugar donors (117). Few enzymes from this superfamily have been structurally characterized, which is unsurprising considering that this fold constitutes only ~15% of the GTs deduced from genomic data and they are inherently difficult to study because they contain a membrane-integrated domain.

In 2014, the first member of the GT-D superfamily was discovered in domain of unknown function 1792 (DUF1792) that mediates the third step in the glycosylation of the bacterial adhesin Fap1 (115,121). DUF1792 crystallized with the typical extended β-sheet like other GTs, but it had very low sequence and structural homology to other structurally characterized GT enzymes (115). The domain crystallized with its UDP product and a Mn^{2+} cofactor, which revealed the presence of a canonical DXE motif within a Rossmann-like fold and provided a structural explanation for why its activity depends on a Mn^{2+} cofactor (115). The structure of the fifth and most recently characterized GT superfamily, GT-E, was first discovered in the TagA glycosyltransferase that

mediates WTA linkage unit synthesis (70). The work presented in chapters 2 and 3 of this dissertation detail the structural and biochemical characterization of TagA.

1.6 The TagA glycosyltransferase

WTA biosynthetic enzymes are attractive targets for the development of novel antibacterial agents. TagA is a particularly strong candidate because when it is genetically deleted, it eliminates WTA display without compromising cell viability (63,66). The removal of WTAs also impairs the cooperative action of PBPs, which impart β-lactam antibiotic resistance in MRSA, and sensitizes the bacteria to β-lactam antibiotics once again (22,87,122). WTAs are essential for virulence of bacteria in animal models, so it stands to reason that suppressing WTA biosynthesis would create noninfectious strains that are also susceptible to antibiotic treatment (123-125). In addition, TagA does not share significant sequence or structural homology with any known eukaryotic protein, suggesting that few off-target effects will occur when it is targeted.

TagA is an inverting GT that catalyzes the transfer of ManNAc to the C4 hydroxyl of GlcNAc on lipid-α by a β-1,4 glycosidic bond (84,126). The enzyme preferentially utilizes synthetic acceptor ligands with long poly-prenyl chains which suggests that it associates with the cell membrane in order to form favorable hydrophobic interactions with the lipid tail of lipid-α (126). TagA appears to function without the use of a divalent cation cofactor and it lacks a canonical DXD/DXE motif (126). Presumably, basic residues (Arg or Lys) likely form stabilizing salt bridges with the UDP-ManNAc diphosphate group to reduce the activation energy of an oxocarbenium ion-like transition state. Chapters 2 and 3 of this work revealed the identity of these residues as well as the role of a C-terminal appendage in catalysis.

TagA is a founding member of the GT26 family as defined by the CAZY database based on its primary sequence (70). Chapters 2-4 of this dissertation reveal that this enzyme adopts a structurally distinct GT-E fold in which the helical C-terminal appendage containing key catalytic residues is responsible for the association of the protein with the bacterial cell membrane and its substrates (71). The work described in this dissertation establishes a structural and mechanistic framework of how TagA operates in solution and on the cell membrane. This detailed molecular understanding of the enzyme-substrate interactions is the first step in guiding the rational optimization of next-generation antivirulence drugs.

1.7 Figures

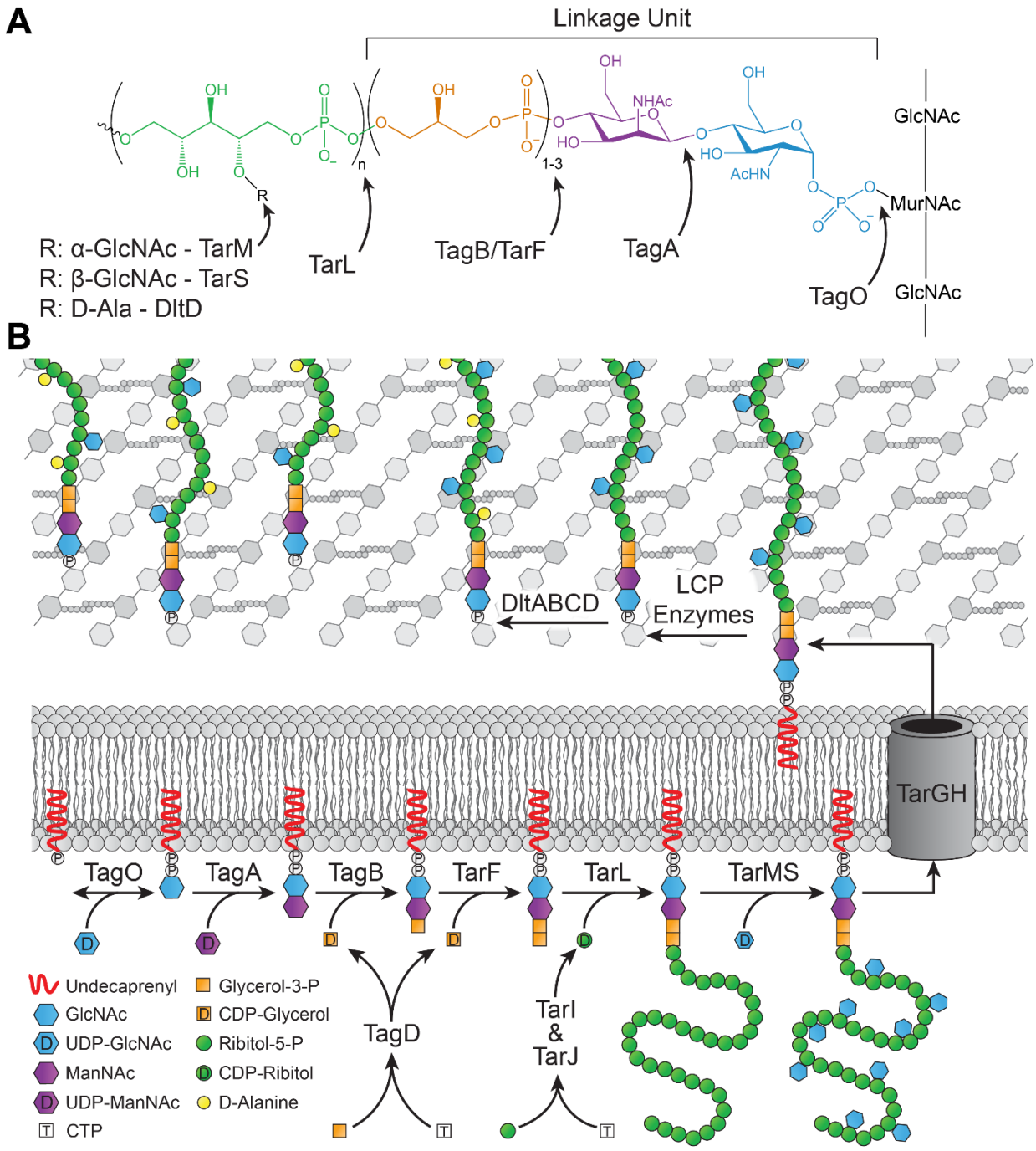

Figure 1.1 The wall teichoic acid biosynthetic pathway. (A) Annotated chemical structure of the *Staphylococcus aureus* wall teichoic acid anchored to the C6 hydroxyl of MurNAc in peptidoglycan. The enzymes used to create the covalent linkages in the polymer are shown. (B)

The biosynthetic enzymes that produce wall teichoic acids localize to the cell membrane. Briefly, the intracellular TagO, TagA, and TagB enzymes synthesize the wall teichoic acid linkage unit, which is composed of GlcNAc (blue), ManNAc (purple), and Gro-P (orange). Then TagF primes the linkage unit for polymerization with an additional Gro-P before TarL polymerizes the glycopolymer with Rbo-P (green). TarM or TarS glycosylate the main chain polymer before the ABC transporter, TarGH, exports the mature wall teichoic acid to the cell wall. Extracellularly, LCP enzymes anchor the polymer to the cell wall and release Und-P for recycling. Finally, four enzymes in the *dlt* operon are implicated in the alanylation of anchored wall teichoic acids, but their mechanisms remain unknown. See main text for component abbreviations.

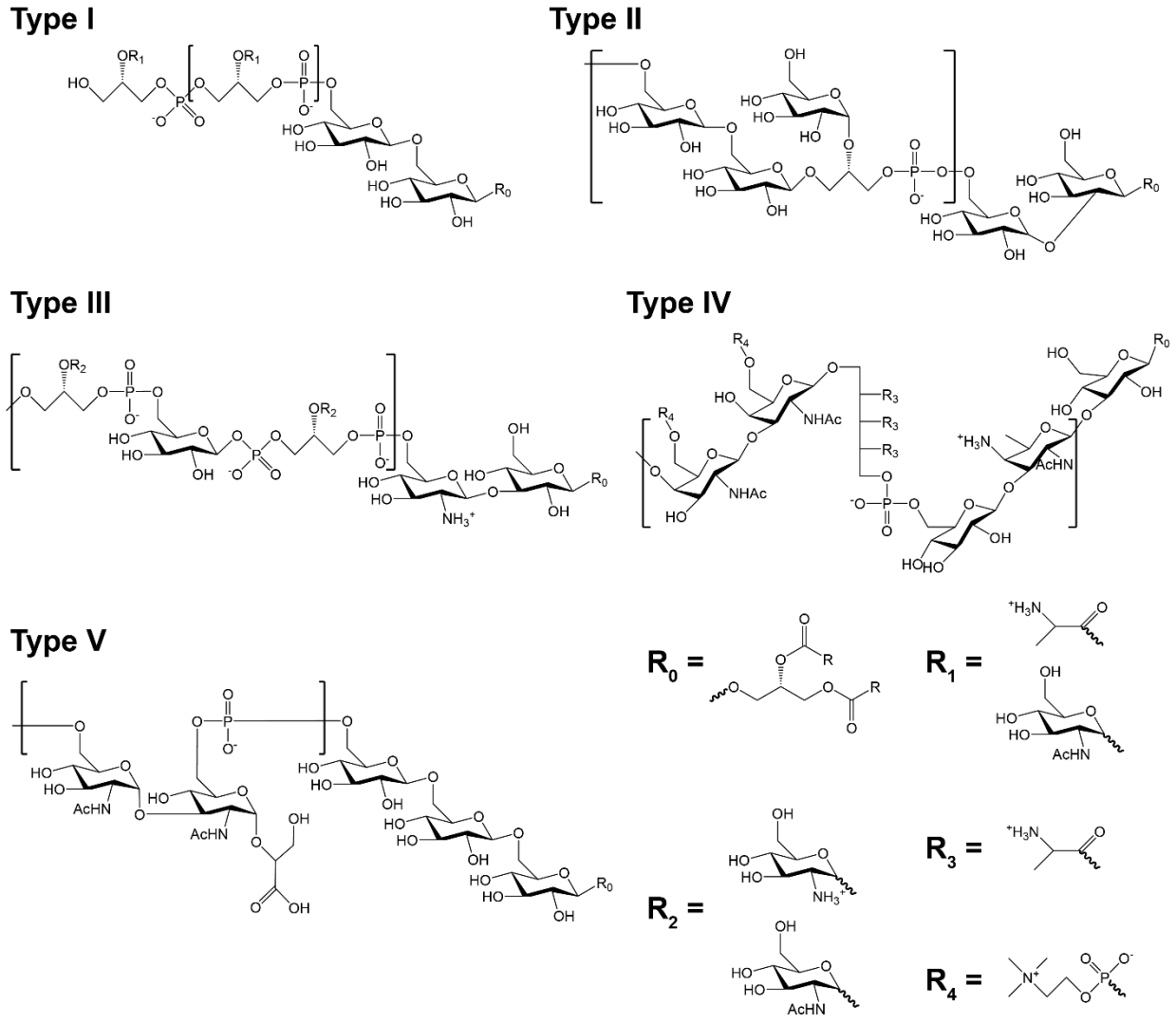

Figure 1.2 Chemical structures of five LTA classes. The chemical structures of Types I-V LTA glycopolymers. Type I was characterized from *Bacillus subtilis*, Type II from *Lactococcus garvieae*, Type III from *Clostridium innocuum*, Type IV from *Streptococcus pneumoniae*, and Type V from *Clostridium difficile*. Repeat units for the polymers are indicated by brackets. R indicates the fatty acid of the glycolipid, R_0 indicates diacylglycerol (DAG), and $R_{1\text{-}4}$ indicates modifications in the LTA polymer. Figure adapted from (30).

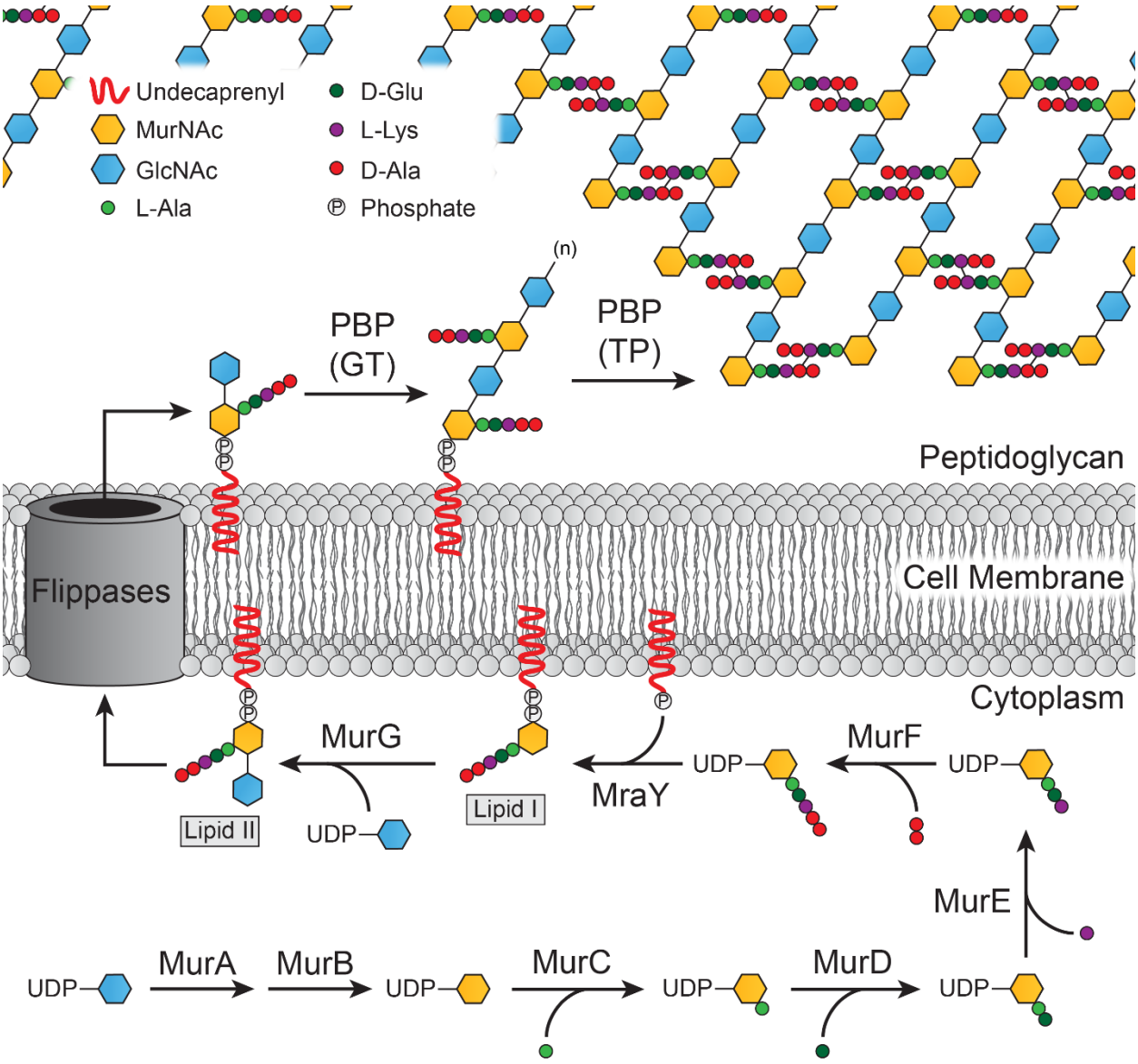

Figure 1.3 The peptidoglycan biosynthesis pathway. The enzymes necessary for the synthesis, polymerization, and crosslinking of peptidoglycan are shown. MurA and MurB catalytically convert UDP-GlcNAc to UDP-MurNAc. Then, the enzymes MurC, MurD, MurE, and MurF append the pentapeptide stem to the MurNAc. The amino acid composition of the pentapeptide can vary among monoderm bacteria. Next, MraY anchors the MurNAc-pentapeptide to an undecaprenyl scaffold, thereby creating lipid I. The MurG glycosyltransferase adds a GlcNAc sugar to MurNAc and finishes the production of the peptidoglycan repeat unit. Unknown flippases

then transport lipid II across the membrane where bifunctional penicillin binding proteins (PBPs) polymerize (via glycosyltransferase (GT) activity) and crosslink (via transpeptidase (TP) activity) nascent peptidoglycan. This figure is adapted from (32) and (52).

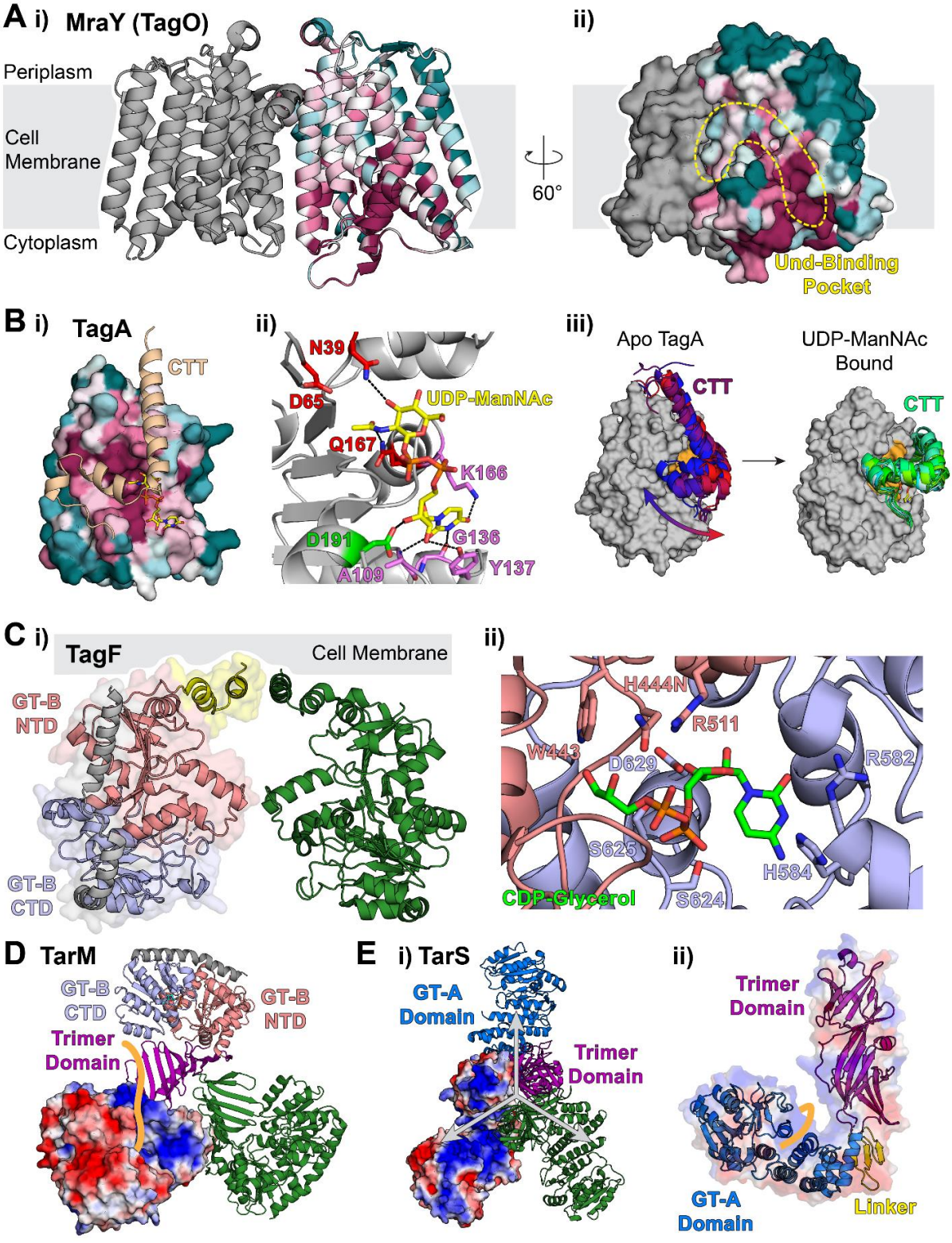

Figure 1.4 Structures of cytoplasmic wall teichoic acid biosynthetic enzymes. (A) (i) The physiological dimeric structure of the peptidoglycan biosynthesis enzyme MraY from *Aquifex*

aeolicus (PDB: 4J72) and its placement in the cell membrane. One protomer of the structure is colored according to its amino acid conservation, with well-conserved residues colored maroon, moderately-conserved residues colored white, and poorly-conserved residues colored teal. MraY is a functionally similar homolog to TagO in wall teichoic acid biosynthesis. (ii) A well-conserved and hydrophobic surface of MraY suggests a likely undecaprenyl-phosphate binding pocket on the cytoplasmic side of the membrane (yellow outline). (B) (i) The structure of the TagA glycosyltransferase from *Thermoanaerobacter italicus* bound to its native sugar donor substrate, UDP-ManNAc (yellow sticks) (PDB: 7N41). The structure lacked density for the flexible and catalytic C-terminal tail (CTT, wheat), but computational models predict that it is positioned near the catalytic pocket. The structure of the core domain is colored by amino acid conservation as in (A). (ii) Several residues in the catalytic pocket of TagA facilitate binding to the uridine (purple and green) or sugar (red) groups of UDP-ManNAc and contribute to substrate stereospecificity. (iii) Molecular dynamics simulations of TagA indicate that the dynamic CTT of TagA is highly flexible in its apo-state (red and blue) and locks into a single conformation when bound to UDP-ManNAc (green and cyan). (C) (i) The structure of the Gro-P polymerase, TagF, from *Staphylococcus epidermidis* and its dimeric position on the cell membrane (PDB: 3L7K). The enzyme adopts a canonical GT-B domain containing two Rossmann-like folds (salmon and light blue) that form a catalytic cleft between them. An additional two-helix appendage (yellow) mediates peripheral membrane association. (ii) A structure of TagF complexed with the CDP-Gro (green sticks) identified substrate-contacting residues in both Rossmann-like folds in the GT-B domain and provide a functional basis for the essential H444. (D) The homotrimeric structure of the wall teichoic acid glycosylation enzyme, TarM, from *Staphylococcus aureus*, which is composed of a GT-B domain (salmon and light blue) and a trimerization domain (purple) (PDB:

4X6L). TarM is responsible for α-GlcNAc modification of the C2 hydroxyl of Rbo-P inside of a cationic pocket in the cleft of the GT-B domain. The electrostatic potential surface is shown for one protomer with negative protein surfaces in red and positive protein surfaces in blue. The poly-Rbo-P wall teichoic acid was modeled to enter the catalytic cleft as seen by the orange arc. (E) (i) The homotrimeric structure of TarS from *S. aureus*, which modifies the C2 hydroxyl of Rbo-P with β-GlcNAc (PDB: 5TZ8). The electrostatic potential surface is shown for one protomer and colored as in (D). TarS contains a GT-A domain (blue) and a trimerization domain (purple) connected by a short linker region (yellow). (ii) The catalytic GT-A domain contains an electropositive surface that may facilitate wall teichoic acid polymer binding (represented by the orange arc).

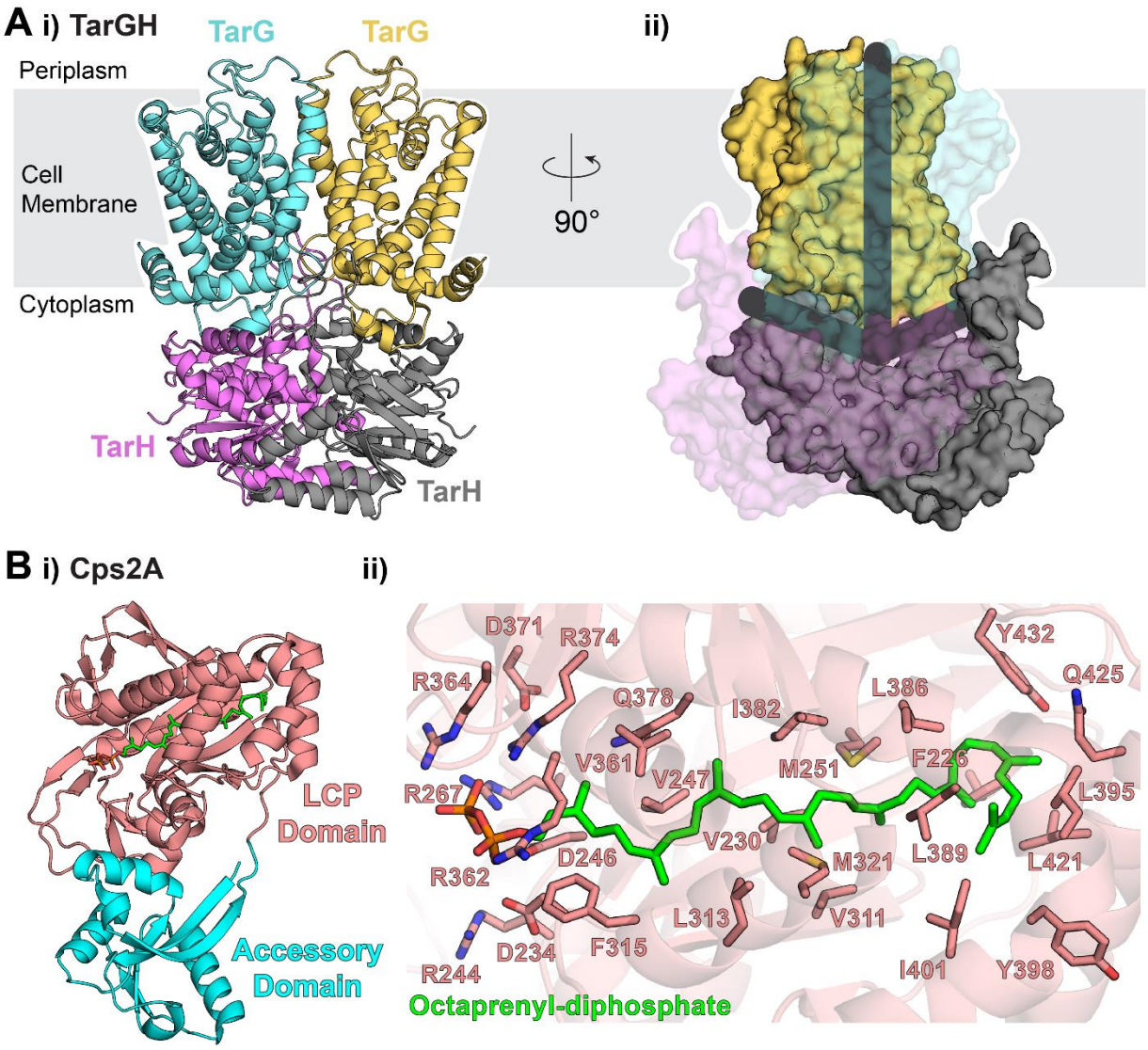

Figure 1.5 Structures of late-stage wall teichoic acid biosynthetic enzymes. (A) (i) The structure of heterotetrameric TarG (blue and yellow) and TarH (gray and purple) from *Aicyclobacillus herbarius* as determined by cryo-electron microscopy (PDB: 6JBH). TarG consists of six transmembrane helices and tightly associates with TarH via two interface helices (IF1 and IF2) and a short coupling helix (CH). TarH possesses a classical nucleotide binding domain that binds to ATP and triggers the conformational change that translocates the lipid-bound substrate across the cell membrane. (ii) The TarGH structure revealed two back-to-back L-shaped tunnels

(represented by the gray bars) that are accessible by three putative gate loops on the cytoplasmic side of the membrane. Cationic residues surround the turn of each tunnel and hydrophobic residues encircle the membrane-embedded side of the tunnel. The complex presumably functions via a crankshaft conrod mechanism that is activated by wall teichoic acid substrate binding and powered by ATP hydrolysis. (B) (i) The structure of the capsule synthesis protein Cps2A from *S. pneumoniae* lacking its transmembrane-spanning domain (Cps2A-ΔTM) and complexed with *cis*-octaprenyl-diphosphate (green sticks) (PDB: 3TEP). Two domains compose Cps2A-ΔTM: an N-terminal accessory domain (teal, residues 111-213) and an LCP (LytR-CpsA-Psr) domain (salmon, residues 214-481). The accessory domain shares little homology to other LCP family proteins. The LCP domain shares ~30% sequence identity to all three *B. subtilis* LCP enzymes and contains a central five-stranded β-sheet sandwiched between helical elements. The LCP domain is responsible for covalently anchoring teichoic acids to peptidoglycan. (ii) The architecture of the LCP domain creates a hydrophobic pocket for lipid substrate binding. Several residues that contain conserved hydrophobicity line a pocket in the LCP domain that leads to a cationic surface. When co-crystallized with a *cis*-octaprenyl-diphosphate substrate analog (green sticks), the lipid oriented such that the cationic residues associate with the diphosphate head and presumably aid in the phosphotransferase activity of the enzyme.

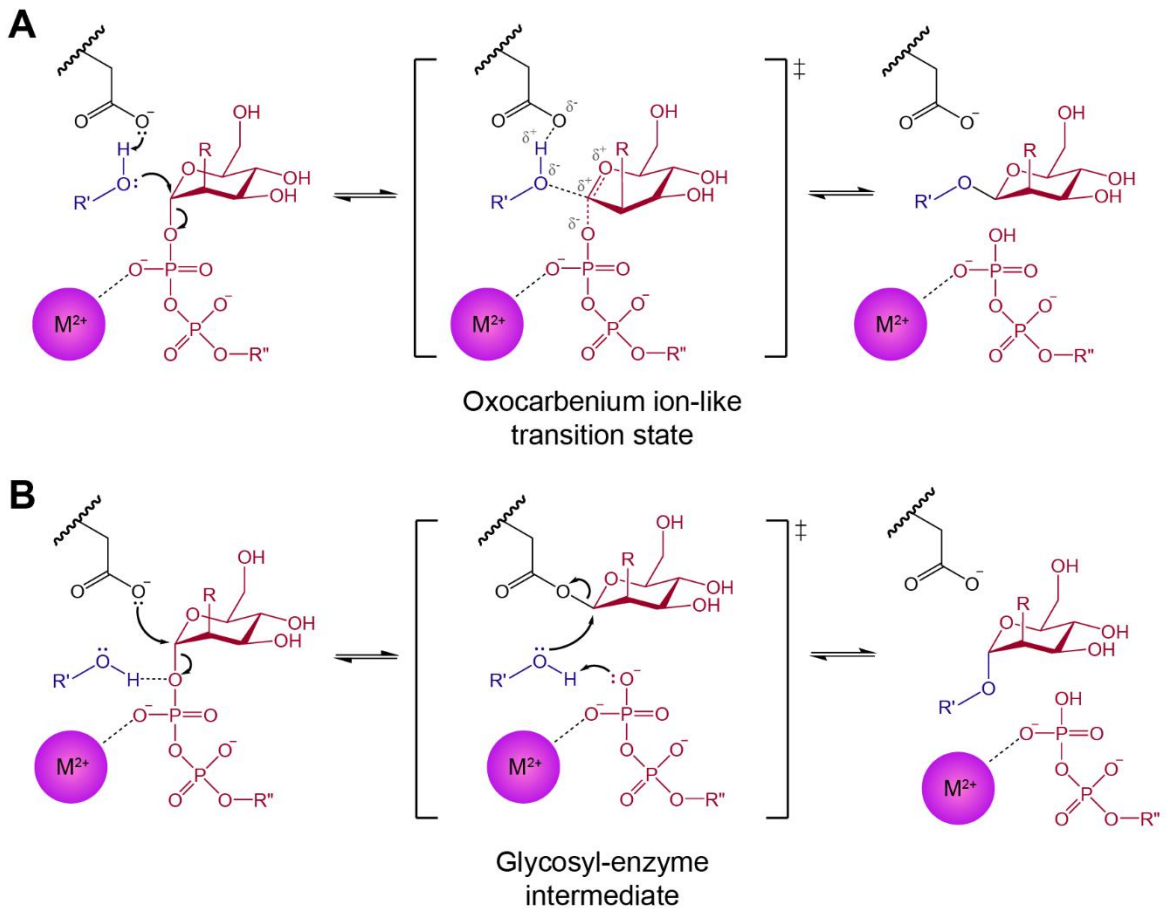

Figure 1.6 Glycosyltransferase enzymatic mechanisms. (A) Inverting glycosyltransferases undergo a direct displacement S_N2-like reaction through an oxocarbenium ion-like transition state that inverts the original stereochemistry of the anomeric carbon of the sugar donor. Divalent metal cofactors or cationic side chains from the protein stabilize the anionic charges of the phosphate leaving group. (B) Retaining glycosyltransferases are proposed to undergo a double displacement mechanism that involves a covalent glycosyl-enzyme intermediate with the enzyme's catalytic glutamate (E) or aspartate (D) residue. These mechanisms also require electrostatic stabilization of the anionic intermediate from a metal cofactor or basic residues. R represents chemical modifications at the C2 of the donor sugar, such as N-acetyl groups in GlcNAc and ManNAc. R'

indicates the acceptor saccharide. R” indicates the nucleotide moiety of the sugar donor. This figure is adapted from (113).

1.8 References

1. Larsson, D. G. J., and Flach, C.-F. (2021) Antibiotic resistance in the environment. *Nature Reviews Microbiology*
2. Andersson, D. I. (2003) Persistence of antibiotic resistant bacteria. *Current Opinion in Microbiology* **6**, 452-456
3. Wang, C.-H., Hsieh, Y.-H., Powers, Z. M., and Kao, C.-Y. (2020) Defeating Antibiotic-Resistant Bacteria: Exploring Alternative Therapies for a Post-Antibiotic Era. *International Journal of Molecular Sciences* **21**, 1061
4. CDC. (2019) Antibiotic Resistance Threats in the United States. U.S. Department of Health and Human Services: Centers for Disease Control and Prevention
5. Peterson, E., and Kaur, P. (2018) Antibiotic Resistance Mechanisms in Bacteria: Relationships Between Resistance Determinants of Antibiotic Producers, Environmental Bacteria, and Clinical Pathogens. *Frontiers in Microbiology* **9**
6. Richardson, L. A. (2017) Understanding and overcoming antibiotic resistance. *PLOS Biology* **15**, e2003775
7. Meredith, H. R., Srimani, J. K., Lee, A. J., Lopatkin, A. J., and You, L. (2015) Collective antibiotic tolerance: mechanisms, dynamics and intervention. *Nature Chemical Biology* **11**, 182-188
8. Allen, R. C., Popat, R., Diggle, S. P., and Brown, S. P. (2014) Targeting virulence: can we make evolution-proof drugs? *Nature Reviews Microbiology* **12**, 300-308
9. El-Halfawy, O. M., Czarny, T. L., Flannagan, R. S., Day, J., Bozelli, J. C., Kuiack, R. C., Salim, A., Eckert, P., Epand, R. M., McGavin, M. J., Organ, M. G., Heinrichs, D. E., and

Brown, E. D. (2020) Discovery of an antivirulence compound that reverses β-lactam resistance in MRSA. *Nature Chemical Biology* **16**, 143-149

10. Rezzoagli, C., Archetti, M., Mignot, I., Baumgartner, M., and Kümmerli, R. (2020) Combining antibiotics with antivirulence compounds can have synergistic effects and reverse selection for antibiotic resistance in Pseudomonas aeruginosa. *PLOS Biology* **18**, e3000805
11. Rice, L. B. (2008) Federal Funding for the Study of Antimicrobial Resistance in Nosocomial Pathogens: No ESKAPE. *The Journal of Infectious Diseases* **197**, 1079-1081
12. Maura, D., Ballok, A. E., and Rahme, L. G. (2016) Considerations and caveats in anti-virulence drug development. *Current Opinion in Microbiology* **33**, 41-46
13. Imperi, F., Fiscarelli, E. V., Visaggio, D., Leoni, L., and Visca, P. (2019) Activity and Impact on Resistance Development of Two Antivirulence Fluoropyrimidine Drugs in Pseudomonas aeruginosa. *Frontiers in Cellular and Infection Microbiology* **9**
14. Rasko, D. A., and Sperandio, V. (2010) Anti-virulence strategies to combat bacteria-mediated disease. *Nature Reviews Drug Discovery* **9**, 117-128
15. Silhavy, T. J., Kahne, D., and Walker, S. (2010) The bacterial cell envelope. *Cold Spring Harb Perspect Biol* **2**, a000414
16. Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G. F., Chater, K. F., and Van Sinderen, D. (2007) Genomics of Actinobacteria : Tracing the Evolutionary History of an Ancient Phylum. *Microbiology and Molecular Biology Reviews* **71**, 495-548
17. Onyenwoke, R., Brill, J., Farahi, K., and Wiegel, J. (2004) Sporulation genes in members of the low G+C Gram-type-positive phylogenetic branch (Firmicutes). *Archives of Microbiology* **182**

18. Weidenmaier, C., and Peschel, A. (2008) Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nat Rev Microbiol* **6**, 276-287

19. Yother, J. (2011) Capsules of Streptococcus pneumoniae and Other Bacteria: Paradigms for Polysaccharide Biosynthesis and Regulation. *Annual Review of Microbiology* **65**, 563-581

20. Holland, L. M., Conlon, B., and O'Gara, J. P. (2011) Mutation of tagO reveals an essential role for wall teichoic acids in Staphylococcus epidermidis biofilm development. *Microbiology* **157**, 408-418

21. Misawa, Y., Kelley, K. A., Wang, X., Wang, L., Park, W. B., Birtel, J., Saslowsky, D., and Lee, J. C. (2015) Staphylococcus aureus Colonization of the Mouse Gastrointestinal Tract Is Modulated by Wall Teichoic Acid, Capsule, and Surface Proteins. *PLOS Pathogens* **11**, e1005061

22. Farha, M. A., Leung, A., Sewell, E. W., D'Elia, M. A., Allison, S. E., Ejim, L., Pereira, P. M., Pinho, M. G., Wright, G. D., and Brown, E. D. (2013) Inhibition of WTA Synthesis Blocks the Cooperative Action of PBPs and Sensitizes MRSA to β-Lactams. *ACS Chemical Biology* **8**, 226-233

23. Brown, S., Santa Maria, J. P., and Walker, S. (2013) Wall Teichoic Acids of Gram-Positive Bacteria. *Annual Review of Microbiology* **67**, 313-336

24. Bush, K. (2018) Past and Present Perspectives on β-Lactamases. *Antimicrobial Agents and Chemotherapy* **62**

25. Poole, K. (2005) Efflux-mediated antimicrobial resistance. *Journal of Antimicrobial Chemotherapy* **56**, 20-51

26. Mulani, M. S., Kamble, E. E., Kumkar, S. N., Tawre, M. S., and Pardesi, K. R. (2019) Emerging Strategies to Combat ESKAPE Pathogens in the Era of Antimicrobial Resistance: A Review. *Frontiers in Microbiology* **10**

27. Fagan, R. P., and Fairweather, N. F. (2014) Biogenesis and functions of bacterial S-layers. *Nature Reviews Microbiology* **12**, 211-222

28. Chateau, A., Van der Verren, S. E., Remaut, H., and Fioravanti, A. (2020) The Bacillus anthracis Cell Envelope: Composition, Physiological Role, and Clinical Relevance. *Microorganisms* **8**, 1864

29. Marrakchi, H., Lanéelle, M.-A., and Daffé, M. (2014) Mycolic Acids: Structures, Biosynthesis, and Beyond. *Chemistry & Biology* **21**, 67-85

30. Percy, M. G., and Gründling, A. (2014) Lipoteichoic Acid Synthesis and Function in Gram-Positive Bacteria. *Annual Review of Microbiology* **68**, 81-100

31. Caveney, N. A., Li, F. K. K., and Strynadka, N. C. J. (2018) Enzyme structures of the bacterial peptidoglycan and wall teichoic acid biogenesis pathways. *Current Opinion in Structural Biology* **53**, 45-58

32. Lovering, A. L., Safadi, S. S., and Strynadka, N. C. J. (2012) Structural Perspective of Peptidoglycan Biosynthesis and Assembly. *Annual Review of Biochemistry* **81**, 451-478

33. Vollmer, W., Blanot, D., and De Pedro, M. A. (2008) Peptidoglycan structure and architecture. *FEMS Microbiology Reviews* **32**, 149-167

34. Pinho, M. G., Kjos, M., and Veening, J.-W. (2013) How to get (a)round: mechanisms controlling growth and division of coccoid bacteria. *Nature Reviews Microbiology* **11**, 601-614

35. Mesnage, S. (2000) Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *The EMBO Journal* **19**, 4473-4484

36. Walter, J., Loach, D. M., Alqumber, M., Rockel, C., Hermann, C., Pfitzenmaier, M., and Tannock, G. W. (2007) d-Alanyl ester depletion of teichoic acids in Lactobacillus reuteri 100-23 results in impaired colonization of the mouse gastrointestinal tract. *Environmental Microbiology* **9**, 1750-1760

37. Kojima, N., Araki, Y., and Ito, E. (1985) Structure of the linkage units between ribitol teichoic acids and peptidoglycan. *Journal of bacteriology* **161**, 299-306

38. Hager, Sütl, Stefanović, Blaukopf, and Schäffer. (2019) Pyruvate Substitutions on Glycoconjugates. *International Journal of Molecular Sciences* **20**, 4929

39. Schäffer, C., and Messner, P. (2005) The structure of secondary cell wall polymers: how Gram-positive bacteria stick their cell walls together. *Microbiology* **151**, 643-651

40. Reichmann, N. T., and Gründling, A. (2011) Location, synthesis and function of glycolipids and polyglycerolphosphate lipoteichoic acid in Gram-positive bacteria of the phylum Firmicutes. *FEMS Microbiology Letters* **319**, 97-105

41. Schneewind, O., and Missiakas, D. (2014) Lipoteichoic Acids, Phosphate-Containing Polymers in the Envelope of Gram-Positive Bacteria. *Journal of Bacteriology* **196**, 1133-1142

42. Denapaite, D., Brückner, R., Hakenbeck, R., and Vollmer, W. (2012) Biosynthesis of Teichoic Acids in Streptococcus pneumoniae and Closely Related Species: Lessons from Genomes. *Microbial Drug Resistance* **18**, 344-358

43. Barreteau, H., Kovač, A., Boniface, A., Sova, M., Gobec, S., and Blanot, D. (2008) Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiology Reviews* **32**, 168-207

44. Yoon, H.-J., Lee, S. J., Mikami, B., Park, H.-J., Yoo, J., and Suh, S. W. (2008) Crystal structure of UDP-N-acetylglucosamine enolpyruvyl transferase fromHaemophilus influenzae in complex with UDP-N-acetylglucosamine and fosfomycin. *Proteins: Structure, Function, and Bioinformatics* **71**, 1032-1037

45. D'Elia, M. A., Millar, K. E., Bhavsar, A. P., Tomljenovic, A. M., Hutter, B., Schaab, C., Moreno-Hagelsieb, G., and Brown, E. D. (2009) Probing teichoic acid genetics with bioactive molecules reveals new interactions among diverse processes in bacterial cell wall biogenesis. *Chem Biol* **16**, 548-556

46. Liu, Y., 刘, 垚., Rodrigues, J. P. G. L. M., Bonvin, A. M. J. J., Zaal, E. A., Berkers, C. R., Heger, M., Gawarecka, K., Swiezewska, E., Breukink, E., and Egmond, M. R. (2016) New Insight into the Catalytic Mechanism of Bacterial MraY from Enzyme Kinetics and Docking Studies. *Journal of Biological Chemistry* **291**, 15057-15068

47. Chung, B. C., Zhao, J., Gillespie, R. A., Kwon, D. Y., Guan, Z., Hong, J., Zhou, P., and Lee, S. Y. (2013) Crystal structure of MraY, an essential membrane enzyme for bacterial cell wall synthesis. *Science* **341**, 1012-1016

48. Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., and Walker, S. (2003) Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. *Proceedings of the National Academy of Sciences* **100**, 845-849

49. Ha, S., Walker, D., Shi, Y., and Walker, S. (2000) The 1.9 Å crystal structure of Escherichia coli MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis. *Protein Science* **9**, 1045-1052

50. Schneider, T., Senn, M. M., Berger-Bächi, B., Tossi, A., Sahl, H.-G., and Wiedemann, I. (2004) In vitro assembly of a complete, pentaglycine interpeptide bridge containing cell wall precursor (lipid II-Gly5) of Staphylococcus aureus. *Molecular Microbiology* **53**, 675-685

51. Hendrickx, A. P. A., Budzik, J. M., Oh, S.-Y., and Schneewind, O. (2011) Architects at the bacterial surface — sortases and the assembly of pili with isopeptide bonds. *Nature Reviews Microbiology* **9**, 166-176

52. Egan, A. J. F., Errington, J., and Vollmer, W. (2020) Regulation of peptidoglycan synthesis and remodelling. *Nature Reviews Microbiology* **18**, 446-460

53. Mohammadi, T., Karczmarek, A., Crouvoisier, M., Bouhss, A., Mengin-Lecreulx, D., and Den Blaauwen, T. (2007) The essential peptidoglycan glycosyltransferase MurG forms a complex with proteins involved in lateral envelope growth as well as with proteins involved in cell division in Escherichia coli. *Molecular Microbiology* **65**, 1106-1121

54. Meeske, A. J., Sham, L.-T., Kimsey, H., Koo, B.-M., Gross, C. A., Bernhardt, T. G., and Rudner, D. Z. (2015) MurJ and a novel lipid II flippase are required for cell wall biogenesis in Bacillus subtilis. *Proceedings of the National Academy of Sciences* **112**, 6437-6442

55. Lovering, A. L., Castro, L. H. d., Lim, D., and Strynadka, N. C. J. (2007) Structural Insight into the Transglycosylation Step of Bacterial Cell-Wall Biosynthesis. *Science* **315**, 1402-1405

56. Sung, M. T., Lai, Y. T., Huang, C. Y., Chou, L. Y., Shih, H. W., Cheng, W. C., Wong, C. H., and Ma, C. (2009) Crystal structure of the membrane-bound bifunctional transglycosylase PBP1b from Escherichia coli. *Proceedings of the National Academy of Sciences* **106**, 8824-8829

57. Yuan, Y., Barrett, D., Zhang, Y., Kahne, D., Sliz, P., and Walker, S. (2007) Crystal structure of a peptidoglycan glycosyltransferase suggests a model for processive glycan chain synthesis. *Proceedings of the National Academy of Sciences* **104**, 5348-5353

58. Tooke, C. L., Hinchliffe, P., Bragginton, E. C., Colenso, C. K., Hirvonen, V. H. A., Takebayashi, Y., and Spencer, J. (2019) β-Lactamases and β-Lactamase Inhibitors in the 21st Century. *Journal of Molecular Biology* **431**, 3472-3500

59. Schwartz, B., Markwalder, J. A., and Wang, Y. (2001) Lipid II: Total Synthesis of the Bacterial Cell Wall Precursor and Utilization as a Substrate for Glycosyltransfer and Transpeptidation by Penicillin Binding Protein (PBP) 1b of Eschericia coli. *Journal of the American Chemical Society* **123**, 11638-11643

60. Munita, J. M., and Arias, C. A. (2016) Mechanisms of Antibiotic Resistance. *Microbiology spectrum* **4**

61. Schirner, K., Marles-Wright, J., Lewis, R. J., and Errington, J. (2009) Distinct and essential morphogenic functions for wall- and lipo-teichoic acids in Bacillus subtilis. *The EMBO journal* **28**, 830-842

62. Formstone, A., Carballido-Lopez, R., Noirot, P., Errington, J., and Scheffers, D. J. (2008) Localization and Interactions of Teichoic Acid Synthetic Enzymes in Bacillus subtilis. *Journal of Bacteriology* **190**, 1812-1821

63. D'Elia, M. A., Henderson, J. A., Beveridge, T. J., Heinrichs, D. E., and Brown, E. D. (2009) The N-acetylmannosamine transferase catalyzes the first committed step of teichoic acid assembly in Bacillus subtilis and Staphylococcus aureus. *J Bacteriol* **191**, 4030-4034

64. Atilano, M. L., Pereira, P. M., Yates, J., Reed, P., Veiga, H., Pinho, M. G., and Filipe, S. R. (2010) Teichoic acids are temporal and spatial regulators of peptidoglycan cross-linking in Staphylococcus aureus. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 18991-18996

65. GrüNdling, A., and Schneewind, O. (2007) Genes Required for Glycolipid Synthesis and Lipoteichoic Acid Anchoring in Staphylococcus aureus. *Journal of Bacteriology* **189**, 2521-2530

66. D'Elia, M. A., Millar, K. E., Beveridge, T. J., and Brown, E. D. (2006) Wall teichoic acid polymers are dispensable for cell viability in Bacillus subtilis. *J Bacteriol* **188**, 8313-8316

67. D'Elia, M. A., Pereira, M. P., Chung, Y. S., Zhao, W., Chau, A., Kenney, T. J., Sulavik, M. C., Black, T. A., and Brown, E. D. (2006) Lesions in Teichoic Acid Biosynthesis in Staphylococcus aureus Lead to a Lethal Gain of Function in the Otherwise Dispensable Pathway. *Journal of Bacteriology* **188**, 4183-4189

68. Soldo, B., Karamata, D., and Lazarevic, V. (2002) tagO is involved in the synthesis of all anionic cell-wall polymers in Bacillus subtilis 168. *Microbiology* **148**, 2079-2087

69. Soldo, B., Lazarevic, V., Pooley, H. M., and Karamata, D. (2002) Characterization of a Bacillus subtilis Thermosensitive Teichoic Acid-Deficient Mutant: Gene mnaA (yvyH) Encodes the UDP-N-Acetylglucosamine 2-Epimerase. *Journal of Bacteriology* **184**, 4316-4320

70. Kattke, M. D., Gosschalk, J. E., Martinez, O. E., Kumar, G., Gale, R. T., Cascio, D., Sawaya, M. R., Philips, M., Brown, E. D., and Clubb, R. T. (2019) Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria. *PLOS Pathogens* **15**, e1007723

71. Martinez, O. E., Mahoney, B. J., Goring, A. K., Yi, S.-W., Tran, D. P., Cascio, D., Phillips, M. L., Muthana, M. M., Chen, X., Jung, M. E., Loo, J. A., and Clubb, R. T. (2021) Insight into the molecular basis of substrate recognition by the wall teichoic acid glycosyltransferase TagA. *Journal of Biological Chemistry*, 101464

72. Swoboda, J. G., Campbell, J., Meredith, T. C., and Walker, S. (2010) Wall teichoic acid function, biosynthesis, and inhibition. *Chembiochem : a European journal of chemical biology* **11**, 35-45

73. Fong, D. H., Yim, V. C. N., D'Elia, M. A., Brown, E. D., and Berghuis, A. M. (2006) Crystal structure of CTP:glycerol-3-phosphate cytidylyltransferase from Staphylococcus aureus: Examination of structural basis for kinetic mechanism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1764**, 63-69

74. Bhavsar, A. P., Truant, R., and Brown, E. D. (2005) The TagB Protein in Bacillus subtilis 168 Is an Intracellular Peripheral Membrane Protein That Can Incorporate Glycerol Phosphate onto a Membrane-bound Acceptor in vitro. *Journal of Biological Chemistry* **280**, 36691-36700

75. Ginsberg, C., Zhang, Y.-H., Yuan, Y., and Walker, S. (2006) In vitro Reconstitution of Two Essential Steps in Wall Teichoic Acid Biosynthesis. *ACS Chemical Biology* **1**, 25-28

76. Gale, R. T., Sewell, E. W., Garrett, T. A., and Brown, E. D. (2014) Reconstituting poly(glycerol phosphate) wall teichoic acid biosynthesis in vitro using authentic substrates. *Chemical Science* **5**, 3823-3823

77. Davison, A. L., and Baddiley, J. (1964) Teichoic Acids in the Walls of Staphylococci : Glycerol Teichoic Acids in Walls of Staphylococcus epidermidis. *Nature* **202**, 874-874

78. Fitzgerald, S. N., and Foster, T. J. (2000) Molecular Analysis of the tagF Gene, Encoding CDP-Glycerol:Poly(glycerophosphate) Glycerophosphotransferase of Staphylococcus epidermidis ATCC 14990. *Journal of Bacteriology* **182**, 1046-1052

79. Brown, S., Meredith, T., Swoboda, J., and Walker, S. (2010) Staphylococcus aureus and Bacillus subtilis W23 make polyribitol wall teichoic acids using different enzymatic pathways. *Chem Biol* **17**, 1101-1110

80. Qian, Z., Yin, Y., Zhang, Y., Lu, L., Li, Y., and Jiang, Y. (2006) Genomic characterization of ribitol teichoic acid synthesis in Staphylococcus aureus: genes, genomic organization and gene duplication. *BMC Genomics* **7**

81. Brown, S., Zhang, Y.-H., and Walker, S. (2008) A Revised Pathway Proposed for Staphylococcus aureus Wall Teichoic Acid Biosynthesis Based on In Vitro Reconstitution of the Intracellular Steps. *Chemistry & Biology* **15**, 12-21

82. Meredith, T. C., Swoboda, J. G., and Walker, S. (2008) Late-Stage Polyribitol Phosphate Wall Teichoic Acid Biosynthesis in Staphylococcus aureus. *Journal of Bacteriology* **190**, 3046-3056

83. Pereira, M. P., and Brown, E. D. (2004) Bifunctional Catalysis by CDP-ribitol Synthase: Convergent Recruitment of Reductase and Cytidylyltransferase Activities in Haemophilus influenzae and Staphylococcus aureus. *Biochemistry* **43**, 11802-11812

84. Yokoyama, K., Miyashita, T., Araki, Y., and Ito, E. (1986) Structure and functions of linkage unit intermediates in the biosynthesis of ribitol teichoic acids in Staphylococcus aureus H and Bacillus subtilis W23. *European journal of biochemistry* **161**, 479-489

85. Lovering, A. L., Lin, L. Y., Sewell, E. W., Spreter, T., Brown, E. D., and Strynadka, N. C. (2010) Structure of the bacterial teichoic acid polymerase TagF provides insights into membrane association and catalysis. *Nat Struct Mol Biol* **17**, 582-589

86. Lomize, A. L., Pogozheva, I. D., Lomize, M. A., and Mosberg, H. I. (2007) The role of hydrophobic interactions in positioning of peripheral proteins in membranes. *BMC Structural Biology* **7**, 44

87. Brown, S., Xia, G., Luhachack, L. G., Campbell, J., Meredith, T. C., Chen, C., Winstel, V., Gekeler, C., Irazoqui, J. E., Peschel, A., and Walker, S. (2012) Methicillin resistance in Staphylococcus aureus requires glycosylated wall teichoic acids. *Proc Natl Acad Sci U S A* **109**, 18909-18914

88. Sobhanifar, S., Worrall, L. J., Gruninger, R. J., Wasney, G. A., Blaukopf, M., Baumann, L., Lameignere, E., Solomonson, M., Brown, E. D., Withers, S. G., and Strynadka, N. C. (2015) Structure and mechanism of Staphylococcus aureus TarM, the wall teichoic acid alpha-glycosyltransferase. *Proc Natl Acad Sci U S A* **112**, E576-585

89. Torii, M., Kabat, E. A., and Bezer, A. E. (1964) SEPARATION OF TEICHOIC ACID OF STAPHYLOCOCCUS AUREUS INTO TWO IMMUNOLOGICALLY DISTINCT SPECIFIC POLYSACCHARIDES WITH α- AND ß-N-ACETYLGLUCOSAMINYL LINKAGES RESPECTIVELY. *Journal of Experimental Medicine* **120**, 13-29

90. Koc, C., Gerlach, D., Beck, S., Peschel, A., Xia, G., and Stehle, T. (2015) Structural and enzymatic analysis of TarM glycosyltransferase from Staphylococcus aureus reveals an

oligomeric protein specific for the glycosylation of wall teichoic acid. *J Biol Chem* **290**, 9874-9885

91. Sobhanifar, S., Worrall, L. J., King, D. T., Wasney, G. A., Baumann, L., Gale, R. T., Nosella, M., Brown, E. D., Withers, S. G., and Strynadka, N. C. (2016) Structure and Mechanism of Staphylococcus aureus TarS, the Wall Teichoic Acid beta-glycosyltransferase Involved in Methicillin Resistance. *PLoS Pathog* **12**, e1006067

92. Endl, J., Seidl, H. P., Fiedler, F., and Schleider, K. H. (1983) Chemical composition and structure of cell wall teichoic acids of staphylococci. *Archives of Microbiology* **135**, 215-223

93. Allison, S. E., D'Elia, M. A., Arar, S., Monteiro, M. A., and Brown, E. D. (2011) Studies of the genetics, function, and kinetic mechanism of TagE, the wall teichoic acid glycosyltransferase in Bacillus subtilis 168. *J Biol Chem* **286**, 23708-23716

94. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research* **37**, D233-D238

95. Rismondo, J., Percy, M. G., and Gründling, A. (2018) Discovery of genes required for lipoteichoic acid glycosylation predicts two distinct mechanisms for wall teichoic acid glycosylation. *Journal of Biological Chemistry* **293**, 3293-3306

96. Chen, L., Hou, W.-T., Fan, T., Liu, B., Pan, T., Li, Y.-H., Jiang, Y.-L., Wen, W., Chen, Z.-P., Sun, L., Zhou, C.-Z., and Chen, Y. (2020) Cryo-electron Microscopy Structure and Transport Mechanism of a Wall Teichoic Acid ABC Transporter. *mBio* **11**

97. Schirner, K., Stone, L. K., and Walker, S. (2011) ABC Transporters Required for Export of Wall Teichoic Acids Do Not Discriminate between Different Main Chain Polymers. *ACS Chemical Biology* **6**, 407-412

98. Gale, R. T., Li, F. K. K., Sun, T., Strynadka, N. C. J., and Brown, E. D. (2017) B. subtilis LytR-CpsA-Psr Enzymes Transfer Wall Teichoic Acids from Authentic Lipid-Linked Substrates to Mature Peptidoglycan In Vitro. *Cell Chemical Biology* **24**, 1537-1546.e1534

99. Li, F. K. K., Rosell, F. I., Gale, R. T., Simorre, J.-P., Brown, E. D., and Strynadka, N. C. J. (2020) Crystallographic analysis of Staphylococcus aureus LcpA, the primary wall teichoic acid ligase. *Journal of Biological Chemistry* **295**, 2629-2639

100. Schaefer, K., Owens, T. W., Kahne, D., and Walker, S. (2018) Substrate Preferences Establish the Order of Cell Wall Assembly in Staphylococcus aureus. *Journal of the American Chemical Society* **140**, 2442-2445

101. Kawai, Y., Marles-Wright, J., Cleverley, R. M., Emmins, R., Ishikawa, S., Kuwano, M., Heinz, N., Bui, N. K., Hoyland, C. N., Ogasawara, N., Lewis, R. J., Vollmer, W., Daniel, R. A., and Errington, J. (2011) A widespread family of bacterial cell wall assembly proteins. *EMBO J* **30**, 4931-4941

102. Schaefer, K., Matano, L. M., Qiao, Y., Kahne, D., and Walker, S. (2017) In vitro reconstitution demonstrates the cell wall ligase activity of LCP proteins. *Nature Chemical Biology* **13**, 396-401

103. Haas, R., Koch, H. U., and Fischer, W. (1984) Alanyl turnover from lipoteichoic acid to teichoic acid inStaphylococcus aureus. *FEMS Microbiology Letters* **21**, 27-31

104. Fisher, N., Shetron-Rama, L., Herring-Palmer, A., Heffernan, B., Bergman, N., and Hanna, P. (2006) The dltABCD Operon of Bacillus anthracis Sterne Is Required for Virulence and

Resistance to Peptide, Enzymatic, and Cellular Mediators of Innate Immunity. *Journal of Bacteriology* **188**, 1301-1309

105. Peschel, A., Vuong, C., Otto, M., and GöTz, F. (2000) The d-Alanine Residues of Staphylococcus aureus Teichoic Acids Alter the Susceptibility to Vancomycin and the Activity of Autolytic Enzymes. *Antimicrobial Agents and Chemotherapy* **44**, 2845-2847

106. Wood, B. M., Santa Maria, J. P., Matano, L. M., Vickery, C. R., and Walker, S. (2018) A partial reconstitution implicates DltD in catalyzing lipoteichoic acid d-alanylation. *Journal of Biological Chemistry* **293**, 17985-17996

107. Kamar, R., Réjasse, A., Jéhanno, I., Attieh, Z., Courtin, P., Chapot-Chartier, M.-P., Nielsen-Leroux, C., Lereclus, D., el Chamy, L., Kallassy, M., and Sanchis-Borja, V. (2017) DltX of Bacillus thuringiensis Is Essential for D-Alanylation of Teichoic Acids and Resistance to Antimicrobial Response in Insects. *Frontiers in Microbiology* **8**

108. Peschel, A., Otto, M., Jack, R. W., Kalbacher, H., Jung, G., and Götz, F. (1999) Inactivation of the dlt Operon inStaphylococcus aureus Confers Sensitivity to Defensins, Protegrins, and Other Antimicrobial Peptides. *Journal of Biological Chemistry* **274**, 8405-8410

109. Fabretti, F., Theilacker, C., Baldassarri, L., Kaczynski, Z., Kropec, A., Holst, O., and Huebner, J. (2006) Alanine Esters of Enterococcal Lipoteichoic Acid Play a Role in Biofilm Formation and Resistance to Antimicrobial Peptides. *Infection and Immunity* **74**, 4164-4171

110. Bärland, N., Rueff, A.-S., Cebrero, G., Hutter, C. A. J., Seeger, M. A., Veening, J.-W., and Perez, C. (2022) Mechanistic basis of choline import involved in teichoic acids and lipopolysaccharide modification. *Science Advances* **8**, eabm1122

111. Rosenow, C., Ryan, P., Weiser, J. N., Johnson, S., Fontan, P., Ortqvist, A., and Masure, H. R. (1997) Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of Streptococcus pneumoniae. *Molecular Microbiology* **25**, 819-829

112. DiRita, V. J., Gosink, K. K., Mann, E. R., Guglielmo, C., Tuomanen, E. I., and Masure, H. R. (2000) Role of Novel Choline Binding Proteins in Virulence of Streptococcus pneumoniae. *Infection and Immunity* **68**, 5690-5695

113. Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: Structures, Functions, and Mechanisms. *Annual Review of Biochemistry* **77**, 521-555

114. Liu, J., and Mushegian, A. (2003) Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Science* **12**, 1418-1431

115. Zhang, H., Zhu, F., Yang, T., Ding, L., Zhou, M., Li, J., Haslam, S. M., Dell, A., Erlandsen, H., and Wu, H. (2014) The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold. *Nature Communications* **5**, 4339-4339

116. Chang, A., Singh, S., Phillips, G. N., and Thorson, J. S. (2011) Glycosyltransferase structural biology and its role in the design of catalysts for glycosylation. *Current Opinion in Biotechnology* **22**, 800-808

117. Bohl, H., Bai, L., and Li, H. (2021) Recent Progress in Structural Studies on the GT-C Superfamily of Protein Glycosyltransferases. Springer International Publishing. pp 259-271

118. Larivière, L., Gueguen-Chaignon, V., and Moréra, S. (2003) Crystal Structures of the T4 Phage β-Glucosyltransferase and the D100A Mutant in Complex with UDP-glucose: Glucose Binding and Identification of the Catalytic Base for a Direct Displacement Mechanism. *Journal of Molecular Biology* **330**, 1077-1086

119. Mulichak, A. M., Losey, H. C., Lu, W., Wawrzak, Z., Walsh, C. T., and Garavito, R. M. (2003) Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. *Proceedings of the National Academy of Sciences* **100**, 9238-9243

120. Mulichak, A. M., Lu, W., Losey, H. C., Walsh, C. T., and Garavito, R. M. (2004) Crystal Structure of Vancosaminyltransferase GtfD from the Vancomycin Biosynthetic Pathway: Interactions with Acceptor and Nucleotide Ligands†,‡. *Biochemistry* **43**, 5170-5180

121. Zhu, F., Zhang, H., Yang, T., Haslam, S. M., Dell, A., and Wu, H. (2016) Engineering and Dissecting the Glycosylation Pathway of a Streptococcal Serine-rich Repeat Adhesin. *Journal of Biological Chemistry* **291**, 27354-27363

122. Lade, H., and Kim, J.-S. (2021) Bacterial Targets of Antibiotics in Methicillin-Resistant Staphylococcus aureus. *Antibiotics* **10**, 398

123. Wanner, S., Schade, J., Keinhörster, D., Weller, N., George, S. E., Kull, L., Bauer, J., Grau, T., Winstel, V., Stoy, H., Kretschmer, D., Kolata, J., Wolz, C., Bröker, B. M., and Weidenmaier, C. (2017) Wall teichoic acids mediate increased virulence in Staphylococcus aureus. *Nature Microbiology* **2**, 16257-16257

124. Weidenmaier, C., Peschel, A., Xiong, Y.-Q., Kristian, S. A., Dietz, K., Yeaman, M. R., and Bayer, A. S. (2005) Lack of Wall Teichoic Acids in Staphylococcus aureus Leads to Reduced Interactions with Endothelial Cells and to Attenuated Virulence in a Rabbit Model of Endocarditis. *The Journal of Infectious Diseases* **191**, 1771-1777

125. Zhu, X., Liu, D., Singh, A. K., Drolia, R., Bai, X., Tenguria, S., and Bhunia, A. K. (2018) Tunicamycin Mediated Inhibition of Wall Teichoic Acid Affects Staphylococcus aureus

and Listeria monocytogenes Cell Morphology, Biofilm Formation and Virulence. *Frontiers in Microbiology* **9**

126. Zhang, Y. H., Ginsberg, C., Yuan, Y., and Walker, S. (2006) Acceptor substrate selectivity and kinetic mechanism of Bacillus subtilis TagA. *Biochemistry* **45**, 10895-10904

Chapter 2

The structural and mechanistic role of the TagA glycosyltransferase core domain in wall teichoic acid synthesis

2.1 Overview

Methicillin-resistant *Staphylococcus aureus* (MRSA) is a clinically important Gram-positive pathogen that accounts for more than 300,000 infections and 10,000 deaths in the United States annually. In *S. aureus* and other Gram-positive bacteria, wall teichoic acid (WTA) glycopolymers constitute a major component of the cell wall and have important functions in cell division, morphology, adhesion, and pathogenesis. WTA biosynthesis has drawn considerable interest in novel antibiotic development because genetic elimination of TagA, which catalyzes the first committed step in WTA biosynthesis, re-sensitizes MRSA to β-lactam antibiotics. Prior to this work, little was known about the structure and mechanism of TagA. This chapter discusses the first crystal structure of the TagA core domain ($TagA^{\Delta C}$) and a mechanistic analysis of key catalytic residues in $TagA^{\Delta C}$. Biochemical and cellular data suggest that TagA activates via interactions with the cell membrane, which trigger a conformational change that structures a dynamic C-terminal appendage and facilitates the formation of a competent active site. My contributions to this manuscript included the identification and alteration of key active site residues, the design of C-terminal tail mutants to stabilize full-length TagA and dissociate the enzyme from the *Bacillus subtilis* cell membrane, the determination of protein oligomerization in solution by sedimentation equilibrium and size exclusion chromatography experiments, data interpretation, figure generation, and preparation of the resulting manuscript.

This chapter is reformatted with permission from a peer-reviewed research article in the journal PLOS Pathogens: “Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria.” Kattke, M.D.*, Gosschalk, J.E.*, Martinez, O.E., Kumar, G., Gale, R.T., Cascio, D., Sawaya, M.R., Phillips, M.L., Brown, E.D., Clubb, R.T., *PLoS Pathogens*. **15**(4), e1007723 (2019).

2.2 Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria

2.2.1 Abstract

Staphylococcus aureus and other bacterial pathogens affix wall teichoic acids (WTAs) to their surface. These highly abundant anionic glycopolymers have critical functions in bacterial physiology and their susceptibility to β-lactam antibiotics. The membrane-associated TagA glycosyltransferase (GT) catalyzes the first-committed step in WTA biosynthesis and is a founding member of the WecB/TagA/CpsF GT family, more than 6,000 enzymes that synthesize a range of extracellular polysaccharides through a poorly understood mechanism. Crystal structures of TagA from *T. italicus* in its apo- and UDP-bound states reveal a novel GT fold, and coupled with biochemical and cellular data define the mechanism of catalysis. We propose that enzyme activity is regulated by interactions with the bilayer, which trigger a structural change that facilitates proper active site formation and recognition of the enzyme's lipid-linked substrate. These findings inform upon the molecular basis of WecB/TagA/CpsF activity and could guide the development of new anti-microbial drugs.

2.2.2 Author Summary

Gram-positive bacteria cause thousands of deaths in the United States each year and are a growing health concern because many bacterial strains have become resistant to commonly used antibiotics. One of the most abundant polymers displayed on the surface of Gram-positive bacteria is wall teichoic acid (WTA), a negatively charged carbohydrate polymer that has critical functions in cell division, morphology, adhesion and pathogenesis. The WTA biosynthetic pathway has

drawn significant interest as a drug target because clinically important methicillin-resistant *S. aureus* (MRSA) strains that lack WTA are defective in host colonization and re-sensitized to β-lactam antibiotics. To understand how bacteria produce WTA, we determined the structure and deduced the enzymatic mechanism of TagA, an important enzyme that is required for WTA synthesis. This research reveals a new method for enzyme regulation, whereby peripheral membrane association enables TagA to adopt its active form as a monomer. As TagA enzymes are highly conserved in bacteria, they can be expected to operate through a similar mechanism. The results of this work provide insight into WTA biosynthesis and could lead to innovative approaches to treat infections caused by pathogenic bacteria.

2.3 Introduction

The thick peptidoglycan (PG) sacculus that surrounds Gram-positive bacteria maintains cellular integrity and is affixed with proteins and glycopolymers that have important roles in microbial physiology and host-pathogen interactions. In *Staphylococcus aureus* and other Gram-positive bacteria, wall teichoic acids (WTAs) are a major component of the cell wall, constituting up to 60% of its dry mass [1]. WTAs have essential functions, including regulating PG biosynthesis, morphogenesis, autolysin activity, immune evasion, resistance to host cationic antimicrobial peptides, and pathogenesis [2–9]. The WTA biosynthetic pathway has drawn considerable interest as an antibiotic target, as genetically eliminating WTA production in clinically important Methicillin-resistant *Staphylococcus aureus* (MRSA) re-sensitizes it to β-lactam antibiotics and attenuates its virulence [2, 3].

WTA polymers are constructed from polymerized alditol-phosphate subunits that are attached to the cell wall via a disaccharide-containing linkage unit. While the chemical structure

of the main chain of the polymer can vary, the structure of the linkage unit is highly conserved across different species of Gram-positive bacteria and is composed of an *N*-acetylmannosamine (ManNAc) (β1→4) *N*-acetylglucosamine (GlcNAc) disaccharide appended to one to three glycerol-3-phosphate (GroP) groups [10] (Fig. 2.1A). The linkage unit performs a key function in WTA display, connecting the WTA polymer to the C6 hydroxyl of PG's *N*-acetylmuramic acid (MurNAc) [11]. WTA is synthesized on the cytoplasmic face of the cell membrane by modifying a membrane-embedded undecaprenyl-phosphate (C_{55}-P) carrier. In *Bacillus subtilis* and *S. aureus*, the conserved GlcNAc-ManNAc-GroP linkage unit is first synthesized by the sequential action of the TagO, TagA, and TagB enzymes (originally designated TarOAB in *S. aureus*). TagO initiates WTA synthesis by transferring GlcNAc from the UDP-activated sugar to the C_{55}-P carrier to produce lipid-α[12]. The TagA glycosyltransferase (GT) then appends ManNAc from a UDP-ManNAc donor, producing a C_{55}-PP-GlcNAc-ManNAc disaccharide-lipid product (lipid-β) [13, 14].

Linkage unit synthesis is then completed by TagB, which appends a single GroP to lipid-β using a CDP-glycerol substrate that is produced by the TagD enzyme [13]. In *S. aureus*, the linkage unit is then primed by TarF, which attaches at least one additional GroP. The TarIJL enzymes then construct the main chain of the polymer by adding 40–60 ribitol-5-phosphate (RboP) units [15–17]. After being modified with GlcNAc by TarM and TarS, the TarGH ABC-like transporter exports the polymer to the cell surface, where it is further modified with D-alanine to tune its electrostatic properties [18]. The assembled polymer is then covalently attached to the cell wall by an LCP ligase, which catalyzes a phosphotransfer reaction that joins WTA via its linkage unit to PG's MurNAc [19–21]. *B. subtilis* also uses functionally analogous enzymes to produce strain-specific GroP (spp. 168) or RboP (spp. W23) WTA polymers. Recent structural studies have begun

to reveal the mechanism through which bacteria produce WTA, including how the polymerization is primed (TarF), and how WTA is modified with GlcNAc (TarM and TarS) and attached to the cell wall (LCP ligases) [22–24]. However, it remains unknown exactly how bacteria produce the highly-conserved linkage unit that connects WTA to the cell wall.

The TagA N-acetylmannosamine transferase catalyzes the first committed step in WTA biosynthesis. It is an attractive target for new therapeutics aimed at treating MRSA infections, as *tagA*- strains are attenuated in virulence and re-sensitized to methicillin, imipenem, and ceftazidime. [25–28]. TagA is also a founding member of the WecB/TagA/CpsF family of GTs (PFAM03808; CAZy GT26), which has over 6,000 members [29, 30]. In addition to WTA, these enzymes synthesize a range of important surface-associated and secreted glycopolymers that function as virulence factors, including capsular polysaccharides of Group B Streptococcus (GBS) and the enterobacterial common antigen present in the outer-membrane of *Escherichia coli* and other Gram-negative bacteria [31–33]. Industrially, this family is important as it includes the GumM GT, an essential enzyme in xanthan gum synthesis in *Xanthomonas campestris* [34]. WecB/TagA/CpsF GTs are distinguished by their ability to elaborate membrane-embedded polyprenol substrates, but the molecular basis of their function remains unknown. Here, we report the crystal structure and biochemical studies of TagA from *Thermoanaerobacter italicus*, a close homolog of *S. aureus* TagA. Our results reveal that WecB/TagA/CpsF enzymes adopt a unique GT-E fold and shed considerable light onto their mechanism of catalysis. We propose that membrane association activates TagA by triggering a unique dimer to monomer quaternary structural change that facilitates lipid-α recognition and the formation of a catalytically competent active site. The results of these structural and mechanistic studies represent a major advancement

in our understanding of WTA biosynthesis that could facilitate the discovery of new antibiotics that work by disrupting the synthesis of this important bacterial surface polymer.

2.4 Results and Discussion

2.4.1 WecB/TagA/CpsF Enzymes are Structurally Novel Glycosyltransferases

To gain insight into the mechanism of catalysis, we determined the 2.0 Å crystal structure of the TagA enzyme from *Thermoanaerobacter italicus* ($TagA^{\Delta C}$, residues Met1-Gly195) (Fig. 2.1B). $TagA^{\Delta C}$ contains the highly conserved amino acid region that defines the WecB/TagA/CpsF family (Fig. S2.1), but lacks 49 C-terminal residues that target the protein to the membrane (*see below*). $TagA^{\Delta C}$ is much better suited for structural analyses, since unlike the full-length protein, it does not require high concentrations of salt and glycerol to be solubilized. Selenomethionine (SeMet)-labeled $TagA^{\Delta C}$ in its apo-form crystallized in the $P2_1$ space group as a dimer, with eight molecules per asymmetric unit. The structure was determined using the multiple anomalous dispersion (MAD) method and is well-defined by continuous electron density. The subunits in the dimer are related by two-fold non-crystallographic symmetry and possess similar atomic structures; their heavy atom coordinates can be superimposed with an RMSD of 0.11 Å. All four dimers in the asymmetric unit also superimpose with similarly low RMSD (average of 0.41 Å). Complete data collection and structural statistics are provided in Table 2.1.

TagA adopts a unique α/β tertiary structure that differs markedly from previously described GTs [35, 36]. Each protomer consists of eight β–strands and nine α-helices that form two distinct regions (Fig. 2.1C). The N-terminal region of TagA is formed by helices H2 to H4 that pack against a β-hairpin constructed from strands β1 and β2, while its larger C-terminal region is comprised of

a six-stranded parallel β-sheet (β–strands β4, β3, β8, β7, β5, and β6) that is surrounded by seven helices (helices H1, H3, and H5 to H9). The arrangement of the six parallel strands forming the β–sheet resembles a Rossmann fold commonly found in nucleotide-binding proteins. The regions are interconnected, with the N-terminal β-hairpin forming a single backbone-backbone hydrogen bond to strand β4 within the C-terminal region (between the amide of Asp13 (β2) and the carbonyl of Asn63 (β3)). In the dimer, the C-terminal helix H8 in one subunit packs against helices H2 and H4 located in the N-terminal region of the adjacent protomer, burying 1,212 $Å^2$ of solvent accessible surface area to produce a narrow pore (detailed in Fig. S2.6). Analytical ultracentrifugation (AUC) experiments indicate that $TagA^{\Delta C}$ dimerizes in solution with relatively weak affinity, with a monomer-dimer dissociation constant (K_d) of only 7.4 ± 0.7 μM (Fig. S2.3A). The modest dimer affinity may in part be due to the fact that the dimer interface is discontinuous (Fig. 2.1D).

To determine the biological relevance of the oligomeric interface observed in the crystal structure we analyzed the SeMet-labeled $TagA^{\Delta C}$ structure with the Evolutionary Protein-Protein Classifier (EPPIC) program [37]. EPPIC indicates with 99% certainty that the dimeric interface is biologically relevant. Just as SeMet-$TagA^{\Delta C}$ crystallizes as a dimer, native $TagA^{\Delta C}$ is dimeric in solution (Fig. S2.4C-D). Based on the crystal structure of the SeMet-labeled $TagA^{\Delta C}$ protein the side chains of Val43 and Ala72 reside at the dimer interface, burying 99% and 84% of their surface area (Fig. S2.6). Consistent with the native and SeMet-labeled proteins adopting similar quaternary structures, introduction of V43E and A72R mutations into $TagA^{\Delta C}$ disrupts dimerization (Fig. S2.4C-D). Combined, these results substantiate the conclusion that the dimeric interface visualized by crystallography is biologically relevant and that it is present in solution in the native form of the protein. As WecB/TagA/CpsF enzymes exhibit related primary sequences, it is expected that they will adopt tertiary structures that are similar to that observed for TagA.

Glycosylation reactions catalyzed by GT enzymes play a central role in biology, creating an enormous array of biologically important oligosaccharides and glycoconjugates. Interestingly, the array of enzymatic machinery used to perform glycosylation is surprisingly simple, and only four distinct GT protein folds have been identified that are capable of glycosyltransferase activity (termed GT-A, GT-B, GT-C, and GT-D enzymes) [32, 36, 38]. TagA$^{\Delta C}$ differs markedly from all of these enzymes based on its tertiary structure and how its secondary structural elements are arranged. Notably, TagA lacks the canonical Asp-X-Asp motif found in GT-A enzymes that participate in nucleotide binding and it differs substantially from GT-B and GT-C class enzymes that adopt multi-domain structures [32, 36]. Interestingly, TagA does exhibit limited structural homology with DUF1792 (PDB ID: 4PFX), the founding member of the GT-D family that transfers glucose from UDP-glucose onto protein O-linked hexasaccharides [38]. The backbone coordinates of a subset of residues within the TagA$^{\Delta C}$ and DUF1792 structures can be superimposed with an RMSD of 3.7 Å (Fig. S2.2A). However, consistent with these enzymes sharing only 15% sequence identity, the arrangement, number, and topology of their secondary structural elements are distinct (Fig. S2.2B). Furthermore, the enzymes have different catalytic mechanisms, as TagA exhibits ion-independent glycosyltransferase activity and it lacks the conserved Asp-X-Glu motif present in DUF1792 that coordinates an Mn^{2+} ion cofactor [14]. Thus, TagA and related members of the WecB/TagA/CpsF family adopt a novel glycosyltransferase fold, which we term GT-E.

2.4.2 Active site architecture

To define the enzyme active site, we determined the structure of the SeMet-labeled TagA$^{\Delta C}$:UDP complex at 3.1 Å resolution (Fig. 2.2A). The crystal structure visualizes the enzyme-

product complex, as steady-state kinetics studies of *B. subtilis* TagA have shown that the enzyme operates via an ordered Bi-Bi mechanism in which UDP-ManNAc binds first and UDP is released last [14]. The complex crystallizes as a dimer of trimers in the $P2_1$ space group, with six molecules in each asymmetric unit (Fig. S2.5). The coordinates of the apo- and UDP-bound forms of $TagA^{\Delta C}$ are nearly identical (RMSD of 0.18 Å), suggesting that the enzyme binds UDP through a lock-and-key mechanism. In each protomer, UDP contacts the β7-H8 motif within the Rossmann-like fold, as well as the C-terminal edge of strand β5 (Fig. 2.2A). The uracil base is engaged in pi stacking with Tyr137 (β6-H7 loop), while Asp191 (H9) contacts the ribose sugar. The trimeric oligomer observed in the structure of the complex is presumably an artifact of crystallization, as AUC experiments using the native $TagA^{\Delta C}$ performed in the presence of saturating amounts of UDP indicate that similar to apo-$TagA^{\Delta C}$, the UDP-bound protein fits best to a monomer-dimer equilibrium model rather than monomer-trimer or dimer-trimer equilibrium models (K_d = 3.5 ± 0.4 μM in the presence of 10:1 UDP:TagA) (Fig. S2.3A). Moreover, an EPPIC analysis of the trimer structure indicates that its interfaces are unlikely to be biologically relevant (14% confidence level). In the complex, each UDP molecule is primarily contacted by a single subunit. However, the C2 hydroxyl group in each UDP molecule forms a hydrogen bond with the backbone carbonyl of Val192 (H9) located in the adjacent subunit, which may explain why the complex crystallized as a trimer.

Intriguingly, positioned immediately adjacent to the UDP binding site on each protomer is a large pocket that harbors several phylogenetically conserved amino acids that are important for catalysis (Fig. 2.2B). The pocket resides near the C-terminal end of the parallel β3 and β8 strands and has walls that are formed by residues located in helices H5 and H8, as well as residues within the polypeptide segments that connect strand β7 to helix H8, and strand β3 to helix H2. Several

highly-conserved residues are located within the pocket and its periphery, including Thr37, Asn39, Asp65, Arg83, Gln167, and Glu168 (Fig. 2.2A and Fig. S2.1). Interestingly, in the crystal structure of the TagA$^{\Delta C}$:UDP complex, the β-phosphate group of UDP extends inward toward the pocket. As UDP is a competitive inhibitor of the UDP-ManNAc substrate, it is likely that these ligands bind to the same site on the enzyme, such that the ManNAc moiety within UDP-ManNAc is projected into the pocket where it can interact with the conserved side chains of residues Thr37, Gln167 or Glu168 [14]. In an effort to better understand how TagA processes its substrates, we modeled how lipid-α and UDP-ManNAc might bind to TagA. Binding of UDP-ManNAc was modeled using the coordinates of the TagA$^{\Delta C}$:UDP crystal structure to position the uracil and ribose components, enabling manual placement of the ManNAc moiety. Lipid-α was then docked *in silico* using Autodock vina (see Materials and Methods). These docking experiments suggest that the UDP-ManNAc and lipid-α substrates bind to opposite sides of the conserved pocket [39]. In models of the enzyme-substrate ternary complex, the GlcNAc and diphosphate portion of lipid-α are positioned near residues that connect strands β4 to helix H4, while the undecaprenyl chain of lipid-α exits near the C-terminus of the TagA$^{\Delta C}$ protomer (Fig. 2.2C). In this binding mode, the sugar acceptor's C-4 hydroxyl group is positioned near the side chain of Asp65, while the highly-conserved side chains of Arg83 and Asn39 are adjacent to lipid-α's diphosphate and GlcNAc, respectively.

To investigate the importance of conserved pocket residues, we reconstituted its GT activity *in vitro* using UDP-ManNAc and a lipid-α analog that replaces its undecaprenyl chain with tridecane, as previously reported [14, 40]. The full-length TagA enzymes from *T. italicus* and *S. aureus* exhibit similar transferase activities *in vitro*, producing the UDP product at a rate of 1.5 μM min^{-1} and 1.0 μM min^{-1} at 200 nM enzyme concentration, respectively (Fig. 2.2D). This is

expected, as all TagA homologs presumably catalyze the synthesis of the ManNAc(β1→4)GlcNac glycosidic bond within the conserved linkage unit. Thr37Ala and Asp65Ala mutations in TagA cause the largest decreases in activity (0.41 μM min^{-1} and 0.052 μM min^{-1}, respectively), compatible with these residues residing within the enzyme's active site. As discussed later, the carboxyl side chain of Asp65 is poised to function as general base that deprotonates the nucleophilic C4 hydroxyl group in GlcNAc, while the side chain of Thr37 may stabilize the orientation of ManNAc by interacting with its C5 hydroxyl group.

2.4.3 A conserved C-terminal appendage is required for catalysis

Surprisingly, the mutational analysis reveals that only the full length TagA protein is enzymatically active *in vitro*, while the truncated $TagA^{\Delta C}$ protein used for crystallography and modeling is catalytically inactive (Fig. 2.2D). This is compatible with the high level of primary sequence conservation of C-terminal residues within WecB/TagA/CpsF enzymes and suggests that the deleted appendage may be necessary to construct a catalytically competent active site (Fig. S2.1). To gain insight into the function of the appendage, we utilized the structure of the TagA protein modeled using Generative Regularized Models of Proteins (GREMLIN), a recently developed protein modeling server that predicts tertiary structure by exploiting sequence conservation and amino acid co-evolutionary patterns [41, 42]. Only GREMLIN-predicted structures in which TagA was assumed to be monomeric yielded favorable results, and is substantiated by important correlations between residues within the body of $TagA^{\Delta C}$ and the appendage (in *T. italicus*: Arg83 and Arg205, Ala72 and Val235, Gln47 and Lys234, Val197 and Glu218/Lys217, and Lys198 and Lys217 couplings). In general, the tertiary structures of the GREMLIN-predicted model of monomeric TagA ($TagA^{GM}$) and the experimentally determined

crystal structures of $TagA^{\Delta C}$ are similar. However, only in $TagA^{GM}$ is the C-terminal appendage present, which forms three α–helices (H10 to H12) that pack against the active site harboring the catalytically important Asp65 and Thr37 residues. The C-terminal appendage is presumably unstructured in dimeric forms of the enzyme, as, importantly, the contact surface used to engage the appendage is occluded by inter-subunit interactions in the crystal structures of $TagA^{\Delta C}$.

The model of the $TagA^{GM}$ monomer provides insight into why the C-terminal appendage is critical for catalysis, as it predicts several highly-conserved arginine residues project their side chains into the enzyme's active site (Arg214, Arg221 and Arg224) (Fig. 2.3A). Of particular interest is the side chain of Arg221 within helix H11 of the appendage, as modeling of $TagA^{GM}$ bound to its substrates suggests that the Arg221 guanidino group may stabilize the β-phosphate of the UDP leaving group (Fig. 2.4A). Indeed, Arg221Glu mutation in TagA significantly abates TagA GT activity, strongly suggesting that only the monomeric form of TagA is enzymatically active (Fig. 2.2D). While not tested in this study, the side chains of Arg214 and Arg224 may also be important for catalysis, as some cation-independent GTs use more than one positively charged residue to stabilize phosphotransfer reaction intermediates [43].

2.4.4 Membrane-induced structural changes likely facilitate the recognition of bilayer-embedded polyprenol substrates

To build the linkage unit, TagA should associate with the cytoplasmic membrane where it attaches ManNAc to its lipid-α substrate. Intriguingly, only the monomeric form of TagA contains a non-polar surface patch that is suitable for interacting with the lipid bilayer. In monomeric $TagA^{GM}$, helices H11 and H12 within the C-terminal appendage project several non-polar side chains into the solvent for potential membrane interactions (e.g. Leu212, Ile216, and Ile233, in *T.*

italicus TagA) (Fig. 2.3B) [44–48]. To investigate the role of the C-terminal appendage in membrane binding, we determined how this structural element affected TagA localization in *B. subtilis*, since unlike *T. italicus*, robust tools are available to genetically manipulate this model Gram-positive bacterium. Importantly, the TagA homologs in these organisms are related (34% sequence identity) and both contain the conserved C-terminal appendage (Fig. S2.1). *B. subtilis* cells expressing hexahistidine-tagged TagA (^{Bs}TagA) were fractionated and analyzed by Western blotting. ^{Bs}TagA peripherally associates with the membrane, since it is present in the membrane fraction, but released into the soluble fraction after adding either potassium hydroxide or the chaotropic salt sodium iodide (Fig. 2.3C-D, respectively). Interestingly, TagA lacking the C-terminal appendage (^{Bs}TagA$^{\Delta C}$, residues Met1-Val196) primarily partitions into the soluble fraction, indicating that this structural element tethers the protein to the membrane (Fig. 2.3C).

To identify specific residues required for peripheral membrane binding and to validate the structural model of the full-length TagA^{GM} protein, we constructed a ^{Bs}TagA mutant that replaces surface-exposed hydrophobic residues with hydrophilic substitutes within helix H11 (^{Bs}TagA$^{\Delta H11}$, ^{Bs}TagA containing L208Q, F211K, L215E mutations, yellow colored residues in Fig. 2.3B). Fractionation studies reveal that ^{Bs}TagA$^{\Delta H11}$ is significantly more solubilized into the cytoplasm compared to wild-type ^{Bs}TagA, which is exclusively found in the membrane (Fig. 2.3C). These results substantiate the predicted structure of the monomer and suggest that monomeric TagA associates with the membrane via surface-exposed non-polar residues located within helix H11. Other bacterial enzymes are targeted to the membrane via terminal helices (e.g. TagB, FtsA, MinD, PBP enzymes), but TagA is novel because its helices likely form an integral part of the active site upon membrane association [44–48].

Biochemical studies suggest that prior to engaging the membrane, full-length TagA in the cytoplasm is primarily dimeric and in equilibrium with its monomeric form. This is supported by chemical crosslinking studies of cells expressing *T. italicus* TagA that show the enzyme exists as a mixture of dimeric and monomeric species (Fig. S2.3C). It is also substantiated by SEC experiments that reveal the full-length protein is primarily dimeric in aqueous solvent, but that monomeric and higher-order species are also present (Fig. S2.4A-B). Interestingly, our results reveal that this equilibrium can be shifted toward the monomeric form by reducing the hydrophobicity of the membrane-targeting C-terminal appendage. We constructed a full-length *T. italicus* TagA mutant that contained four amino acid substitutions: I203E, L209Q, L212K and I216E. Based on the $TagA^{GM}$ model, these alterations are expected to increase the polarity of the membrane-targeting patch that is formed by helices H10-H12. Unlike the dimeric native protein, the mutant is primarily monomeric according to SEC (Fig. S2.4A-B). This finding is compatible with native TagA in aqueous solution being in equilibrium between dimeric and monomeric forms. The mutations shift the equilibrium toward the monomeric state, presumably by reducing the entropic cost associated with clustering these non-polar residues together, which is only expected to occur in the monomer. Thus, we conclude that in the cell, TagA toggles between at least two distinct states: 1) membrane associated enzymatically active monomers, in which the C-terminal appendage contributes Arg221 to the active site to facilitate catalysis, and 2) inactive dimers in the cytoplasm, in which inter-subunit interactions mask an incompletely formed active site.

Collectively, our data suggest that TagA is activated through a unique membrane-association mechanism that is mediated by residues at its C-terminus. TagA is in equilibrium between monomeric and dimeric states. Removed from the membrane, it is primarily dimeric with interfacial interactions obstructing the binding site for its C-terminal appendage and holding the

enzyme in an inactive, dimeric state (Fig. 2.4B). However, upon encountering the membrane, we posit that non-polar side chains within the appendage are inserted into the bilayer, thereby nucleating its folding and the formation of a monomeric, catalytically functional enzyme in which the appendage contributes key active site residues. In the $TagA^{GM}$ monomer, a gap is located between helices H11 and H12 in the C-terminal appendage, forming a short pore that connects the active site to the protein's membrane binding surface. Thus, membrane induced folding of the C-terminal appendage may also facilitate recognition of the lipid-α by providing an additional binding site for several of its non-polar prenyl groups. While the catalytic activities of other membrane-associated enzymes are regulated by lipid bilayer-induced quaternary structural changes or allosteric mechanisms, to the best of our knowledge, the TagA activation mechanism outlined here is unique [49].

In conclusion, our results provide direct insight into how TagA enzymes synthesize the conserved linkage unit used to attach WTA to the cell wall and, more generally, how members of the large WecB/TagA/CpsF GT family produce a range of important surface-associated and secreted bacterial glycopolymers. TagA is a structurally unique GT that we propose defines a new GT-E fold. Removed from the membrane, it forms catalytically inactive dimers that are presumably incapable of mediating spurious GT reactions or hydrolyzing UDP-ManNAc. However, upon encountering the membrane containing its lipid-α substrate, conserved C-terminal residues in TagA fold into an essential active site appendage, stabilizing the monomeric and catalytically-active form of the enzyme. From our structural and biochemical data, a working model of the catalytic mechanism can be proposed (Fig. 2.4C). Previous studies have shown that catalysis likely occurs via an S_N2-like mechanism that inverts the anomeric stereochemistry of ManNAc [14]. It seems likely that Asp65 functions as a base, deprotonating the GlcNAc C4

hydroxyl in lipid-α. This would facilitate its nucleophilic attack at the anomeric carbon of ManNAc, resulting in an oxocarbenium ion-like transition state that is stabilized by electrostatic interactions between Arg221, donated by the C-terminal appendage, and the diphosphate moiety of UDP. Thr37 within the conserved pocket may play an important role in orienting the ManNAc moiety of the sugar donor, poising its electrophilic center for glycosidic bond formation. A complete understanding of the catalytic mechanism will require the structure determination of additional key reaction intermediates. As WTA and other bacterial glycopolymers are critical components of the cell wall, the results of these studies are of fundamental importance and could facilitate the discovery of GT-E (WecB/TagA/CpsF) specific enzyme inhibitors that could be useful antibiotics.

2.5 Materials and Methods

2.5.1 Cloning, expression, protein purification, and crystallization

Bacterial strains and plasmids used in this study are listed in SI Materials and Methods, Table S2.1. Protocols for protein purification, crystallization, and structure determination are detailed in SI Materials and Methods.

2.5.2 Mutagenesis and activity assays

Conservation analysis was conducted using the Consurf Server [50–52]. The lipid-α or UDP-ManNAc substrates were generated *in silico* using the Phenix electronic Ligand Builder and Optimization Workbench (Phenix.eLBOW) [53]. Cis-trans configuration and stereochemistry were confirmed or corrected using the Phenix Restraints Editor Especially Ligands

(Phenix.REEL). The substrates were docked to the 2.0 Å resolution protomer structure using Autodock vina with a 25 x 25 x 25 Å search space and an exhaustiveness of 18 [39]. TagA *in vitro* enzyme activity was determined using an anion-exchange HPLC system and the following assay conditions: 0.2 μM TagA; 100 μM lipid-α; 100 μM UDP-GlcNAc; 3 μM MnaA; 50 mM Tris-HCl, pH 7.5; and 250 mM NaCl [40]. Reactions were incubated for 40 minutes before quenching with 3 M urea. The reactions were separated with a DNAPak PA200 anion-exchange column using a buffer gradient of 100% Buffer A (20 mM NH4HCOO3, and 10% MeCN, pH 8.0) to 90% Buffer A plus 10% Buffer B (20 mM NH4HCOO3, 10% MeCN, and 1 M NaCl, pH 8.0) over 10 minutes. UDP-ManNAc or UDP elution peaks were monitored at 271 nm and integrated to determine turnover rate.

2.5.3 Cell fractionation

Overnight *B. subtilis* cultures containing selective antibiotics were diluted 1:100 into 1 L of fresh LB broth containing 1 mM IPTG. Cultures were incubated at 37°C and 250 rpm until an OD_{600} of 1.0 was reached. Cells were pelleted at room temperature, washed once with PBS, and then frozen at -80°C. Cells were fractioned as previously reported with several exceptions [48]. Cells were re-suspended in 10 mL of Lysis Buffer (PBS, pH 7.3; 1 mM EDTA; 1 mM DTT; 10 μg/mL RNase; 10 μg/mL DNase; 2 mM PMSF, and 100 μL Protease inhibitor cocktail) and sonicated on ice for eight minutes, with one minute "on" (one second pulses for 60 seconds) and one minute "off" (no pulses for 60 seconds). Ten milliliters of lysate were divided into two 5 mL volumes and centrifuged in a Beckman type 50 TI rotor in a Beckman XPN-100 preparative ultracentrifuge. The lysate was centrifuged at 9,600 rpm for 10 minutes to remove cellular debris. The supernatant was collected and centrifuged at 21,000 rpm. The supernatant was removed and

spun for a third time at 40,000 rpm for one hour. This final supernatant fraction was considered to be the soluble fraction, and the remaining pellet, which contained the membrane fraction was re-suspended in 1 mL of ice-cold Lysis Buffer. The membrane fraction was diluted ten-fold for comparison with the supernatant and samples were run on an SDS-PAGE gel for 50 minutes at 170 V and transferred to a PVDF membrane using an iBlot transfer device (ThermoFisher Scientific). The membrane was fixed in methanol for five minutes, briefly washed in water, and then blocked over-day in TBST Blocking Buffer (20 mM Tris-HCl, pH 7.5; 500 mM NaCl; 0.05% Tween; and 5% w/v nonfat milk). The membrane was washed, incubated with primary antibody (Invitrogen #MA-21215, mouse anti-6xHis, 1:1000 dilution in TBST Blocking Buffer) overnight at 4°C, washed again, and incubated with secondary antibody (Sigma-Aldrich #A9044, anti-mouse horseradish peroxidase). The membrane was incubated with SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific #34080) for five minutes and exposed to radiography film for thirty seconds to ten minutes.

2.5.4 Chaotropic agent analysis

Membrane fractions were prepared as described above and re-suspended to 600 µL. Fractionation was performed as previously reported, with some minor changes [54]. 200 µL of membrane resuspension was diluted 1:3 in PBS containing 1.5 M sodium iodide, 0.1 N potassium hydroxide, or PBS alone. Samples were loaded onto a 2.4 mL sucrose cushion (PBS and chaotrope, 0.5 M sucrose) and centrifuged for 30 minutes at 40,000 rpm. The top (800 µL) and middle (2.4 mL) fractions were removed, and the pellet was re-suspended in 200 µL of PBS. The top and middle fractions were precipitated with 10% trichloroacetic acid on ice for 30 minutes, centrifuged at 20,000 *g* for 10 minutes, and then re-suspended in 200 uL of 8 M urea. Samples were mixed 1:1

with 2X SDS loading dye (100 mM Tris base, 200 mM DTT, 4% SDS, 0.2% bromophenol blue, 20% glycerol). Immunoblotting was performed as described above.

2.6 Figures

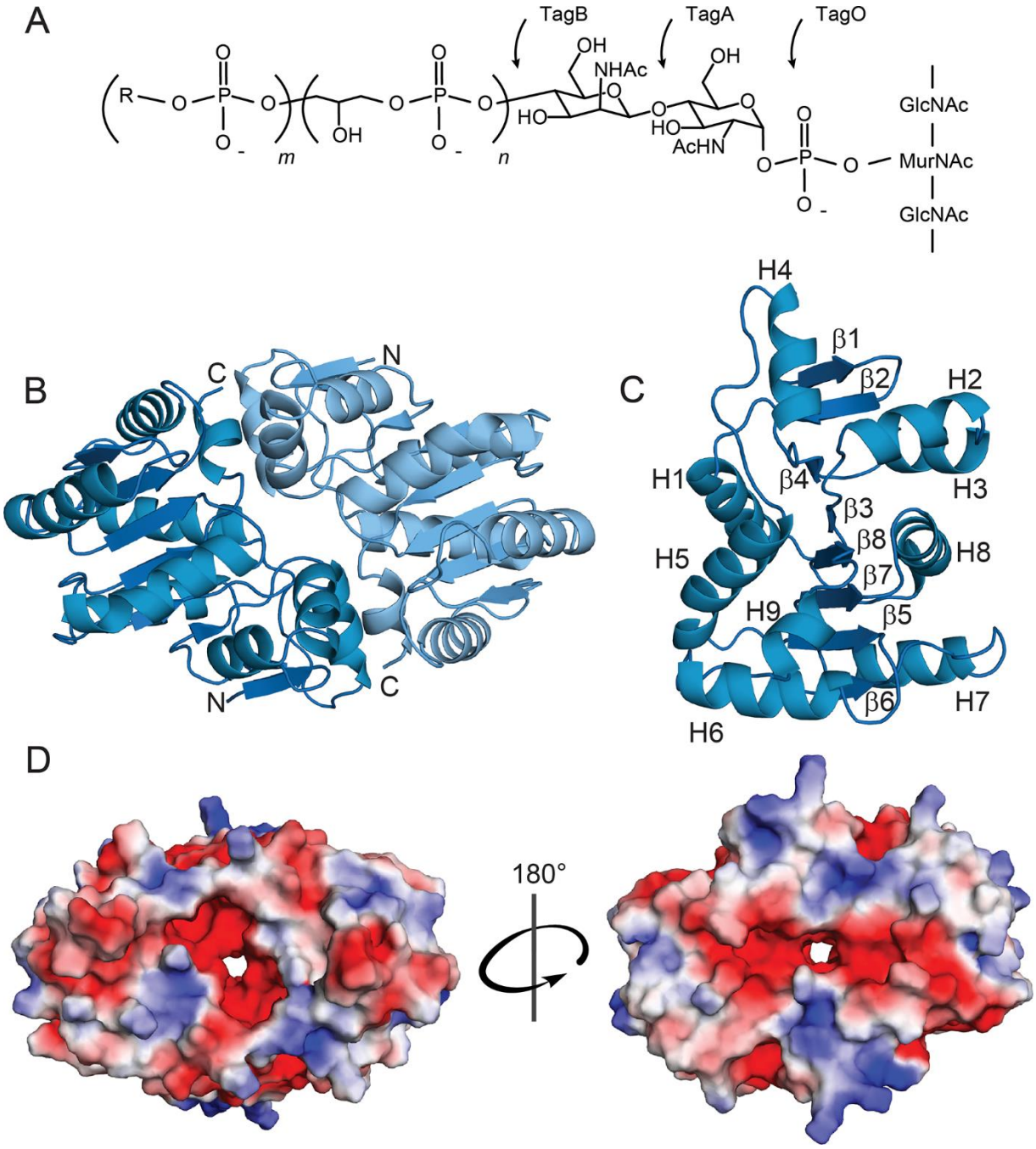

Figure 2.1 The WTA linkage unit and TagA structural characteristics. (A) WTA linkage unit and polymer. The linkage unit is attached to the peptidoglycan via the C6-hydroxyl of *N*-acetylmuramic acid and is composed of a GlcNAc (TagO-catalyzed), ManNAc (TagA-catalyzed), and n = 2–3 glycerol phosphates (TagB- and TagF-catalyzed). R = glycerol or ribitol, m = 40–60.

(B) Cartoon ribbon representation of TagA$^{\Delta C}$ from *T. italicus*. Apo-TagA$^{\Delta C}$ crystallizes as a dimer. The dimer interface is over 1000 Å in surface area and is formed by buried hydrophobic residues. (C) TagA$^{\Delta C}$ protomer with secondary structural elements indicated. H = alpha helix, β = beta strand. (D) Electrostatic surface representation of the TagA$^{\Delta C}$ dimer. Negatively charged (*red*), neutral (*white*), and positively charged (*blue*) residues are indicated. Rotation of 180° about the dimer interface allows visualization of the pore.

Data collection	**SeMet ^{Ti}TagA$^{\Delta C}$**	**SeMet ^{Ti}TagA$^{\Delta C}$-UDP**
PDB Code	5WB4	5WFG
Space group	$P2_1$	$P2_1$
Cell dimensions		
a, *b*, *c* (Å)	77.0, 107.8, 89.4	64.9, 104.2, 90.1
α, β, γ (˚)	90.0, 98.3, 90.0	90.0, 108.3, 90.0
Resolution (Å)	88.51–2.00 (2.05–2.00)[b]	85.6–2.9 (3.0–2.9)[b]
Wavelength[a]	0.9791	
R_{merge} (%)	15.3 (61.8)	7.4 (64.6)
I / σI	7.2 (2.5)	7.2 (1.0)
$CC_{1/2}$	98.9 (84.3)	99.5 (59.4)
Completeness (%)	98.1 (92.8)	88.7 (79.1)
Redundancy	4.8 (4.5)	1.6 (1.6)
Wilson B-factor (Å^2)	21.7	69.7
Refinement		
Resolution (Å)	88.51–2.00	33.03–2.90
No. reflections	96634	24132
R_{work} / R_{free} (%)[c]	21.9/24.6	20.0/24.7
No. atoms	12043	8456
Protein	11591	8216
Ligand/ion	45	216
Water	407	24
B-factors (all atoms)	25.0	81.0
Protein	24.9	80.5
Ligand/ion	29.4	102.9
Water	25.9	43.3
R.m.s. deviations		
Bond lengths (Å)	0.010	0.010
Bond angles (˚)	1.04	1.14
Ramachandran favored (%)	89.7	90.4
Ramachandran allowed (%)	9.4	8.9
Ramachandran generally allowed (%)	0.9	0.4
Ramachandran outliers (%)	0.0	0.3

[a] Selenium peak.

[b] Values in parentheses are for highest-resolution shell.

[c] R_{free} calculated using 5% of the data.

https://doi.org/10.1371/journal.ppat.1007723.t001

Table 2.1 Crystal data collection and structure refinement statistics.

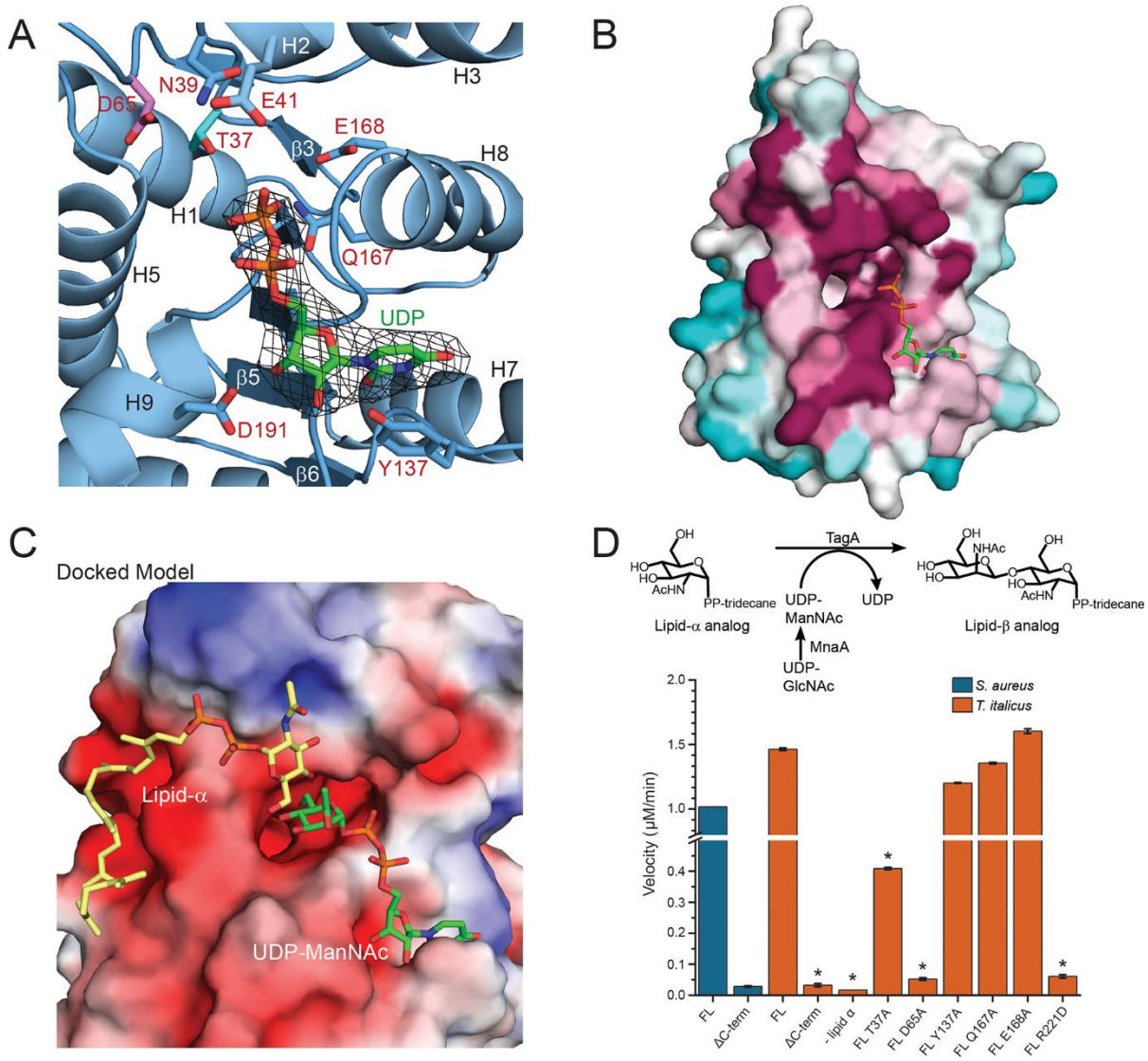

Figure 2.2 *In silico* substrate binding and mechanistic studies of TagA. (A) Structure of the UDP:TagA$^{\Delta C}$ complex. A simulated annealing omit map contoured at 3σ reveals that the β-phosphate of UDP (*orange*) in the UDP:TagA$^{\Delta C}$ structure is adjacent to the proposed catalytic base, Asp65 (*pink*), and putative ManNAc stabilizing residue, Thr37 (*cyan*). The uracil nucleoside is pi-stacked over Tyr137. (B) Consurf analysis reveals that UDP projects its phosphates (*orange*) into the pocket and is coordinated by a pocket of highly conserved residues. Highly conserved (*magenta*), moderately conserved (*white*) and weakly conserved (*teal*) residues are indicated. (C) *In silico* generated model of TagA$^{\Delta C}$ bound to its substrates. The model contains a lipid-α analog

(*yellow*) and UDP-ManNAc (*green*). The electrostatic surface of the protein is shown. Negatively charged residues (*red*), neutral residues (*white*), positively charged residues (*blue*) are shown. The coordinates of the model were generated using a two-step procedure. First, the coordinates of the ligand-bound UDP:TagA$^{\Delta C}$ crystal structure were used to restrain the positioning of UDP-ManNAc. Autodock vina was then used to dock lipid-α. The docking results positioned the non-reducing end of GlcNAc toward the C4 in ManNAc when bound to the TagA$^{\Delta C}$ dimer. (D) The upper panel indicates the reaction scheme used for the *in vitro* TagA activity assay. 200 nM TagA enzyme was incubated at 30°C with 100 μM lipid-α substrate analog and UDP-ManNAc produced *in situ* from UDP-GlcNAc by the epimerase MnaA, followed by quenching with 4M urea. Conversion of UDP-ManNAc to UDP is monitored at 271 nm using a DNAPak PA200 anion exchange column. The lower panel indicates activity measurements of *T. italicus* and *S. aureus* TagA enzymes from the *in vitro* TagA activity assay, as described above. Reactions with error bars were performed in triplicate, and asterisks indicate $p<0.005$ by Student's T-test.

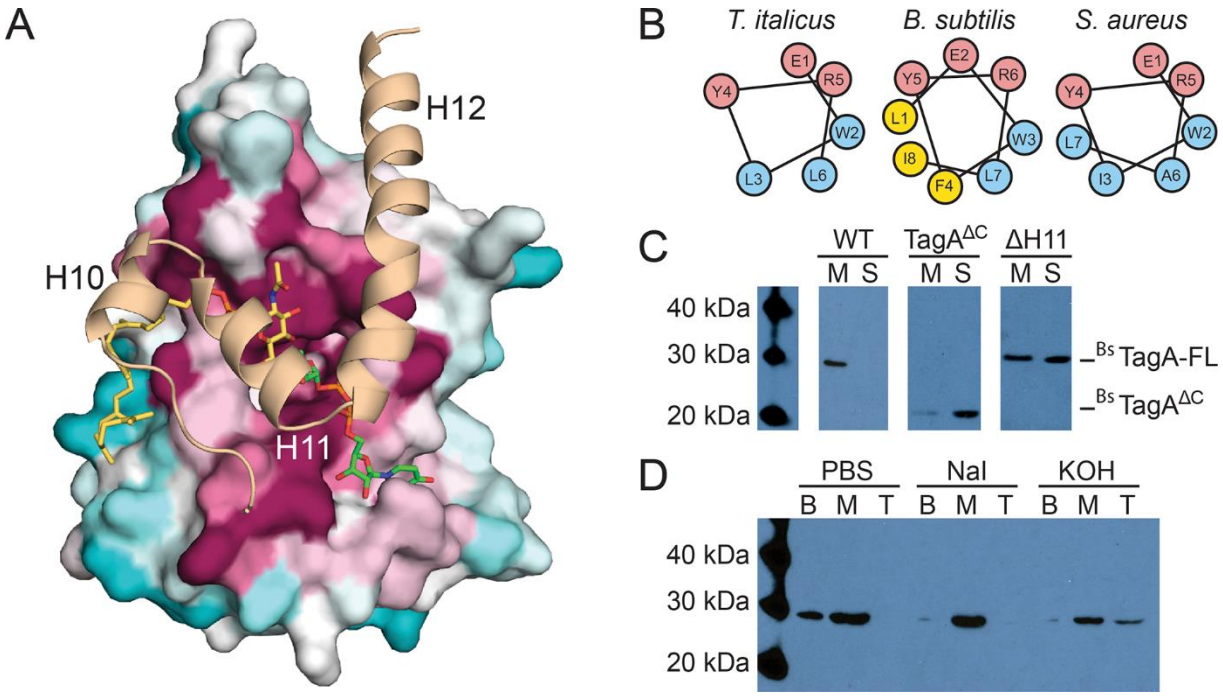

Figure 2.3 Computational and biochemical studies of the TagA enzyme inform cellular localization. (A) A model of full-length TagA was constructed with experimentally determined TagA$^{\Delta C}$ (*surface representation*) and the C-terminal domain, which was modeled by GREMLIN structural prediction (*cartoon representation*). Three C-terminal helices appear to complete the active site and obstruct the dimeric interface of TagA$^{\Delta C}$. Highly conserved (*magenta*), moderately conserved (*white*) and weakly conserved (*teal*) residues are indicated for the TagA$^{\Delta C}$ crystal structure. (B) Helical wheel projections of helix H11 in TagA homologs predict a putative amphipathic helix. (C) TagA associates with the bacterial cell membrane. Immunoblots of cellular fractionation indicate that *B. subtilis* TagA is exclusively localized to the membrane (*M*), while TagA$^{\Delta C}$ is primarily localized in the supernatant (*S*). Samples were fractionated by ultracentrifugation identically and the ^{Bs}TagA-FL blot was exposed for 10 minutes, the ^{Bs}TagA-V196 blot was exposed for 1 minute, and the ^{Bs}TagA-ΔH11 blot was exposed for 30 seconds. (D) TagA is a peripheral membrane protein. Chaotropic and alkaline treatments of *B. subtilis* TagA reveal that the enzyme is peripherally associated with the membrane and is more effectively

displaced by alkaline treatment. Treated membrane fractions were loaded onto a sucrose cushion, centrifuged, and carefully separated into bottom (*B; pellet*), middle (*M; sucrose cushion volume*), and top (*T; sample volume*).

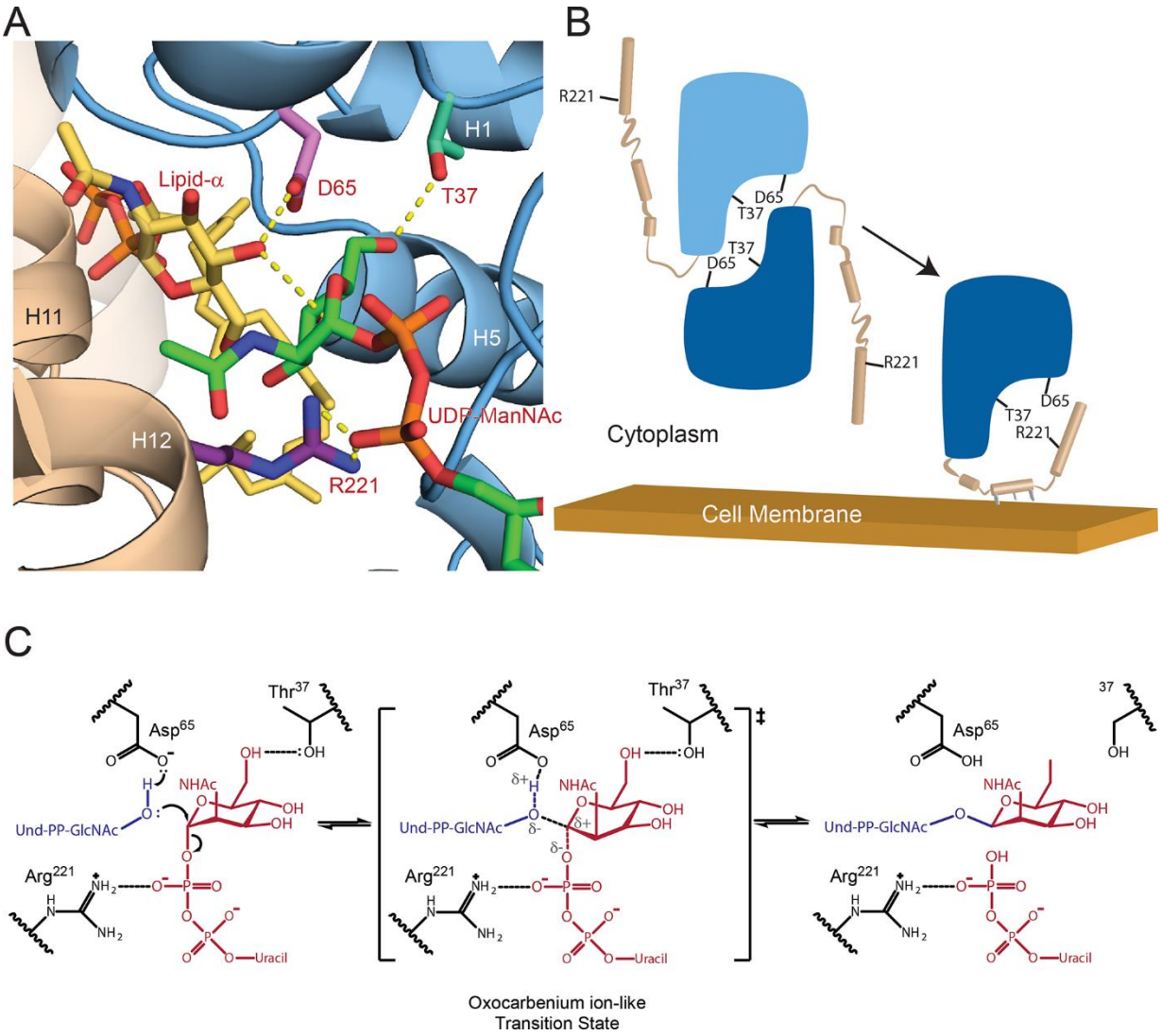

Figure 2.4 Model of the TagA enzyme-substrate complex and mechanism of catalysis. (A) The proposed active site of TagA co-localizes residues D65, T37, and R221. Lipid-α (*yellow*) is activated by the catalytic base D65 (*pink*), while ManNAc (*green*) is positioned by contacts between its C6 hydroxyl and T37 (*cyan*). The C-terminal helices (*tan*) are modeled according to GREMLIN structural predictions and place R221 (*purple*) adjacent to the phosphates of UDP-ManNAc (*orange*) to putatively stabilize the leaving group. (B) The TagA molecular mechanism is proposed to utilize a dimer to monomer transition to regulate glycosyltransferase activity. TagA is stabilized as a soluble dimer. Upon interaction with the cell membrane, the C-terminus adopts

an ordered state and disrupts the dimer interface, which produces a competent active site by co-localizing D65, T37, and R221 to coordinate the soluble UDP-ManNAc and membrane-bound lipid-α substrates. (C) TagA reveals the catalytic mechanism of the GT26 family. Asp65 activates lipid-α, which proceeds to attack UDP-ManNAc in an S_N2-like mechanism. Coordination between Arg221 and the phosphates of UDP stabilize the leaving group, permitting the oxocarbenium ion-like transition state. The mechanism is completed by glycosidic bond-formation between GlcNAc and ManNAc to form lipid-β.

2.7 Supporting Information

2.7.1 Supplemental Methods

2.7.1.1 Cloning, expression, and protein purification

The N-terminal domain of TagA from *T. italicus* (^{Ti}TagA$^{\Delta C}$, residues Met1-Gly195) or *S. aureus* (^{Sa}TagA$^{\Delta C}$, residues Ala10-Ala204) was expressed from a pMAPLe4 plasmid in *Escherichia coli* BL21(DE3) cells (Table S2.1). Standard methods were employed, with cultures grown in the presence of 50 µM kanamycin at 37°C until an OD_{600} of 0.6-0.8 was reached. Protein expression was initiated by adding isopropyl-β-D-1-thiogalactopyranoside (IPTG) to 1 mM, followed by overnight protein expression at 18°C. A four-liter cell culture was harvested by centrifugation at 7000 rpm in a Beckman JA-10 rotor, and the pellet was re-suspended in 40 mL of Buffer A (50 mM Tris, pH 7.5; 500 mM NaCl; 40 mM CHAPS) with 400 µL of protease inhibitor cocktail (Sigma) and 2 mM phenylmethanesulfonylfluoride (PMSF). The cells were then lysed using an EmulsiFlex high pressure homogenizer (Avestin). Cell lysates were fractionated by centrifugation at 15,000 rpm in a Beckman JA-20 rotor, and the soluble portion was applied to a gravity column containing 10 mL of suspended His-Pure Co^{2+} resin (Life Technologies) that was pre-equilibrated with Buffer A. The resin was washed with 20 mL aliquots of Buffer B (50 mM Tris, pH 7.5; 500 mM NaCl; 0.5% CHAPS) that contained 0, 25, or 50 mM imidazole. His-tagged TagA$^{\Delta C}$ was eluted using Buffer B with 500 mM imidazole, and the fractions were pooled and concentrated using a 10 kDa MWCO Amicon Ultra-15 centrifugal filter (Millipore). To remove His_6-tag from the protein, TEV protease was added to TagA$^{\Delta C}$, and the solution was dialyzed in a 3.5 kDa MWCO Slide-A-Lyzer dialysis cassette (ThermoFisher Scientific) against Buffer C (50

mM Tris, pH 7.5; 200 mM NaCl) at 4 °C overnight. TEV protease was then separated from $TagA^{\Delta C}$ by binding 10 mL of suspended His-Pure Co^{2+} resin (Life Technologies) that was pre-equilibrated with Buffer C; cleaved $TagA^{\Delta C}$, which lacked the His_6-tag, was eluted from the resin using Buffer C. Cleaved $TagA^{\Delta C}$ was further purified by gel filtration chromatography using a Sephacryl size-exclusion column (GE Healthcare Life Sciences) that was equilibrated with Buffer C. Purified $TagA^{\Delta C}$ was then pooled, concentrated to 55 mg/mL, and stored at 4°C. Analytical SEC experiments were performed using a Superdex 75 10/300 Increase column (GE) using Buffer D (50mM Tris, pH 7.5; 500 mM NaCl).

Selenomethionine (SeMet) labeled protein was prepared with cultures grown in M9 minimal media in the presence of kanamycin at 37°C until an OD_{600} of 0.6-0.8 was reached. Protein expression was then initiated by adding IPTG to 1 mM followed by overnight protein expression at 18°C. Protein purification was performed as described for native protein.

2.7.1.2 Structure determination

Recombinant $TagA^{\Delta C}$ at a concentration of 50 mg/mL in Buffer C was used for crystal screening. Screening was performed with the JCSG+ broad matrix suite (Molecular Dimensions) at room temperature in a sitting-drop vapor diffusion format (200 nL drop size). SeMet-labeled protein crystals grew over the course of three days in the presence of 200 mM lithium sulfate; 100 mM phosphate citrate, pH 4.2; and 20% PEG 1000. For X-ray data collection, $TagA^{\Delta C}$ crystals were cryoprotected using reservoir solution containing 35% glycerol. Diffraction datasets were collected at the Advanced Photon Source (APS) beamline 24-1D-C equipped with a Pilatus-6M detector. All data were collected at 100 K. Data were collected at the detector distance of 300 mm, with 0.25° oscillations, and at a 0.9791 Å wavelength. Multi-wavelength anomalous dispersion

(MAD) experiment was collected at peak (12663.0 eV), inflection (12660.3 eV), and high remote (12763.0 eV) energy wavelengths.

The TagA$^{\Delta C}$ crystals diffracted X-rays to 2.0 Å resolution. The XDS/XSCALE package was used to index, integrate and scale data in $P2_1$ space group [55]. The asymmetric unit of the crystal contained eight protein molecules, yielding a Matthews coefficient of 2.11 Å/Da and a 41.79% solvent content in the unit cell. The SHELX suite was used to locate the heavy atom substructure, which identified a total of 56 selenium atom sites [56]. The quality of the phases calculated with the peak, inflection, and high remote energy diffraction datasets were improved using SHARP and the wARP suite (Global Phasing Limited). The heavy atom parameters were refined with MLPhare using the CCP4i suite [57]. Density modification and non-crystallographic symmetry averaging was performed with the CCP4i suite to improve the quality of the electron density map. Automated model building was performed with BUCCANEER, followed by refinement with BUSTER [58]. Modeling of the additional electron density was confirmed using $2F_o$-F_c omit maps generated using BUSTER. Complete refinement and structure statistics are reported in Table 2.1 (5WB4).

A second crystal form was produced with recombinant TagA$^{\Delta C}$ in the presence of UDP and ManNAc ligands. TagA$^{\Delta C}$ at a concentration of 45 mg/mL in Buffer C with 10 mM UDP and 10 mM ManNAc was used for crystal screening with the JCSG+ broad matrix suite (Molecular Dimensions) as described above. TagA$^{\Delta C}$-ligand co-crystals grew over the course of two days in the presence of 200 mM calcium acetate; 100 mM sodium cacodylate, pH 6.5; and 40% PEG 300. A single wavelength diffraction dataset for a non-cryoprotected crystal was collected at the APS beamline 24-1D-C equipped with a Pilatus-6M detector as described above. The crystals diffracted X-rays to 2.9 Å resolution. The XDS/XSCALE package was used to index, integrate and scale

data in $P2_1$ space group [55]. The asymmetric unit of the crystal contained six protein molecules, yielding a Matthews coefficient of 2.22 Å/Da and a 44.66% solvent content in the crystal unit cell. The PHASER program in the CCP4i suite was used for molecular replacement, employing the coordinates of the apo-TagA$^{\Delta C}$ (5WB4) [59]. Molecular replacement yielded a single solution, which was refined in iterative runs using Buster software. Additional electron density resembling the UDP ligand was observed using $2F_o$-F_c omit maps generated by BUSTER [58]. Complete refinement and structure statistics are reported in Table 2.1 (5WFG).

TagA carboxyl terminal structure was modeled using Generative Regularized Models of Proteins (GREMLIN), accessed online at http://gremlin.bakerlab.org. TagA structure could not be modeled as a dimer in its full length (personal communications). Chain A of TagA$^{\Delta C}$ and TagA^{GM} were aligned with an RMSD of 2.2 Å.

2.7.1.3 Oligomeric analysis

The dissociation constant and oligomeric determination of apo- and UDP-bound TagA$^{\Delta C}$ were determined by equilibrium sedimentation analytical ultracentrifugation (AUC) on an Optima XL-A analytical ultracentrifuge (Beckman Coulter, Brea, CA). Data regression analysis was performed using the Beckman-Coulter Optima Analytical Ultracentrifuge Origin Data Analysis Package. The data were fit to multi-exponential models and represented best by a monomer-dimer equilibrium that was calculated using the predicted monomeric molecular weight of 21633 Da by the ExPASy ProtParam tool. The dissociation constant (K_d) was determined to be the inverse of $K_{a(conc)}$ using equation 1 [60]:

$$K_{a(conc)} = K_{a(abs)} \frac{(\varepsilon l)^{n-1}}{n} \quad (1)$$

The molar extinction coefficient (ε) was determined to be 12950 $cm^{-1}M^{-1}$ using the ExPASy ProtParam tool, where l is the path length of 1.2 cm, n is the order of oligomerization, and $K_{a(abs)}$ is the absorbance association constant, as determined by the nonlinear regression of the monomer-dimer multiexponential model using the analytical software program mentioned above.

Chemical crosslinking experiments with disuccinyl suberate (DSS) (ThermoFisher Scientific) was performed according to manufacturer guidelines, with the following variations. RC21tagAFL and RC21tagAG195 overnight cultures were diluted 1:100 into LB, grown at 37°C and induced with 1mM IPTG at OD_{600} = 0.4. After two hours, cells were collected, washed three times with buffer (PBS, pH 8.0) and finally re-suspended at 3X concentration. Ten microliters of sample were loaded onto an SDS-PAGE gel, and western blotting was performed as described above.

2.7.1.4 *Bacillus subtilis* cloning

Antibiotic concentrations used in this study, unless otherwise indicated, were 100 µg/mL ampicillin, 1 µg/mL erythromycin and 100 µg/mL spectinomycin. The full-length *B. subtilis tagA* gene was amplified from purified genomic DNA from *B. subtilis* 168 (*Bacillus* genetic stock center) and sub-cloned into the pBL113 shuttle vector (Table S2.1) using *E. coli* XL10 (New England BioLab) to create pHisTagAFL. The hexahistidine-tag was incorporated during Gibson assembly [61]. The truncated TagA construct (residues Met1-Val196) was constructed by amplifying the first 588 nucleotides of *tagA* and cloning into the pBL113 shuttle vector to create plasmid pHisTagAV196. *B. subtilis* was made competent as previously reported and transformed with 5-10 µL of pure plasmid to create strains RC168tagAFL and RC168tagAV196 [62]. Homologous double-crossover at the *thrC* locus was verified using tryptophan/threonine

auxotrophy and sequencing (Laragen Sequencing). The *tagA* gene was removed via allelic replacement using plasmid p*tagA::spec*, which was constructed by cloning 1 kb of DNA upstream and downstream of the *tagA* gene to flank the spectinomycin resistance cassette in the pIC56 plasmid; a portion of the spectinomycin cassette that was predicted to form a stem loop was removed to ensure that the native operon was not disrupted. Strain RC168tagAFL was made competent as previously described, transformed with an approximately 3 kb linear PCR product from p*tagA::spec*, and plated on LB agar plates with spectinomycin at 37°C to produce strain RC168tagAFLΔtagA. The *tagA* knockout was determined via colony PCR, and the DNA sequence was confirmed (Laragen Sequencing).

2.7.2 Supporting Figures

Strain	Genotype/Description	Source
RC168	*B. subtilis* Wild type *trpC2*	BGSC
RC168tagAFL	RC168, *thrC*::pBL113-^{His}*tagA*	Current study
RC168tagAFLΔtagA	RC168tagAFL, *thrC*::pBL113-^{His}*tagA tagA*::*spec*	
RC168tagAV196	RC168 *thrC*::pBL113-^{His}*TagA-V196*	Current study
RC168tagAΔH11	RC168 *thrC*::pBL113-^{His}*TagA*ΔH11	Current study
RC21tagAFL	*E. coli BL21(DE3)* cells with pMAPLe4 containing *T. italicus* TagA	Current study
RC21tagAG195	*E. coli BL21(DE3)* cells with pMAPLe4 containing *T. italicus* TagA (M1-G195)	Current study
RC21tagAΔSrf4	*E. coli BL21(DE3)* cells with pMAPLe4 containing *T. italicus* TagA and I203E, L209Q, L212K, I216E mutations	Current study

Plasmid	Genotype/Description	Source
pMAPLe4	*E. coli* plasmid that expresses proteins with N-terminal maltose binding protein fusion	[64]
pBL113	*B. subtilis – E. coli* shuttle vector derived from pRDC19 (3) that integrates into the thrC locus in the *B. subtilis* genome, Amp^R, Erm^R, IPTG inducible	[65]
pHisTagAFL	*pBL113-containing B. subtilis tagA* gene with a 5' hexa-his tag.	Current study
pHisTagAV196	*pBL113-containing B. subtilis tagA* gene with a 5' hexa-his tag truncated at G618 to produce a V196 translated product	Current study
RC21tagAΔSrf4	*E. coli BL21(DE3)* cells with pMAPLe4 containing *T. italicus* TagA and I203E, L209Q, L212K, I216E mutations	Current study
pIC156	*E. coli/B. subtilis* shuttle vector with spectinomycin^R cassette.	[65]
p*tagA*::spec	pIC156 with genomic 1kb DNA flanking *tagA* from *B. subtilis* 168 inserted 5' and 3' to spec^R cassette	Current study

Table S2.1 Strains and plasmids used in this study.

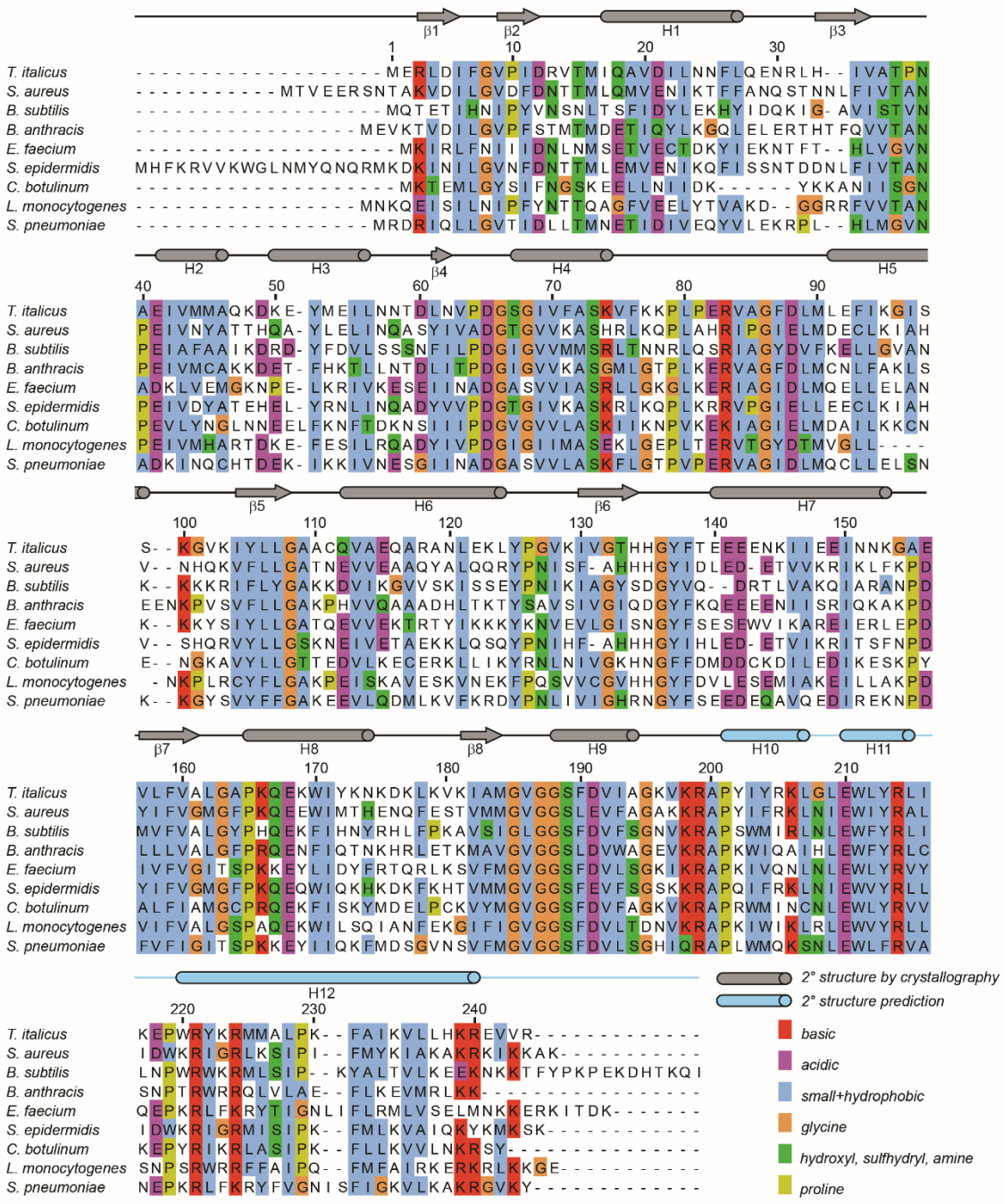

Figure S2.1 TagA primary sequence homology. The National Center for Biotechnology Institute's Basic Local Alignment Tool (BLAST) was used to determine TagA sequence homologs with high sequence identity. The Clustal Omega multiple sequence alignment tool was used to generate a sequence alignment [63]. Secondary structure is shown above the sequence and coloring is indicated in the key.

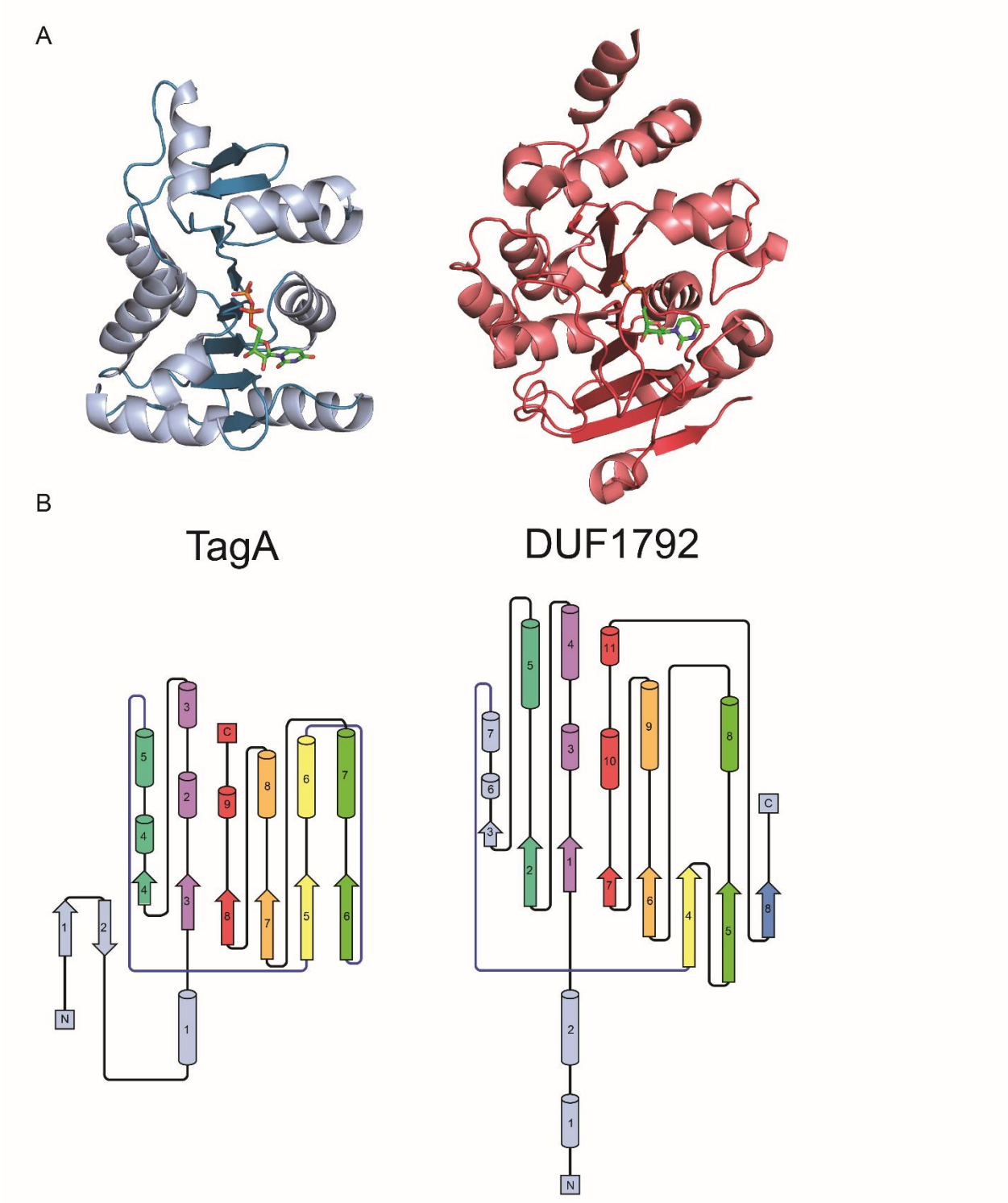

Figure S2.2 TagA structural comparison with DUF1792, a GT-D enzyme. (A) TagA aligns with DUF1792 with a DALI Z-score of 7.2 and an RMSD of 3.7 Å. Direct comparison reveals that

a β-sheet composed of parallel β-strands is the main component of structural similarity. The tertiary organization of secondary structural elements between TarA and DUF1792 is significantly different. (B) Cartoon representation of secondary structure topology highlights that TagA has fewer β-strands in its sheet than DUF1792 and reinforces the dissimilarity in the order of secondary structural elements.

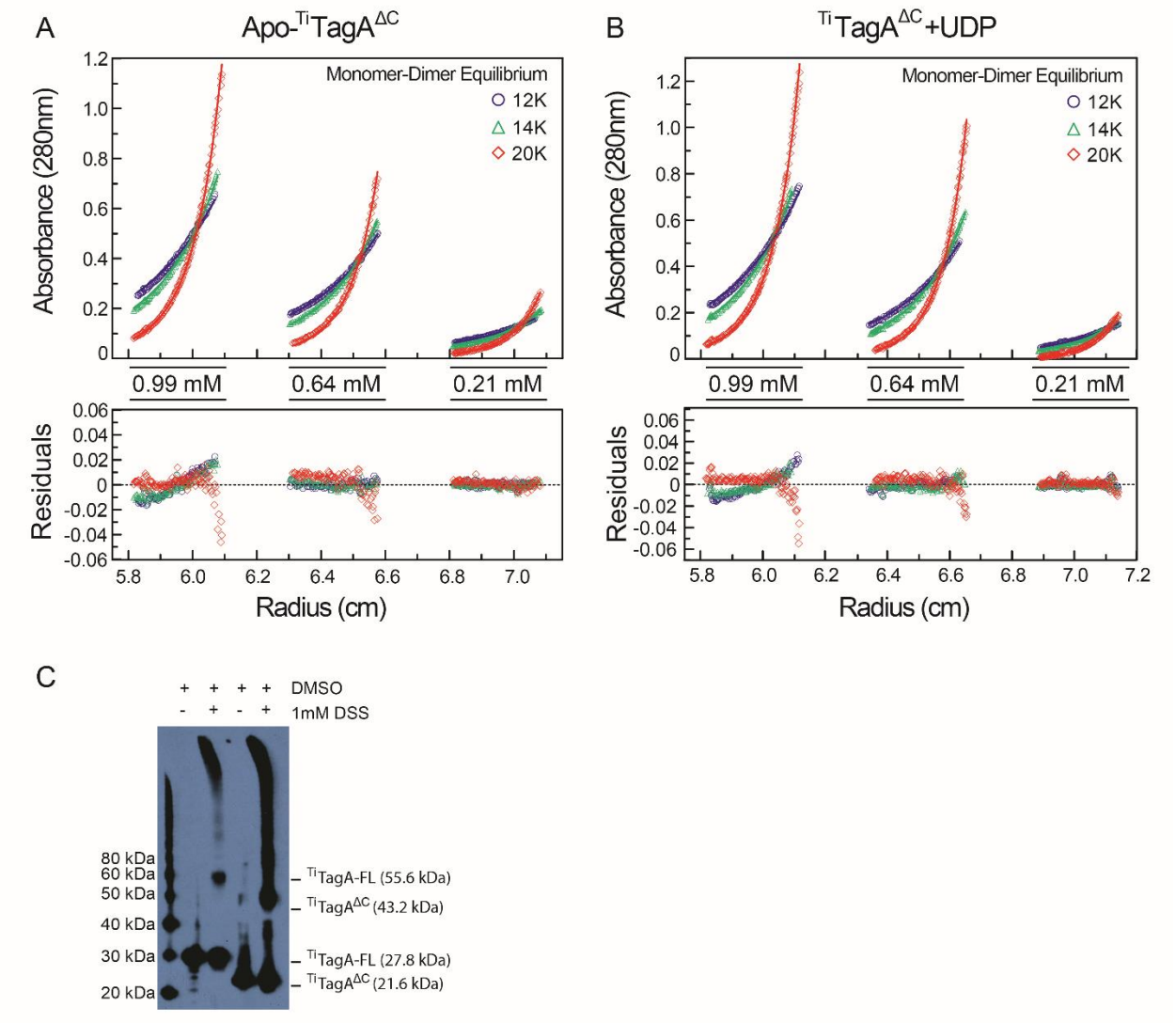

Figure S2.3 TagA oligomerization. The dissociation constant for TagA oligomerization was determined by equilibrium sedimentation analytical ultracentrifugation. The concentration distribution of TagA for three rotor speeds (12k, 14k, and 20k rpm) at three protein concentrations (0.99, 0.64, and 0.21 mM) for (A) apo-state TagA and (B) UDP-bound TagA. The lower panel shows the regression residuals for each protein concentration and centrifugal speed. The data were collected at 280 nm at 4°C and referenced against 50 mM Tris-HCl, pH 7.5, and 200 mM NaCl. (C) Crosslinking studies with disuccinimidyl suberate (DSS) in *E. coli* cells expressing *T. italicus* TagA constructs confirm that a dimer species is formed in the context of the cell. Both TagA and

TagA$^{\Delta C}$ are monomeric under denaturing conditions (+ DMSO,—DSS); however, the addition of 1 mM DSS (+ DMSO, + DSS) produced a band corresponding to dimer species.

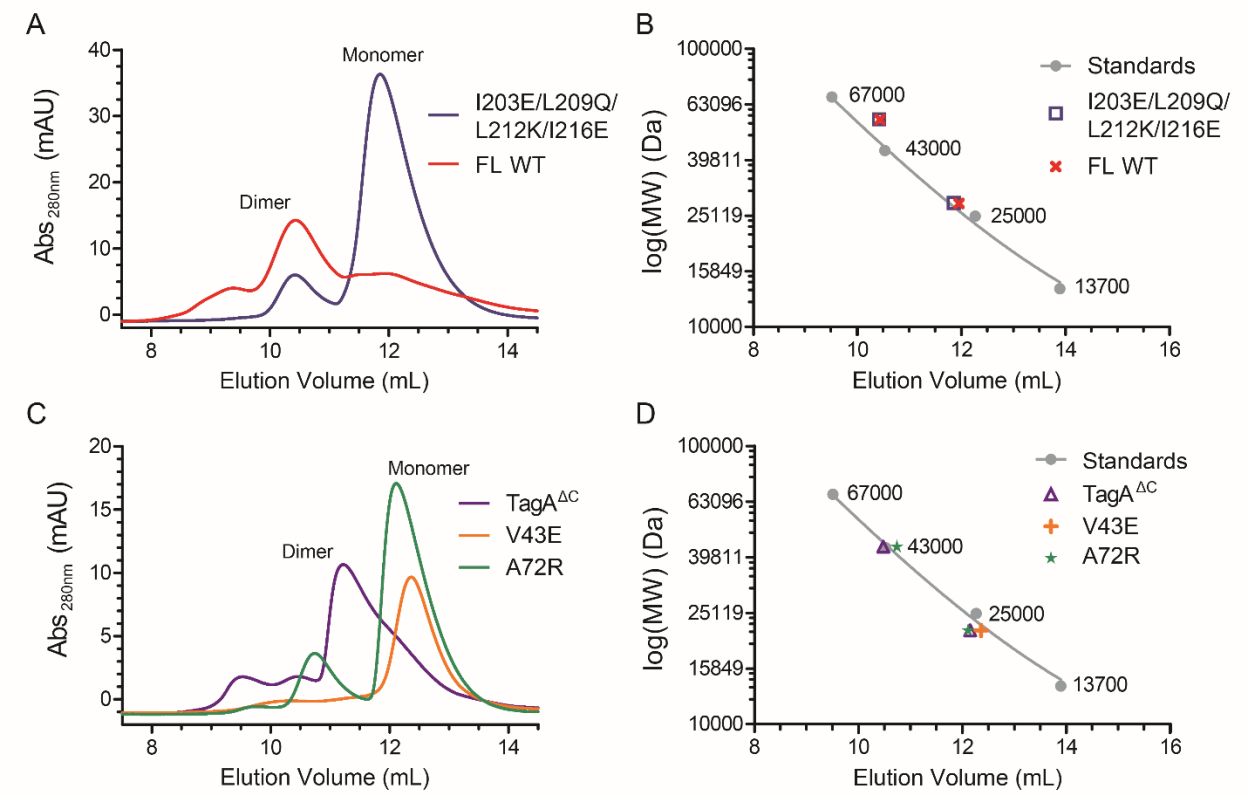

Figure S2.4 Size exclusion chromatography of full-length and C-terminal truncated TagA. (A) SEC chromatograms of *T. italicus* TagA^{FL} (FL WT; full length wild-type, *red*) and TagA^{FL} containing mutations in its C-terminal membrane-targeting appendage (I203E/L209Q/L212K/I216E mutant, *blue*). Based on the monomeric TagA^{GM} model, the four mutations increase the polarity of the hydrophobic surface formed by helices H10 and H11. These non-polar residues form a continuous patch in TagA^{GM} that we have shown to be important for membrane binding. The mutations stabilize formation of the monomer, presumably by reducing unfavorable entropic changes associated with solvating the hydrophobic surface of the native protein. (B) SEC calibration curve used to assign the oligomeric states of WT and mutant forms of TagA. A plot of the log of the molecular weight versus elution volume is shown. Elution position of molecular weight standards are shown in grey and were obtained in a separate experiment. (C) SEC chromatograms of TagA$^{\Delta C}$ and mutant forms of TagA$^{\Delta C}$: V43E (*orange*) and A72R (*green*).

Based on the crystal structure of apo- TagA$^{\Delta C}$, these mutations are at the interface and impede dimerization, which we confirm here via SEC. (D) Identical to (B) repeated for consistency.

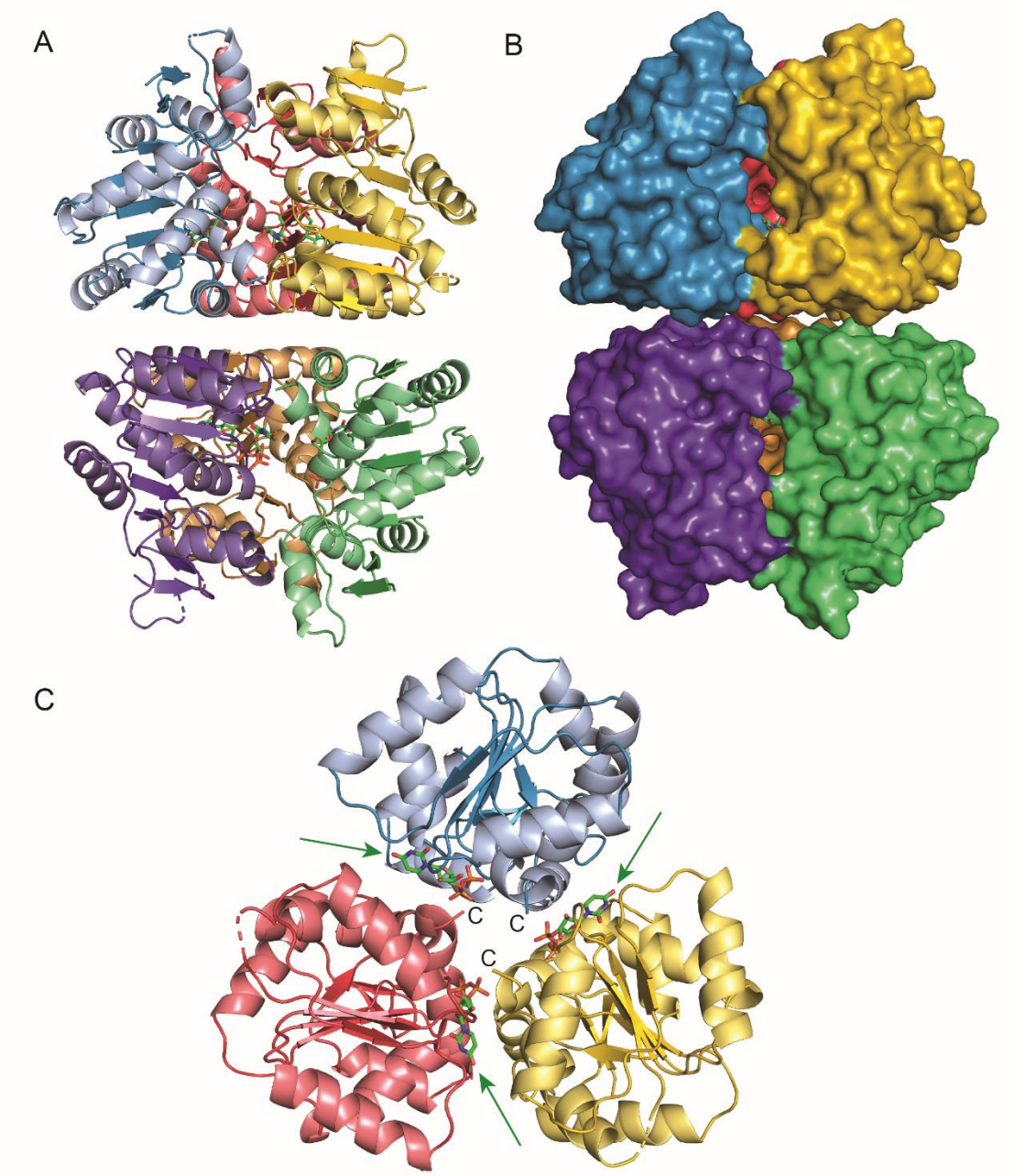

Figure S2.5 TagA$^{\Delta C}$:UDP complex crystallizes as a dimer of trimers. (A) Cartoon representation of the TagA$^{\Delta C}$:UDP complex. (B) Surface representation of the TagA$^{\Delta C}$:UDP complex. (C) View of one trimer unit within the crystallographic dimer. The C-termini are projected inward toward the center of symmetry. Green arrows indicate UDP, which can be seen at the interface between protomers of the trimer.

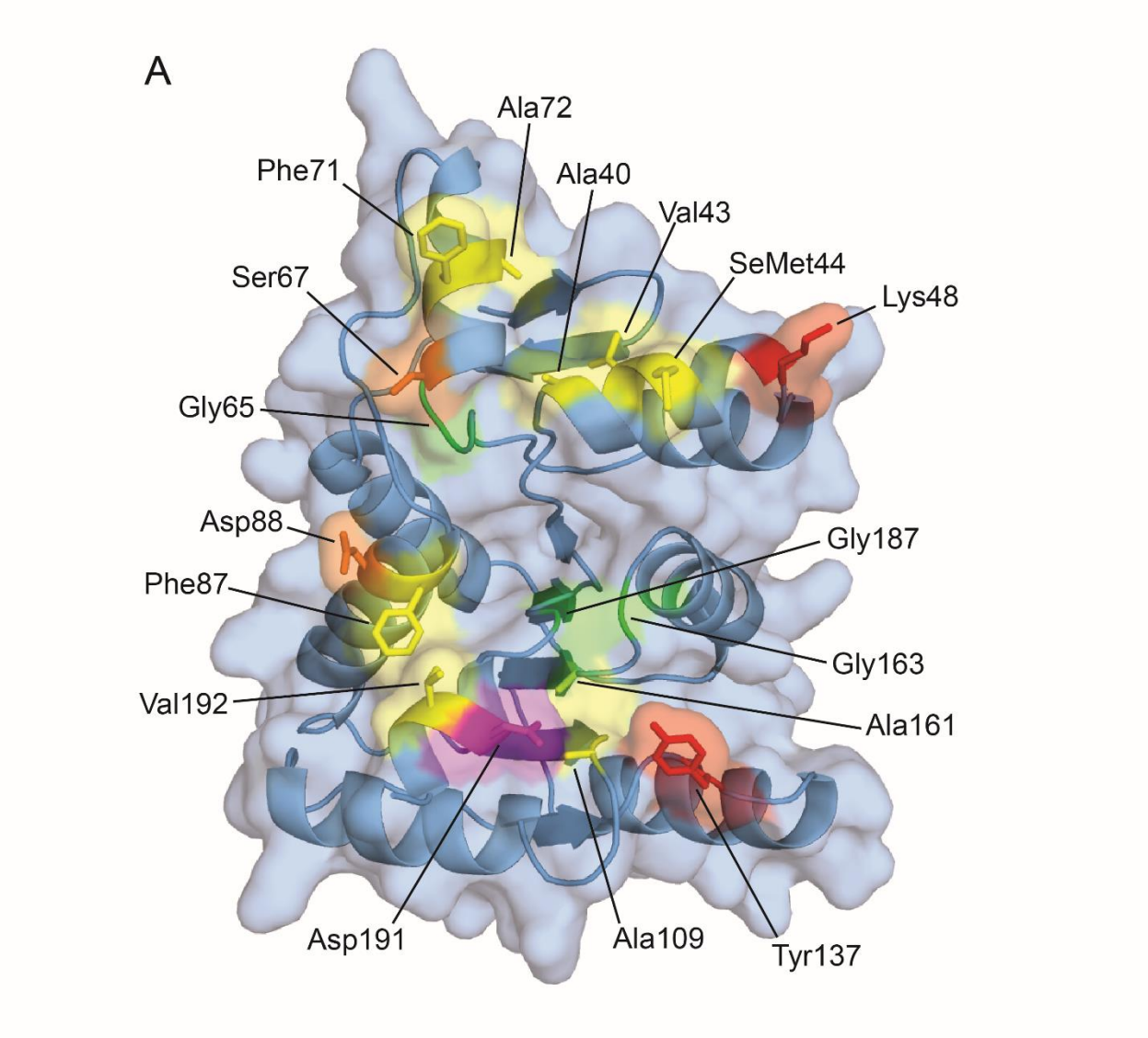

Figure S2.6 Buried residues at the TagA$^{\Delta C}$ dimeric interface. An EPPIC analysis of PDB 5WB4 identified fourteen residues with a buried surface area ≥ 75% in either protomer. An additional three residues engaged in polar bonds with lower percent buried surface area are shown. The side chains of these residues are shown and color-coded as follows: hydrophobic (*yellow*, Ala40, Val43, SeMet44, Phe71, Ala72, Phe87, Ala109, Ala161 and Val192), glycines (*green*, Gly65, Gly163 and Gly187), hydrogen bonds (*orange*, Ser67 and Asp88), polar (magenta, Asp191), and cation-pi stacking interactions (*red*, Lys48 and Tyr137). Mutation of residues Val43 and Ala72 were shown to disrupt dimerization in Fig. S2.4C.

2.8 References

1. Formstone A, Carballido-Lopez R, Noirot P, Errington J, Scheffers DJ. Localization and Interactions of Teichoic Acid Synthetic Enzymes in Bacillus subtilis. Journal of Bacteriology. 2008;190(5):1812–21. pmid:18156271
2. Brown S, Xia G, Luhachack LG, Campbell J, Meredith TC, Chen C, et al. Methicillin resistance in Staphylococcus aureus requires glycosylated wall teichoic acids. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(46):18909–14. pmid:23027967
3. Farha MA, Leung A, Sewell EW, D'Elia MA, Allison SE, Ejim L, et al. Inhibition of WTA synthesis blocks the cooperative action of pbps and sensitizes MRSA to β-lactams. ACS Chemical Biology. 2013;8(1):226–33. pmid:23062620
4. Heckels BJE, Lambert PA, Baddiley J. Binding of Magnesium Ions to Cell Walls of Bacillus subtilis W23 containing Teichoic Acid or Teichuronic Acid. 1977;162:359–65.
5. Holland LM, Conlon B, Gara JPO. Mutation of tagO reveals an essential role for wall teichoic acids in Staphylococcus epidermidis biofilm development. 2017;(2011):408–18.
6. Lambert BPA, Hancock IANC, Baddiley J. The Interaction of Magnesium Ions with Teichoic Acid. 1975:519–24. pmid:994
7. Misawa Y, Kelley KA, Wang X, Wang L, Park WB. Staphylococcus aureus Colonization of the Mouse Gastrointestinal Tract Is Modulated by Wall Teichoic Acid, Capsule, and Surface Proteins. 2015:1–21. pmid:26201029
8. Wanner S, Schade J, Keinhörster D, Weller N, George SE, Kull L, et al. Wall teichoic acids mediate increased virulence in Staphylococcus aureus. Nature Microbiology. 2017;2:16257–. pmid:28112716

9. Weidenmaier C, Kokai-Kun JF, Kristian Sa, Chanturiya T, Kalbacher H, Gross M, et al. Role of teichoic acids in Staphylococcus aureus nasal colonization, a major risk factor in nosocomial infections. Nature medicine. 2004;10(3):243–5. pmid:14758355

10. Yokoyama K, Miyashita T, Araki Y, Ito E. Structure and functions of linkage unit intermediates in the biosynthesis of ribitol teichoic acids in Staphylococcus aureus H and Bacillus subtilis W23. European journal of biochemistry. 1986;161(2):479–89. pmid:3096735

11. Brown S, Santa Maria JP, Walker S. Wall Teichoic Acids of Gram-Positive Bacteria. Annual Review of Microbiology. 2013;67(1):313–36. pmid:24024634

12. Soldo B, Karamata D, Lazarevic V. tagO is involved in the synthesis of all anionic cell-wall polymers in Bacillus subtilis 168. Microbiology. 2002;148(7):2079–87. pmid:12101296

13. Ginsberg C, Zhang Y-H, Yuan Y, Walker S. In vitro Reconstitution of Two Essential Steps in Wall Teichoic Acid Biosynthesis. ACS Chemical Biology. 2006;1(1):25–8. pmid:17163636

14. Zhang Y-H, Ginsberg C, Yuan Y, Walker S. Acceptor substrate selectivity and kinetic mechanism of Bacillus subtilis TagA. Biochemistry. 2006;45(36):10895–904. pmid:16953575

15. Swoboda JG, Campbell J, Meredith TC, Walker S. Wall teichoic acid function, biosynthesis, and inhibition. Chembiochem: a European journal of chemical biology. 2010;11(1):35–45. pmid:19899094

16. Lovering AL, Lin LYC, Sewell EW, Spreter T, Brown ED, Strynadka NCJ. Structure of the bacterial teichoic acid polymerase TagF provides insights into membrane association and catalysis. Nature structural & molecular biology. 2010;17(5):582–9. pmid:20400947

17. Schertzer JW, Brown ED. Purified, Recombinant TagF Protein from Bacillus subtilis 168 Catalyzes the Polymerization of Glycerol Phosphate onto a Membrane Acceptor in Vitro. Journal of Biological Chemistry. 2003;278(20):18002–7. pmid:12637499

18. Schirner K, Stone LK, Walker S. ABC Transporters Required for Export of Wall Teichoic Acids Do Not Discriminate between Different Main Chain Polymers. ACS Chemical Biology. 2011;6(5):407–12. pmid:21280653

19. Chan YG, Frankel MB, Dengler V, Schneewind O, Missiakas D. Staphylococcus aureus mutants lacking the LytR-CpsA-Psr family of enzymes release cell wall teichoic acids into the extracellular medium. J Bacteriol. 2013;195(20):4650–9. pmid:23935043; PubMed Central PMCID: PMCPMC3807444.

20. Gale RT, Li FKK, Sun T, Strynadka NCJ, Brown ED. B. subtilis LytR-CpsA-Psr Enzymes Transfer Wall Teichoic Acids from Authentic Lipid-Linked Substrates to Mature Peptidoglycan In Vitro. Cell Chemical Biology. 2017;24(12):1537–46.e4. pmid:29107701

21. Schaefer K, Matano LM, Qiao Y, Kahne D, Walker S. In vitro reconstitution demonstrates the cell wall ligase activity of LCP proteins. Nature Chemical Biology. 2017;13(4):396–401. pmid:28166208

22. Lovering AL, Lin LYC, Sewell EW, Spreter T, Brown ED, Strynadka NCJ. Structure of the bacterial teichoic acid polymerase TagF provides insights into membrane association and catalysis. Nature Structural &Amp; Molecular Biology. 2010;17:582. pmid:20400947

23. Sobhanifar S, Worrall LJ, King DT, Wasney GA, Baumann L, Gale RT, et al. Structure and Mechanism of Staphylococcus aureus TarS, the Wall Teichoic Acid β-glycosyltransferase Involved in Methicillin Resistance. PLOS Pathogens. 2016;12(12):e1006067. pmid:27973583

24. Sobhanifar S, Worrall LJ, Gruninger RJ, Wasney GA, Blaukopf M, Baumann L, et al. Structure and mechanism of Staphylococcus aureus TarM, the wall teichoic acid α-glycosyltransferase. Proceedings of the National Academy of Sciences. 2015;112(6):E576–E85. pmid:25624472

25. D'Elia MA, Henderson JA, Beveridge TJ, Heinrichs DE, Brown ED. The N-acetylmannosamine transferase catalyzes the first committed step of teichoic acid assembly in Bacillus subtilis and Staphylococcus aureus. Journal of Bacteriology. 2009;191(12):4030–4. pmid:19376878

26. Farha MA, Czarny TL, Myers CL, Worrall LJ, French S, Conrady DG, et al. Antagonism screen for inhibitors of bacterial cell wall biogenesis uncovers an inhibitor of undecaprenyl diphosphate synthase. Proceedings of the National Academy of Sciences. 2015;112(35):11048.

27. D'Elia MA, Millar KE, Beveridge TJ, Brown ED. Wall teichoic acid polymers are dispensable for cell viability in Bacillus subtilis. Journal of Bacteriology. 2006;188(23):8313–6. pmid:17012386

28. D'Elia MA, Pereira MP, Chung YS, Zhao W, Chau A, Kenney TJ, et al. Lesions in teichoic acid biosynthesis in Staphylococcus aureus lead to a lethal gain of function in the otherwise dispensable pathway. Journal of Bacteriology. 2006;188(12):4183–9. pmid:16740924

29. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Research. 2014;42(D1):D490–D5. pmid:24270786

30. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Research. 2016;44(D1):D279–D85. pmid:26673716

31. Haft RF, Wessels MR, Mebane MF, Conaty N, Rubens CE. Characterization of cpsF and its product CMP-N-acetylneuraminic acid synthetase, a group B streptococcal enzyme that can function in K1 capsular polysaccharide biosynthesis in Escherichia coli. Molecular microbiology. 1996;19(3):555–63. pmid:8830246

32. Liu J, Mushegian A. Three monophyletic superfamilies account for the majority of the known glycosyltransferases. Protein Science. 2003;12(7):1418–31. pmid:12824488

33. Rahman A, Barr K, Rick PD. Identification of the Structural Gene for the TDP-Fuc4NAc:Lipid II Fuc4NAc Transferase Involved in Synthesis of Enterobacterial Common Antigen in Escherichia coli K-12. Journal of Bacteriology. 2001;183(22):6509–16. pmid:11673418

34. Katzen F, Ferreiro DU, Oddo CG, Ielmini MV, Becker A, Pühler A, et al. Xanthomonas campestris pv. campestris Mutants: Effects on Xanthan Biosynthesis and Plant Virulence. Journal of Bacteriology. 1998;180(7):1607. pmid:9537354

35. Breton C, Šnajdrová L, Jeanneau C, Koča J, Imberty A. Structures and mechanisms of glycosyltransferases. Glycobiology. 2006;16(2):29R–37R. pmid:16037492

36. Lairson LL, Henrissat B, Davies GJ, Withers SG. Glycosyltransferases: Structures, Functions, and Mechanisms. Annual Review of Biochemistry. 2008;77(1):521–55. pmid:18518825

37. Bliven S, Lafita A, Parker A, Capitani G, Duarte JM. Automated evaluation of quaternary structures from protein crystals. PLOS Computational Biology. 2018;14(4):e1006104. pmid:29708963

38. Zhang H, Zhu F, Yang T, Ding L, Zhou M, Li J, et al. The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold. Nature Communications. 2014;5:4339-. pmid:25023666

39. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31(2):455–61. pmid:19499576; PubMed Central PMCID: PMCPMC3041641.

40. Gale RT, Sewell EW, Garrett TA, Brown ED. Reconstituting poly(glycerol phosphate) wall teichoic acid biosynthesis in vitro using authentic substrates. Chemical Science. 2014;5(10):3823–.

41. Balakrishnan S, Kamisetty H, Carbonell JG, Lee SI, Langmead CJ. Learning generative models for protein fold families. Proteins. 2011;79(4):1061–78. pmid:21268112.

42. Kamisetty H, Ovchinnikov S, Baker D. Assessing the utility of coevolution-based residue–residue contact predictions in a sequence- and structure-rich era. Proceedings of the National Academy of Sciences. 2013;110(39):15674–9. pmid:24009338

43. Pak JE, Arnoux P, Zhou S, Sivarajah P, Satkunarajah M, Xing X, et al. X-ray crystal structure of leukocyte type core 2 beta1,6-N-acetylglucosaminyltransferase. Evidence for

a convergence of metal ion-independent glycosyltransferase mechanism. J Biol Chem. 2006;281(36):26693–701. pmid:16829524.

44. Pichoff S, Lutkenhaus J. Tethering the Z ring to the membrane through a conserved membrane targeting sequence in FtsA. Molecular Microbiology. 2005;55(6):1722–34. pmid:15752196

45. Szeto TH, Rowland SL, Rothfield LI, King GF. Membrane localization of MinD is mediated by a C-terminal motif that is conserved across eubacteria, archaea, and chloroplasts. Proceedings of the National Academy of Sciences. 2002;99(24):15693–8. pmid:12424340

46. Hu Z, Lutkenhaus J. A conserved sequence at the C-terminus of MinD is required for binding to the membrane and targeting MinC to the septum. Molecular Microbiology. 2003;47(2):345–55. pmid:12519187

47. Street C. Multiple mechanisms of membrane anchoring of Escherichia coli penicillin-binding proteins. 1994;13:1–12. pmid:8117464

48. Bhavsar AP, D'Elia MA, Sahakian TD, Brown ED. The amino terminus of Bacillus subtilis TagB possesses separable localization and functional properties. Journal of Bacteriology. 2007;189(19):6816–23. pmid:17660278

49. Mouchlis VD, Bucher D, McCammon JA, Dennis EA. Membranes serve as allosteric activators of phospholipase A2, enabling it to extract, bind, and hydrolyze phospholipid substrates. Proceedings of the National Academy of Sciences. 2015;112(6):E516–E25. pmid:25624474

50. Celniker G, Nimrod G, Ashkenazy H, Glaser F, Martz E, Mayrose I, et al. ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. Israel Journal of Chemistry. 2013;53(3-4):199–206.

51. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Research. 2010;38(suppl_2):W529–W33. pmid:20478830

52. Berezin C, Glaser F, Rosenberg J, Paz I, Pupko T, Fariselli P, et al. ConSeq: the identification of functionally and structurally important residues in protein sequences. Bioinformatics. 2004;20(8):1322–4. pmid:14871869

53. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D. 2010;66(2):213–21. pmid:20124702

54. Bhavsar AP, Truant R, Brown ED. The TagB protein in Bacillus subtilis 168 is an intracellular peripheral membrane protein that can incorporate glycerol phosphate onto a membrane-bound acceptor in vitro. Journal of Biological Chemistry. 2005;280(44):36691–700. pmid:16150696

55. Kabsch W. XDS. Acta Crystallographica Section D. 2010;66(2):125-132.

56. Thorn A. Experimental Phasing: Substructure Solution and Density Modification as Implemented in SHELX. Protein Crystallography: Methods and Protocols, eds Wlodawer A, Dauter Z, & Jaskolski M (Springer New York, New York, NY), 2017;pp 357-376.

57. Winn MD, et al. Overview of the CCP4 suite and current developments. Acta Crystallographica Section D. 2011;67(4):235-242.

58. Bricogne G. Direct phase determination by entropy maximization and likelihood ranking: status report and perspectives. Acta Crystallogr D Biol Crystallogr. 1993;49(Pt 1):37-60.

59. McCoy AJ, et al. Phaser crystallographic software. Journal of Applied Crystallography. 2007;40(Pt 4):658-674.

60. McRorie DK & Voelker P. Self-associating systems in the analytical ultracentrifuge. Beckman Instruments, Fullerton, CA. 1993.

61. Gibson DG, et. al. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods. 2009;6:343.

62. Anagnostopoulos C & Spizizen J. REQUIREMENTS FOR TRANSFORMATION IN BACILLUS SUBTILIS. Journal of Bacteriology. 1961;81(5):741.

63. Sievers F, et. al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology. 2011;7(1).

64. Arbing MA, et. al. Heterologous Expression of Mycobacterial Esx Complexes in Escherichia coli for Structural Studies Is Facilitated by the Use of Maltose Binding Protein Fusions. PLOS ONE. 2013;8(11):e81753.

65. Steinmetz M & Richter R. Plasmids designed to alter the antibiotic resistance expressed by insertion mutations in Bacillus subtilis, through in vivo recombination. Gene. 1994;142(1):79-83.

Chapter 3

Insight into the molecular basis of substrate recognition by the TagA glycosyltransferase

3.1 Overview

Gram-positive bacteria synthesize, secrete, and anchor wall teichoic acid (WTA) glycopolymers to their cell wall where they perform several functions for the cell, including localizing cell elongation machinery, enabling biofilm formation, maintaining the cell wall architecture, and mediating β-lactam antibiotic resistance. Enzymes involved in the WTA biosynthetic pathway are exciting targets for novel antibiotic development, but the most promising target is the conserved linkage unit glycosyltransferase TagA. TagA catalyzes the first committed step in WTA synthesis and, when genetically deleted in *Staphylococcus aureus*, reduces virulence without compromising cell viability. Thus, TagA-specific inhibitors may reduce the selective pressure for the development of resistance mechanisms while also re-enabling the arsenal of antibiotics that have been rendered ineffective by the modern antibiotic resistance crisis.

In Chapter 2, *in vitro* biochemical data and *in vivo* cellular data suggested that the flexible C-terminal tail (CTT) of TagA was essential for catalysis and membrane association for the protein. However, the mechanisms underlying how it directly participated in catalysis and recognized the enzyme's substrates remained unclear. In this chapter, we engineered a solubility-enhanced construct of TagA ($TagA^{FL*}$) and determined co-crystal structures of $TagA^{FL*}$ with its native substrate, UDP-ManNAc, and an epimer of the substrate, UDP-GlcNAc. A comparison of the two structures revealed clear stereospecific interactions that TagA makes with its native sugar donor substrate and not its epimer. In addition, we performed molecular dynamics simulations on modeled full-length TagA with and without the UDP-ManNAc. These simulations revealed that dynamic motions of the CTT became rigid due to salt bridges between conserved Arg residues in the CTT and the diphosphates of the ligand. Indeed, these Arg residues are essential for catalysis and likely play a role in stabilizing the CTT on the core domain of TagA. Finally, native mass

spectrometry data of native TagA with increasing amounts of detergent confirmed that the monomeric form of the protein binds to membranes, which supports our model that the soluble dimer breaks due to favorable protein-membrane interactions from the CTT. The data here support a model where the membrane-associated and enzymatically active form of TagA is mediated by the CTT. On the membrane, the motions between the core domain and CTT are quenched by UDP-ManNAc binding and the enzyme is primed to bind to its second substrate, lipid-α. My contributions to this work included engineering $TagA^{FL*}$, determining $TagA^{FL*}$ protein-ligand structures, characterizing biochemical properties of $TagA^{FL*}$, analyzing the computationally generated models of full-length TagA, performing molecular dynamics simulations, interpreting data, generating figures, and preparing the resulting manuscript.

This chapter is reformatted with permission from a peer-reviewed editors' pick research article in the Journal of Biological Chemistry: "Insight into the molecular basis of substrate recognition by the wall teichoic acid glycosyltransferase TagA." Martinez, O.E., Mahoney, B.J., Goring, A.K., Yi, S., Tran, D.P., Cascio, D., Phillips, M.L., Muthana, M.M., Chen, X., Jung, M.E., Loo, J.A., Clubb, R.T., *J Biol Chem*. **298**(2), 101464 (2021).

3.2 Insight into the molecular basis of substrate recognition by the wall teichoic acid glycosyltransferase TagA

3.2.1 Abstract

Wall teichoic acid (WTA) polymers are covalently affixed to the Gram-positive bacterial cell wall and have important functions in cell elongation, cell morphology, biofilm formation, and β-lactam antibiotic resistance. The first committed step in WTA biosynthesis is catalyzed by the TagA glycosyltransferase (also called TarA), a peripheral membrane protein that produces the conserved linkage unit, which joins WTA to the cell wall peptidoglycan. TagA contains a conserved GT26 core domain followed by a C-terminal polypeptide tail that is important for catalysis and membrane binding. Here, we report the crystal structure of the *Thermoanaerobacter italicus* TagA enzyme bound to UDP-*N*-acetyl-D-mannosamine, revealing the molecular basis of substrate binding. Native MS experiments support the model that only monomeric TagA is enzymatically active and that it is stabilized by membrane binding. Molecular dynamics simulations and enzyme activity measurements indicate that the C-terminal polypeptide tail facilitates catalysis by encapsulating the UDP-*N*-acetyl-D-mannosamine substrate, presenting three highly conserved arginine residues to the active site that are important for catalysis (R214, R221, and R224). From these data, we present a mechanistic model of catalysis that ascribes functions for these residues. This work could facilitate the development of new antimicrobial compounds that disrupt WTA biosynthesis in pathogenic bacteria.

3.3 Introduction

Staphylococcus aureus and other Gram-positive bacteria are surrounded by a thick murein sacculus that is densely functionalized with wall teichoic acid (WTA) polymers (1, 2, 3, 4, 5). These anionic glycopolymers are essential components of the bacterial cell and involved in cell elongation, morphogenesis, cation homeostasis, pathogenesis, and autolysin localization (6, 7, 8, 9, 10, 11, 12, 13). They are polymers of alditol phosphate repeating units that are covalently joined to the surface peptidoglycan *via* a conserved linkage unit that is composed of 1 to 3 glycerol-3-phosphate (GroP) groups appended to an *N*-acetyl-D-mannosamine (ManNAc) (β1 → 4) *N*-acetyl-D-glucosamine (GlcNAc) disaccharide monophosphate (2, 14, 15, 16). The WTA biosynthetic pathway has drawn considerable interest as a drug target, as genetically eliminating WTA production in clinically important methicillin-resistant *S. aureus* resensitizes it to β-lactam antibiotics and attenuates its virulence (6, 7).

WTA polymers are synthesized on the cytoplasmic face of the cell membrane by enzymes that sequentially elaborate a membrane-embedded undecaprenyl phosphate carrier molecule (2, 3, 5). In *Bacillus subtilis* and *S. aureus*, the linkage unit is synthesized by the sequential action of the TagO, TagA, and TagB enzymes (originally designated as Tar enzymes in *S. aureus*). WTA synthesis is initiated by TagO, which catalyzes the reversible transfer of GlcNAc-1-P from UDP–GlcNAc to the undecaprenyl (Und) phosphate scaffold to produce lipid-α (GlcNAcα–PP-Und) (17). The TagA ManNAc transferase then appends ManNAc from a sugar nucleotide donor, UDP–ManNAc, producing a ManNAc (β1 → 4) GlcNAcα–PP-Und product, called lipid-β (18, 19, 20). The linkage unit synthesis is completed by TagB, which adds a GroP to lipid-β using a CDP–glycerol donor substrate (18, 21). Polymerase enzymes then extend the (GroP)–ManNAc (β1 → 4) GlcNAcα-1-P linkage unit from the terminal GroP at the nonreducing end to construct the body

of the polymer with either ribitol- or glycerol-phosphate repeating units, which can vary substantially among species (22, 23). The mature WTA is modified with monosaccharides at free hydroxyl groups on the main chain polymer (24). WTAs are then transported across the membrane to the cell surface by the TagGH transporter (2, 25). Finally, in the cell wall, the GlcNAcα-1-P at the reducing end of the linkage unit is covalently attached to the 6′ hydroxyl of the *N*-acetylmuramic acid component in the peptidoglycan *via* a phosphodiester bond, and in some instances, further tailored with d-alanine modifications (2, 15, 26, 27).

The TagA enzyme is a promising drug target because it catalyzes the first committed step in WTA biosynthesis and methicillin-resistant *S. aureus* strains in which it is genetically deleted are attenuated in virulence and resensitized to β-lactam antibiotics (7, 28). Biochemical studies have shown that *B. subtilis* TagA is a metal-independent inverting glycosyltransferase (GT) that catalyzes the transfer of ManNAc from UDP–ManNAc to the 4′ hydroxyl of GlcNAc in lipid-α (19). Catalysis occurs *via* a Bi–Bi mechanism in which the UDP–ManNAc donor first binds to the enzyme, and the UDP product is released last (19). *S. aureus* TagA is part of the WecB/TagA/CpsF GT family (PFAM03808), whose >6000 members catalyze the synthesis of WTAs and other important surface glycopolymers such as capsular polysaccharides in group B *Streptococcus* and the enterobacterial common antigen in *Escherichia coli* (29, 30, 31). Members of this large family are classified as GT26 enzymes in the Carbohydrate-Active enZYmes database (www.cazy.org), but little is known about the molecular basis of their enzymatic activity (32, 33, 34, 35). Recently, we reported the structure of the TagA enzyme from *Thermoanaerobacter italicus*, which represents the only known structure of a GT26 enzyme (36, 37). The GT26 domain adopts a unique fold that is distinct from other GTs (*e.g.*, GT-A, GT-B, GT-C, and GT-D folds) and is followed by a 49-amino acid C-terminal tail (CTT) that targets the enzyme to the cell membrane (36).

Functional studies showed that the CTT is essential for activity and identified D65 and R221 as important determinants for catalysis as their alteration causes more than a 20-fold reduction in activity as compared with the wildtype enzyme (36). However, the molecular basis of substrate recognition and catalysis remains incompletely understood.

In this study, we used a combination of biochemical, structural, and computational approaches to determine how TagA recognizes its substrates and how flexibility in the CTT mediates substrate binding and catalysis. Using a solubility-enhanced enzyme variant, we determined the structure of TagA bound to its natural substrate, UDP–ManNAc. Computational modeling, molecular dynamics (MD) simulations, and biochemical experiments of the full-length enzyme provide insight into the role of the CTT in catalysis, revealing that it presents conserved amino-acid side chains to contact the UDP–ManNAc substrate. Paired with native MS of TagA in the presence of micelles, we provide additional support for a membrane-induced activation mechanism of the enzyme. Bilayer association *via* the CTT appears to stabilize the monomeric form of the enzyme so that it can glycosylate its membrane-embedded glycophospholipid substrate. These results could provide a foundation for the discovery of new antibiotics that work by inhibiting WTA biosynthesis.

3.4 Results

3.4.1 Solution-state studies of TagA

In previously reported crystal structures, *T. italicus* $TagA^{\Delta C}$ (TagA, residues M1-G195) adopted both dimeric and trimeric oligomeric states (36). To probe its oligomerization state in solution, we used NMR. $TagA^{\Delta C}$ contains the core domain but is missing 49 residues at its C

terminus that have been shown to promote membrane interactions and facilitate catalysis (Fig. 3.1*A*) (36). A 0.6 mM sample of [U-^{2}H (70%), ^{13}C, ^{15}N] TagA$^{\Delta C}$ was overexpressed and purified, and the chemical shifts of its backbone atoms were assigned using triple resonance methods. The ^{1}H–^{15}N transverse relaxation optimized spectroscopy (TROSY)–heteronuclear single quantum coherence (HSQC) spectrum of TagA$^{\Delta C}$ is reasonably well resolved, enabling ~61% of the amino acids in its primary sequence to be specifically assigned (Fig. 3.2*A*). An analysis of the secondary chemical shifts reveals that the positioning of the secondary structural elements in the protein match those visualized in the crystal structure. Based on ^{15}N NMR relaxation measurements, TagA$^{\Delta C}$ has a molecular correlation time (τ_c) of 27 ns, corresponding to a dimer. This is consistent with previously reported sedimentation equilibrium–analytical ultracentrifugation (SE–AUC) experiments, which showed that apo-TagA$^{\Delta C}$ exists in a monomer–dimer equilibrium in solution (dissociation constant [K_D] = 7.4 ± 0.7 μM) (36). Interestingly, backbone amide signals for 81 of 195 (74 of 188 non-proline) residues in the primary sequence of TagA$^{\Delta C}$ are absent in the NMR data. Two structures of TagA$^{\Delta C}$ have been reported, a dimeric structure of the protein in its apo-state (Protein Data Bank [PDB]: 5WB4) and a trimeric structure in which it is bound to its UDP product (PDB: 5WFG) (37). Mapping the location of the missing amide signals onto the crystal structure of dimeric TagA$^{\Delta C}$ reveals that they are localized at the intersubunit interface, suggesting that signals for these residues are broadened beyond detection because of conformational exchange caused by dimer dissociation (Fig. 3.2*B*). In contrast, when the missing residues are mapped onto the trimeric structure of the TagA$^{\Delta C}$:UDP complex, many of the missing residues are not located at the intersubunit interface, suggesting that this form of the protein is not as prevalent. Thus, we conclude from the NMR data that in solution the dimeric form of TagA$^{\Delta C}$ predominates and that it resembles the previously reported crystal structure of apo-TagA$^{\Delta C}$.

We next used NMR and SE–AUC experiments to investigate full-length TagA (M1–R244), the functional form of the enzyme that associates with the membrane *via* residues within its CTT (Fig. 3.2*C*) (36). SE–AUC experiments indicate that TagA exists as a mixture of aggregated heterogeneous species, as its weight-averaged molecular weight decreases as the centrifugation speed increases (Fig. S3.1) (38). NMR studies of a 0.3 mM sample of [U-^{15}N] TagA are consistent with the SE–AUC data, as its ^{1}H–^{15}N TROSY–HSQC spectrum exhibits severe line broadening (Fig. 3.2*C*, *middle*). The positions of the cross peaks in the spectra of TagA$^{\Delta C}$ and TagA are similar, indicating that the structure of the core domain is retained. Notably, signals for the 49 residues in the CTT are absent in the spectra of TagA, and new signals in the indole region of the spectrum that presumably originate from the side chains of W211 and W220 within the CTT are significantly broadened. From these data, we conclude that the full-length protein adopts a conformation that is similar to that observed in the crystal structure of dimeric TagA$^{\Delta C}$, but residues in the CTT are likely disordered and mediate nonspecific interactions that promote aggregation in solution.

3.4.2 Crystal structure of solubility-enhanced TagA (TagA^{FL}∗) bound to its sugar–nucleotide substrate

Only the structure of TagA$^{\Delta C}$ lacking the CTT has been determined because the full-length protein exhibits low solubility. To overcome this problem, we modeled the structure of full-length TagA to identify potential locations on the protein's surface that might promote aggregation. We first generated a model of full-length *B. subtilis* TagA using the Generative Regularized Models of Proteins (GREMLIN) method, which predicts protein structures by integrating sequence conservation and coevolutionary patterns in a multiple sequence alignment (36, 39). The full-length *T. italicus* TagA computational model (TagA^{CM}) used in this study was then generated by

homology modeling using the GREMLIN-derived structure as a template for input into the program I-TASSER (Iterative Threading Assembly Refinement) (40). Coevolution analyses are a robust approach to model the structure of full-length *T. italicus* TagA, as structures similar to $TagA^{CM}$ are generated using coevolution-derived residue–residue distance probabilities implemented in the RaptorX program (Fig. S3.2) (41, 42). $TagA^{CM}$ is monomeric with the CTT forming three helices (H10', H11', and H12') that pack against the oligomerization interface observed in the crystal structure and is positioned against the body of the protein by several pairs of coevolving residues (Fig. 3.3, *A* and *B* and Table S3.1). Inspection of $TagA^{CM}$ reveals that the CTT contains many nonpolar amino acid side chains that colocalize to form a hydrophobic patch that might promote protein aggregation in solution. Prior cell fractionation studies of *B. subtilis* revealed that several residues within the patch are essential for membrane association (36). To improve protein solubility in solution, we constructed $TagA^{FL}*$, a variant of TagA in which four nonpolar residues in the patch are exchanged with hydrophilic amino acids (I203E/L209Q/L212K/I216E) (Figs. 3.1*B* and 3.3*C*). Indeed, $TagA^{FL}*$ exhibits markedly improved solubility and no longer requires detergents to remain soluble. Moreover, unlike native full-length TagA, which nonspecifically aggregates (Fig. S3.1), $TagA^{FL}*$ is in equilibrium between monomeric and dimeric states; SE–AUC experiments indicate that the $TagA^{FL}*$ dimer has a K_D of 210 ± 50 μM (Fig. 3.3*D*). This supports previous size-exclusion chromatography experiments that asserted that $TagA^{FL}*$ contains both monomeric and dimeric species in aqueous solution, and the monomeric species is more abundant compared with native TagA (36). $TagA^{FL}*$ also has *in vitro* enzymatic activity that is superior to $TagA^{\Delta C}$, albeit lower than that of the native enzyme (discussed below).

To gain insight into the molecular basis of substrate binding, we determined the crystal structure of the solubility-enhanced $TagA^{FL}*$ enzyme bound to its UDP–ManNAc substrate. The crystal structure of the $TagA^{FL}*$:UDP–ManNAc complex was resolved to 3.3 Å and solved using molecular replacement with the coordinates of $TagA^{\Delta C}$ (PDB: 5WB4) as a search model (data collection and refinement statistics in Table 3.1). The diffraction dataset contained moderate anisotropy, so we applied anisotropic scaling and isotropic *B*-factor sharpening prior to molecular replacement (43). The complex crystallized with three molecules in its asymmetric unit, with each intersubunit interface burying only ∼430 $Å^2$ of surface area (Fig. 3.4*A*). These small protein–protein interfaces are not deemed to be biologically relevant based on an analysis using the Evolutionary Protein–Protein Interface Classifier server (44). Only coordinates for the core domain could be modeled (residues M1–R199), as no interpretable electron density for the CTT was observable. The reason for its absence is unclear, but a large volume of nonuniform density is located adjacent to the C terminus of each protein molecule, suggesting that the CTT may be disordered in the crystal. It is also conceivable that during the crystallization process, which took approximately 1 month, the CTT was degraded by proteolysis. The positioning of the UDP–ManNAc substrate is best defined by the electron density data in chain A and is discussed further (Figs. 3.4*B* and S3.3*A*) (45). The uracil base of UDP–ManNAc is stabilized by hydrogen bonding with the terminal amino group of K166, the backbone carbonyl of G136, and the backbone amide of A109, along with π-stacking interactions with the Y137 tyrosyl ring (Fig. 3.4*C*). The ribose in the substrate is stabilized by two strong hydrogen bonds, one between the 3′ hydroxyl hydrogen donor and the terminal carboxylate acceptor of D191 and the other between the 2′ hydroxyl and the terminal hydroxyl of Y137 (Fig. 3.4*C*). Finally, the ManNAc sugar accepts two hydrogen bonds donated from the side-chain amino groups of N39 and Q167 to the substrate 3′ hydroxyl

oxygen and the *N*-acetyl carbonyl oxygen of ManNAc, respectively (Fig. 3.4*C*). The enzyme surface that binds the substrate is near residue D65, which has been shown to be important for catalysis, consists of an electronegative portion that contacts ManNAc and an electropositive region that interacts with the uracil and ribose rings (Fig. 3.4*D*) (36).

To investigate how enzyme contacts to the substrate's *N*-acetyl group affect sugar donor binding, we also determined the structure of $TagA^{FL}*$ bound to UDP–GlcNAc, a C2' epimer of UDP–ManNAc where the *N*-acetyl group is inverted. The $TagA^{FL}*$:UDP–GlcNAc complex crystallized with three molecules in the asymmetric similar to the $TagA^{FL}*$:UDP–ManNAc complex and was resolved to 3.0 Å. The structures of the core domain in each complex are nearly identical with a C_α RMSD of 0.28 Å. However, in the $TagA^{FL}*$:UDP–GlcNAc complex, the GlcNAc adopts two distinct conformations: conformer 1 in which it is displaced from the body of the protein (60% abundance) and conformer 2 that resembles the positioning of the substrate in the UDP–ManNAc complex (40% abundance) (Figs. 3.4*B* and S3.3*B*). However, as compared with the $TagA^{FL}*$:UDP–ManNAc structure, the most abundant GlcNAc conformer is positioned farther away from the catalytically important D65 residue, and both GlcNAc conformers form distinct contacts to the enzyme from their *N*-acetyl groups (Fig. S3.3, *C* and *D*). As TagA can only utilize UDP–ManNAc as a substrate, it seems likely that core domain contacts to the *N*-acetyl and 3′ hydroxyl groups in the substrate from the side chains of Q167 and N39 are important specificity determinants.

3.4.3 MD simulations provide insight into the function of the CTT

The crystal structure of $TagA^{FL}*$ bound to UDP–ManNAc lacks density for the CTT that is required for efficient catalysis *in vitro* (36). A comparison of the $TagA^{FL}*$:UDP–ManNAc

complex and TagA^{CM} structures suggests that the CTT may be involved in substrate binding. To gain insight into the function of the CTT, we performed MD simulations of full-length monomeric TagA in its apo and UDP–ManNAc bound states. The program MODELLER was used to generate starting models for MD using the coordinates of the TagA^{FL}∗:UDP–ManNAc complex and TagA^{CM} (described in the Experimental procedures section) (46). The simulations were equilibrated until the protein's backbone RMSD indicated good convergence, requiring 20 ns for the apo state and 200 ns for the complex. Production simulations were then performed for a duration of 1.0 μs for both systems. For simulations of the complex, the positioning of the substrate relative to the core domain was held fixed using distance restraints.

The simulation of the apoenzyme shows that the CTT remains packed against the body of the protein but transiently moves to expose the UDP–ManNAc binding site on the core domain. The position of the CTT against the core domain resembles the TagA^{CM} model derived solely from coevolutionary restraints, suggesting that it represents a low-energy state. In particular, H12' primarily rests in a groove formed by core domain helices H2 and H4, and helix H11' is positioned adjacent to the surface formed by residues in the β4–H4 and H4–H5 loops. Interestingly, the root mean square fluctuation of the apoprotein's backbone atoms during the trajectory reveals significant mobility in the H10' and H11' helices, as their backbone atom coordinates fluctuate by as much as 4 Å (Fig. 3.5*A*). Large coordinate variations are also observed in residues located at the C-terminal end of H12' (H238–R244), which are poorly conserved and transiently unwind. To gain deeper insight into the nature of the H10'–H11' fluctuations, we clustered conformers in the trajectory that adopted similar structures (47). More than 95% of conformers in the MD trajectory can be grouped into six clusters of related structures, in which members of each cluster exhibit backbone coordinate RMSDs that are less than 0.9 Å. Comparing the central structure from each

cluster reveals that helices H10' and H11' in the CTT move as a unit that toggles between two extrema: an “in state” in which they are positioned nearer the β4–H4 and H4–H5 loop elements on the core domain and an “out state” in which they are further displaced from the body of the protein (Fig. 3.5, *B–D*). As expected, the two extrema are in fairly low abundance (2.2% and 1.2% of the trajectory for the “in state” and the “out state,” respectively), and the vast majority of the simulation (>90% of the trajectory) falls in transient stages between the extremes. Upon closer inspection, multiple periods of transition occur between the “in state” and “out state” for the H10' and H11' helices throughout the simulation (Fig. S3.4*A*). Transient enzyme toggling between states exposes the surface that binds UDP–ManNAc in the crystal structure of the complex, suggesting that these motions may have a functional role in promoting substrate access to the active site (Fig. 3.5, *C* and *D*).

Simulations of the TagA^{CM}:UDP–ManNAc complex reveal that the CTT plays a prominent role in binding UDP–ManNAc and that it becomes immobilized upon substrate binding. Substrate-dependent CTT immobilization is evident from a plot of the root mean square fluctuation values for the backbone atoms during the simulation of the complex, as they generally exhibit small magnitudes throughout the protein (Fig. 3.5*A*). Moreover, more than 95% of the conformers in the trajectory (represented by eight clusters) contain a CTT that adopts an “in state” similar to that observed in the apoenzyme simulation, and the coordinates of the H10'–H11' unit are stable throughout the simulation (Figs. 3.5*E* and S3.4*B*). In addition, the C-terminal H12' helix is displaced upon substrate binding, moving from the groove formed by core helices H2 and H4 in the apoform to a groove formed by helices H2 and H8 in the complex (Fig. 3.5*F*). As helix H12' no longer packs against helix H4 of the core domain, the newly freed H4 and its following H4–H5 loop exhibit increased coordinate fluctuations between residues 74 and 82 (Fig. 3.5*A*). The most

deviant coordinate fluctuations appear again for the C terminus of helix H12' (H238–R244), which transiently unwind similar to the apo simulation. Conformational ordering of the CTT appears to be triggered by substrate contacts from two residues located at the N terminus of helix H12', R221, and R224, which form ionic and hydrogen bond interactions with the diphosphate group in UDP–ManNAc. In the representative clusters for the complex simulation, R221 and R224 work in tandem to stabilize the anionic charges of the substrate (Fig. 3.5*G*). R221 most often makes contacts with the β-phosphate oxygen of UDP–ManNAc and is sometimes proximal to the α-phosphate oxygen atoms. In the eight clusters representing 95% of the trajectory, R224 appears to occupy two different side-chain conformations throughout the simulation. The most populated orientation of R224 (represented by cluster 1) participates in cation–π interactions with F231, and the ε-amine donates a hydrogen bond to an α-phosphate oxygen (Figs. 3.5*G* and S3.5*A*). The minor orientation (represented by cluster 2) makes ionic interactions near α- and β-phosphate oxygen atoms by a terminal amino group, whereas the ε-amine donates a new hydrogen bond to the 6′ hydroxyl oxygen of ManNAc (Figs. 3.5*G* and S3.5*B*). The ionic interactions of these arginine residues with the substrate may be important for catalysis, as they are highly conserved in the primary sequences of TagA homologs (Fig. 3.1*B*). Substrate binding causes the CTT to adopt a more compact state as it engages the core domain; in more than 95% of the conformers in the trajectory, helix H10' packs against the H4–H5 loop, and helix H11' is placed near helix H4. Substrate-induced closure of the CTT over the nucleotide substrate positions several highly conserved amino acids in CTT near the active site whose significance in catalysis is investigated later (Figs. 3.1*B* and S3.5).

3.4.4 Highly conserved residues within the CTT are required for catalysis

TagA homologs contain five highly conserved amino acids in the CTT whose side chains project toward the bound substrate or the core domain in the MD simulation of the enzyme–substrate complex (E210, W211, R214, R221, and R224) (Figs. 3.1*B* and S3.5). To assess their importance in catalysis, we determined the *in vitro* GT activity of TagA variants containing amino-acid substitutions at these sites (Fig. 3.6*A*). A previously described enzyme-coupled end-point activity assay was employed to measure the rate of UDP product formation (19, 36). The TagA variants studied include five single amino-acid mutants that alter conserved amino acids in the CTT (E210A, W211A, R214E, R221E, and R224E), a variant that serves as a negative control because it alters a surface-exposed side chain in the CTT that is not expected to be involved in catalysis (I203E) and the $TagA^{\Delta C}$ and $TagA^{FL}*$ constructs. Severe reductions in activity relative to native TagA occur when the CTT is removed ($TagA^{\Delta C}$ is approximately eightfold less active), or key reactants are withheld from the assay (∼15-fold activity reductions are observed when the lipid substrate, MnaA epimerase, or TagA enzyme are not present) (Fig. 3.6*B*). These reductions in activity also occur when the three conserved arginine residues in the CTT are altered, with R214E and R221E substitutions causing the largest decline (Fig. 3.6*B*). In the complex MD simulation, R221 and R224 directly interact with the UDP–ManNAc substrate, whereas the side chain of R214 may be involved in lipid-α binding (see Discussion section). Altering other conserved residues within the CTT (E210A and W211A) also reduces activity, albeit to a much lesser extent (Fig. 3.6*B*). An I203E variant serves as negative control as its side chain projects into the solvent in the complex MD simulation and, as expected, it exhibits wildtype-level activity. The $TagA^{FL}*$ variant that alters four nonpolar residues in the CTT to improve enzyme solubility has reduced activity; however, it is still nearly three times more active than $TagA^{\Delta C}$ in which the CTT

is removed (Fig. 3.6*B*). The activities of the R221A and TagA$^{\Delta C}$ variants have been measured previously (36). The new data presented here substantiate and expand upon these findings by showing that two additional highly conserved arginine residues in the CTT are important for catalysis.

3.4.5 Native MS suggests that monomeric TagA interacts with membranes

Cell fractionation studies previously revealed that the *B. subtilis* TagA enzyme is a peripheral membrane protein and that nonpolar residues in a hydrophobic patch on CTT affect interactions with the lipid bilayer (36). We investigated how the closely related *T. italicus* TagA interacts with membranes using native MS, which enables the study of intact proteins and their complexes using electrospray ionization MS. A negative ion mode MS analysis of 10 μM TagA in aqueous solution resolves monomeric and dimeric forms of the protein, with a slight enrichment for the dimer (Fig. 3.7*A*). This is consistent with NMR and SE–AUC studies of TagA, which were performed at much higher concentrations that presumably cause the protein to nonspecifically aggregate. The behavior of native TagA in the presence of *n*-dodecyl-β-D-maltoside (DDM), an MS-compatible detergent that forms micelles at higher concentrations, was monitored by native MS (48, 49, 50). With DDM concentrations below its critical micelle concentration, relatively small effects on the TagA monomer–dimer equilibrium are observed by native MS. Interestingly, as the DDM concentration is increased above its critical micelle concentration (ca. 170 μM), there is a shift in TagA toward its monomeric state, suggesting this form of the enzyme is stabilized by the DDM micelles. Combined, the MS, MD, and activity data suggest that the structure of monomeric TagA resembles the TagA^{CM} model, and that this represents the active form of the enzyme that can associate with lipid bilayers *via* its CTT.

3.5 Discussion

TagA ManNAc transferases catalyze the first committed step in WTA biosynthesis and are a promising antimicrobial drug target (6, 7). At present, only the structure of the TagA homolog from *T. italicus* has been reported, revealing that these enzymes contain a unique GT core domain that is followed by a conserved ∼50 amino acid CTT of unknown structure (Fig. 3.1) (36). The CTT targets the enzyme to the cell membrane and is required for catalysis *in vitro* (36). However, its role in catalysis is not well understood because only a structure of a C-terminally truncated variant in which the CTT was removed has been determined ($TagA^{\Delta C}$) (36). $TagA^{\Delta C}$ forms dimeric and trimeric crystals when the protein is in its apo- and UDP-bound states, respectively (36). Our working hypothesis is that only the monomeric form of TagA is biologically active, since a computationally derived model of the monomeric structure of TagA (called $TagA^{CM}$) places several highly conserved and functionally important residues in the CTT near the presumed active site on the core domain (Fig. 3.3*B*). Here, we used a combination of crystallography, molecular modeling, MS, MD simulations, and functional assays to gain insight into how the core domain and CTT in TagA work together to catalyze glycosylation at the cell membrane.

NMR and AUC studies of the full-length protein reveal that the CTT promotes nonspecific aggregation (Figs. 3.2*C* and S3.1). Using the $TagA^{CM}$ model as a guide, we constructed $TagA^{FL}*$, which is significantly more soluble than the native protein because it eliminates four nonpolar residues in the CTT that are predicted to form a hydrophobic patch (Fig. 3.3*C*). Unlike the native protein, $TagA^{FL}*$ is in equilibrium between monomeric and dimeric states (Fig. 3.3*D*). A crystal structure of the $TagA^{FL}*$:UDP–ManNAc complex provides insight into the mechanism of substrate binding. This structure contains a core domain that adopts a Rossmann-like fold in which the UDP–ManNAc sugar donor is positioned at the edge of the sheet (coordinated by residues in loops β5–

H6, β6–H7, β7–H8 and β8–H9, and helices H2 and H9) (Fig. 3.4*C*). The substrate is positioned near residue D65 in the core domain, an essential catalytic residue. UDP–ManNAc binds to TagA similar to its UDP product, which was visualized previously in the crystal structure of the $TagA^{\Delta C}$:UDP complex (36). However, in the product complex, the diphosphate group of UDP contacts the protein's core domain, whereas in the structure of the UDP–ManNAc complex, it does not. TagA cannot utilize UDP–GlcNAc as a substrate, a closely related sugar nucleotide that differs from UDP–ManNAc only in the stereochemistry of the C2' atom that is bonded to the *N*-acetyl group. The structure of the $TagA^{FL}$∗:UDP–ManNAc complex reveals sugar contacts to the TagA core domain that contribute to stereoselectivity. In particular, the *N*-acetyl group of ManNAc is projected into a pocket where its carbonyl oxygen accepts a hydrogen bond from the amino group in Q167, and the 3′ hydroxyl oxygen accepts a hydrogen bond from the amino group in N39 (Fig. 3.4*C*). In contrast, the distinct stereochemistry of UDP–GlcNAc precludes these contacts, and instead the GlcNAc adopts two distinct conformers that are presumably enzymatically inactive (Fig. 3.4*B*).

Our results reveal that the CTT plays a major role in facilitating catalysis by encapsulating the UDP–ManNAc substrate and by contributing catalytically important side chains to the active site. We modeled the structure of the full-length TagA enzyme containing the CTT using coevolution restraints and the coordinates of the $TagA^{FL}$∗:UDP–ManNAc complex. MD simulations of the complex reveal that the CTT covers the UDP–ManNAc molecule, forming three α-helices (H10'–H12') that project five highly conserved residues toward the active site: E210, W211, R214, R221, and R224 (Fig. S3.5). Prior work demonstrated that R221 in the CTT is crucial for catalysis, but the role of other conserved amino acids remained unknown. In this study, we demonstrate that all conserved arginine residues in the CTT are critical for catalysis (R214, R221,

and R224). MD simulations shed light onto the functions of R221 and R224, as both residues form favorable and long-lived ionic interactions with the diphosphate group of UDP–ManNAc (Fig. 3.5*G*). The conserved R214 residue in the CTT is also extremely important for catalysis and may be involved in binding lipid-α (see later). In the MD simulation of the complex, the other conserved residues in the CTT (E210 and W211) pack against the core domain and may stabilize the positioning of the CTT over the UDP–ManNAc substrate. However, these interactions appear to be less important for catalysis as E210A and W211A variants retain ∼80% and ∼50% activity as compared with wildtype TagA, respectively (Fig. 3.6*B*).

The data presented here provide new insight into the molecular basis of catalysis. From this work and a previously published study, a total of four highly conserved residues in TagA have been shown to be important for catalysis, and when altered, they reduce enzyme activity to less than 25% of the wildtype enzyme (36). These residues include R214, R221, and R224 in the CTT and D65 located in the core domain (Fig. 3.6*B*) (36). TagA is a metal-independent inverting GT and therefore presumably catalyzes an S_N2-like direct displacement reaction at the anomeric carbon of the UDP–ManNAc sugar donor (19). In this reaction, the nucleophilic 4′ hydroxyl group of GlcNAc in lipid-α may be deprotonated for catalysis by the side chain of D65 located in a conserved pocket on the core domain. Indeed, the $TagA^{FL}$∗:UDP–ManNAc structure shows that the D65 side chain is near the C1' of ManNAc in a solvent-accessible pocket that the GlcNAc sugar of lipid-α presumably occupies when bound. Following activation by D65, the 4′ oxygen in lipid-α attacks the C1' anomeric carbon in UDP–ManNAc, promoting the breakage of its phosphodiester bond and transfer of ManNAc to lipid-α. Based on our MD simulations of the complex, R221 and R224 in the CTT likely play a key role in this process by stabilizing the UDP-leaving group through ionic and hydrogen-bonding interactions with its diphosphate group (Fig.

3.5*G*). How the enzyme recognizes lipid-α remains unclear, as it may also involve enzyme–bilayer interactions that are mediated by the CTT. In the MD simulation of the complex, a small pore formed between the CTT and core domain could serve as the entry point for lipid-α. Intriguingly, this pore connects the nonpolar surface on the CTT that is predicted to contact the membrane to a cavity that contains the C1' atom in UDP–ManNAc. Moreover, it is lined with the conserved E210 and R214 residues in the CTT as well as D65 located in the core domain that is conserved and has previously been shown to be important for catalysis (Figs. 3.6*B* and S3.5) (36). It is thus tempting to speculate that the pore serves as the entry point for lipid-α whose positioning would be stabilized by favorable electrostatic interactions between the diphosphate group of lipid-α and the guanidinium group of R214. A comparison of the MD trajectories of TagA in its free and UDP–ManNAc bound states reveals that UDP–ManNAc binding quenches motions in the H10' and H11' helices of the CTT (Fig. 3.5*A*). When bound, the substrate locks the CTT in an "in state" conformation, in which it presents highly conserved residues that either interact with the core domain (E210 and W211) or the UDP–ManNAc substrate (R221 and R224). This substrate-induced structural transition may explain the results of steady-state kinetics measurements that concluded that the *B. subtilis* TagA enzyme works *via* an ordered Bi–Bi mechanism in which UDP–ManNAc binds to the enzyme first, followed by lipid-α (19). However, additional studies are needed to investigate the molecular basis of this second binding event.

TagA and other members of the GT26 family are structurally unique GTs that share some commonalities with the recently characterized GT-D type enzymes. Structural analyses have identified four distinct GT superfamilies based on their folds, GT-A through GT-D (33, 51, 52). The majority of GT enzymes form water-soluble structures that fall into the GT-A or GT-B superfamilies or are members of the membrane-integrated GT-C superfamily (51). The GT-D

superfamily contains a single characterized enzyme, DUF1792, which transfers glucose from UDP-glucose to post-translationally modify a serine-rich repeat glycoprotein Fap1 in *Streptococci* and *Staphylococci* (52, 53). As previously described, TagA has a similar β-sheet topology and limited structural similarity to the GT-D domain in DUF1792 (DALI *Z*-score: 8.7) (Fig. S3.6) (36, 54, 55). However, the proteins differ markedly at their C termini. DUF1792 contains an extended loop region between the H10 and H11 helices in its core domain that encloses the substrate binding site and coordinates the catalytically essential Mn^{2+} cation with an acidic glutamate residue (Fig. S3.6) (52). In contrast, TagA enzymes eschew this extended loop in favor of a CTT appendage that encapsulates the donor sugar and contributes conserved arginine residues that promote metal-independent GT activity. Based on primary sequence alignments, the CTT is conserved among GT26 enzymes, with ~81% of these enzymes retaining this structural element and its essential conserved arginine residues. The surface-exposed hydrophobic patch is also conserved in these enzymes, suggesting that like TagA, they glycosylate membrane-embedded lipid substrates. Thus, we conclude that TagA and other GT26 enzymes adopt a novel GT-E type structural fold that is well suited for recognizing membrane-associated sugar acceptors.

The results presented here further support the hypothesis that TagA functions as a monomer on the cytoplasmic membrane. Native MS experiments of TagA at protein concentrations likely to occur within the cell (~10 μM) demonstrate that it is in equilibrium between monomeric and dimeric states (Fig. 3.7*A*). However, in the presence of micelles, the equilibrium is shifted to the monomeric form, presumably because bilayer interactions with the hydrophobic surface of CTTs in the monomer stabilize its structure and association with the core domain. As the CTT and potential intersubunit surfaces on the core domain are coincident, stabilizing CTT binding can be expected to limit oligomerization. This notion is compatible with SE-AUC studies of solubilized

variants of TagA ($TagA^{\Delta C}$ and $TagA^{FL}*$), which show that removing the CTT and its ability to compete for binding to the intersubunit interface promotes dimerization (the K_D for dimer association is 210 ± 50 μM for $TagA^{FL}*$ and 7.4 ± 0.7 μM for $TagA^{\Delta C}$). Finally, MD simulations and biochemical studies further validate the structure of the $TagA^{CM}$ model, which is shown to represent an energetically stable form of the protein that explains the importance of several conserved CTT (R214, R221, and R224) and core domain (D65) residues that are required for catalysis.

In conclusion, our data provide new insight into how TagA enzymes synthesize the linkage unit that joins WTA to the bacterial cell wall. When removed from the membrane and at low protein concentrations, TagA is in an equilibrium between monomeric and dimeric states (36). Upon encountering the cytoplasmic membrane that houses its lipid-α substrate, its monomeric form is stabilized *via* interactions between a hydrophobic patch on the CTT and the lipid bilayer (Fig. 3.7*B*). This form of the protein is enzymatically active as it enables the CTT to encapsulate the UDP–ManNAc substrate that is bound to the surface of the core domain. Both the CTT and core domain are instrumental in constructing the active site that contains UDP–ManNAc, with each presenting conserved residues that are important for catalysis (core domain: D65; CTT: R214, R221, and R224). Steady-state kinetics suggest that UDP–ManNAc binds first, which in MD simulations is shown to stabilize the closure of the CTT and the formation of a potential lipid-α binding surface. Based on modeling studies, we ascribe tentative functions for the conserved residues, with the arginine residues acting to stabilize the phosphate groups in UDP–ManNAc (R221 and R224) and lipid-α (R214) substrates. Meanwhile, D65 in the core domain may act as a general base, which deprotonates the terminal GlcNAc sugar in lipid-α. As other members of the GT26 family also contain a conserved CTT segment, they may glycosylate membrane lipid

substrates using a similar mechanism. The results of this study could facilitate the development of new antimicrobial compounds that work by disrupting TagA function and, in turn, WTA biosynthesis.

3.6 Experimental Procedures

3.6.1 Protein expression and purification

All TagA constructs were cloned and expressed using standard methods as described previously (36). Briefly, each construct contained an N-terminal 6×His-tag and tobacco etch virus protease recognition sequence (ENLYFQS) in the pMAPLe4 expression vector and was expressed in *E. coli* BL21(DE3) cells. Cultures were grown in the presence of 50 μg/ml kanamycin at 37 °C in a shaking incubator to an absorbance of 0.6 to 0.8 at 600 nm, before induction with 1 mM IPTG. Expression proceeded at 17 °C for 16 h. For isotopically labeled samples, the cell pellets were exchanged into M9 media supplemented with $^{15}NH_4Cl$, ^{13}C-glucose, and 70% or 99% deuterium oxide (Cambridge Isotope Laboratories) before induction (56). The cells were pelleted by centrifugation, and cells were resuspended in 40 mM CHAPS, 50 mM Tris–HCl, pH 7.5, 500 mM NaCl supplemented with 400 μl of protease inhibitor cocktail (Sigma), and 2 mM phenylmethanesulfonylfluoride, 1 mg egg white lysozyme, and 0.5 mg *Serratia marcescens* nuclease per liter of culture (57). Resuspended cells were lysed using an Emulsiflex high pressure homogenizer (Avestin), and soluble TagA lysate was clarified by centrifugation.

TagA protein was purified by passing over Co^{2+}–NTA HisPur resin (Fisher) and washed with 8 mM CHAPS, 50 mM Tris–HCl, pH 7.5, and 500 mM NaCl buffer. Protein was eluted from the column using the same buffer supplemented with 200 mM imidazole. Eluted protein was

concentrated using Amicon centrifugal filters (Fisher). The 6×His-tag was proteolytically removed using 0.5 mg tobacco etch virus protease and buffer exchanged into a 10% v/v glycerol, 50 mM Tris–HCl, pH 7.5, and 500 mM NaCl buffer. The protein was passed over Co^{2+}–NTA again and washed off the column with 10% v/v glycerol, 50 mM Tris–HCl, pH 7.5, and 500 mM NaCl buffer. Wash fractions containing protein were concentrated, and the protein was subjected to size-exclusion chromatography using a Superdex 75 preparation grade column (GE). Protein was concentrated prior to storage or use. TagA variants were tested for their *in vitro* activity using a previously described HPLC assay (19, 36).

3.6.2 NMR assignments and relaxation experiments

Isotopically labeled samples were dissolved in NMR buffer at pH 6.8 (50 mM sodium phosphate, pH 6.8, and 100 mM NaCl). $TagA^{\Delta C}$ was uniformly isotopically labeled with ^{15}N or $^{15}N/^{13}C/^{2}H$ (70%), whereas $TagA^{FL}*$ was uniformly isotopically labeled with ^{15}N or $^{15}N/^{13}C/^{2}H$. Spectra were acquired at 303 K on Bruker Avance III HD 600 MHz (14.1 T) and Bruker Avance NEO 800 MHz (18.8 T) spectrometers equipped with triple resonance cryogenic probes. Backbone assignments of $TagA^{\Delta C}$ were determined by carrying out TROSY-enhanced variants of the following experiments: ^{15}N-HSQC, HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, and HN(CO)CACB (58, 59). Data were processed using NMRPipe, and CARA was used to perform sequential assignment (60, 61). Secondary structures were predicted from secondary ^{13}C chemical shifts using TALOS-N (62).

The rotational correlation time (τ_c) of $TagA^{\Delta C}$ was estimated from a series of ^{15}N-TRACT experiments (63). Spectra were acquired using a 0.48 mM sample of ^{15}N-labeled $TagA^{\Delta C}$, with 2048 complex points, 128 transients, and 100 experiments for each spin state, and the relevant

delay incremented by 1.5 ms. The decay of integrated amide proton signals was fitted to an exponential decay resulting in a transverse cross-correlated relaxation rate (η_{xy}) of 28 Hz. This allowed calculation of τ_c ~27 ns *via* algebraic solutions (64).

3.6.3 SE-AUC

SE–AUC experiments were performed on an Optima XL-A analytical ultracentrifuge (Beckman Coulter). Three concentrations for each sample (TagA: 4.2, 13, and 19 μM; $TagA^{FL}$∗: 6.0, 18, and 28 μM) were subjected to three ultracentrifuge speeds (TagA: 12,000, 15,000, and 19,000 rpm; $TagA^{FL}$∗: 15,000, 19,000, and 24,000 rpm) and allowed to reach SE at 4 °C. Data regression analysis was performed using the Beckman-Coulter Optima Analytical Ultracentrifuge Origin Data Analysis Package. The data were fit to multiexponential and single-exponential models. $TagA^{FL}$∗ was best represented by a monomer–dimer multiexponential model that was calculated using the predicted monomeric molecular weight of 27,800 Da by the ExPASy ProtParam tool (65). The dissociation constant (K_D) was determined to be the inverse of $K_{A(conc)}$ using the following equation (66, 67):

$$K_{A(conc)} = K_{A(abs)} ((\varepsilon l)^{n-1}/n)$$

where the molar extinction coefficient (ε) was determined to be 21,720 cm^{-1} M^{-1} using the ExPASy ProtParam tool, l is the path length of 1.2 cm, n is the order of oligomerization, and $K_{A(abs)}$ is the absorbance association constant.

3.6.4 X-ray crystallography

The $TagA^{FL}$∗:UDP–GlcNAc and $TagA^{FL}$∗:UDP–ManNAc complexes were cocrystallized from solutions containing 15 mg/ml or 11 mg/ml $TagA^{FL}$∗, respectively, dissolved in 50 mM Tris–

HCl, pH 7.5, 500 mM NaCl buffer, and supplemented with either UDP–GlcNAc or UDP–ManNAc at fivefold molar excess. The UDP–ManNAc used to form the complex was produced using a chemoenzymatic method as described previously (68). Screening was done using the hanging-drop vapor diffusion method at 4 °C. The $TagA^{FL}*$:UDP–GlcNAc complex crystallized in 100 mM HEPES, pH 7.5, 10% w/v PEG 8000, and 8% v/v ethylene glycol. The $TagA^{FL}*$:UDP–ManNAc complex crystallized in 40 mM potassium phosphate monobasic, pH 7.5, 10% w/v PEG 8000, and 20% v/v ethylene glycol. Crystals were cryoprotected when harvested in reservoir solution containing 45% glycerol. Both datasets were obtained at the Advanced Photon Source beamline 24-ID-C. Data were acquired at 100 K with 0.25° oscillations and detector distances of 300 and 450 mm for the $TagA^{FL}*$:UDP–ManNAc and $TagA^{FL}*$:UDP–GlcNAc complexes, respectively.

XDS/XSCALE was used to perform indexing, integration, and scaling of each dataset in the $P3_121$ space group (69). Dataset anisotropy was corrected using the UCLA Diffraction Anisotropy Server and a -50 Å^2 isotropic *B*-factor sharpening was applied to improve the apparent *B*-factors of the $TagA^{FL}*$:UDP–ManNAc crystal structure (43). For both complexes, the asymmetric unit contained three molecules of $TagA^{FL}*$. Phases were determined by molecular replacement using Phaser and the previously published coordinates of $TagA^{\Delta C}$ (PDB: 5WB4) (70, 71). Structures were iteratively improved by rounds of manual model building in Coot and automated refinement using BUSTER (Global Phasing Ltd) and PHENIX (71, 72, 73). The GRADE server was used to parameterize UDP–ManNAc for subsequent refinement (73). To assess the accuracy of ligand placement, composite iterative-build omit maps were generated by refining the structures of $TagA^{FL}*$:UDP–GlcNAc and $TagA^{FL}*$:UDP–ManNAc using the AutoBuild program such that the resulting electron density in the omitted ligand region is unbiased

by the atomic model (45). Additional simulated-annealing composite omit maps were generated using PHENIX (71). Statistics for data collection and structure refinement are given in Table 3.1.

3.6.5 MD simulations

A model of the full-length *T. italicus* protein for MD simulations was constructed as follows. First, the GREMLIN method was used to model the structure of the full-length *B. subtilis* TagA protein using coevolutionary patterns in a multiple sequence alignment (39). Second, the structure of the full-length *T. italicus* TagA protein was constructed by homology modeling using the program I-TASSER and the GREMLIN-derived structure of the *B. subtilis* enzyme as a template (40). Finally, because the coordinates of the core domain in *T. italicus* TagA are known, the final structure used in the MD simulations contains the experimentally derived coordinates of the core domain in the $TagA^{FL}*$:UDP–GlcNAc crystal structure and the coordinates of the CTT derived from $TagA^{CM}$. The program MODELLER was used to combine the coordinates (residues 2–186 and 140–244 were combined using the coordinates of the $TagA^{FL}*$:UDP–GlcNAc crystal structure and $TagA^{CM}$, respectively) (46). The final merged coordinates were used in the MD simulations of the apoenzyme. The structure of the complex used in the simulations was obtained by positioning the ligand in a manner similar to that observed in the $TagA^{FL}*$:UDP–ManNAc crystal structure.

Explicit-solvent MD simulations of the generated model were run using GROMACS 2021 using the CHARMM36m force field (74, 75). The system was solvated with TIP3P, energy minimized, and equilibrated in two stages (NVT followed by NPT). For the complex model, the equilibration was further divided into two steps with both macromolecules held *via* position restraints and then restraints on the protein alone. 1.0 μs production MD simulations were then

performed on both the apo and complexed systems. For the complex, pairwise distances between the protein and UDP–ManNAc as observed in the crystal structure were restrained using three pull groups (N39 terminal amine nitrogen to UDP–ManNAc sugar 3′ oxygen, E41 C_α to UDP–ManNAc *N*-acetyl oxygen and Y137 C_ε to UDP–ManNAc uracil C5' carbon). Global backbone RMSD calculated over the course of the simulations revealed that they reached convergence after 20 ns (apo simulation) or 200 ns (complex simulation) of NPT simulation. Clustering was performed using the GROMACS cluster module with the GROMOS method and a 0.9 Å cutoff threshold where the least squares fit was calculated for backbone atoms of the core domain only (residues 1–195) (47).

3.6.6 Enzyme-coupled activity assay

The TagA enzyme-coupled assay was described previously and is briefly summarized here (19, 36). The reaction was analyzed using a paired ion chromatography HPLC assay and under the following conditions: 0.2 μM TagA; 100 μM lipid-α analog, 100 μM UDP–GlcNAc, 3 μM UDP-*N*-acetyl-d-glucosamine 2-epimerase (MnaA), 50 mM Tris–HCl, pH 7.8, and 250 mM NaCl (19, 76). Reactions were preincubated without TagA enzyme for 15 min at 30 °C. Then the TagA enzyme was added and incubated for 40 min at 30 °C before quenching with 4 M urea. The end-point reactions were separated with an Agilent Eclipse Plus C18 column using a buffer gradient of 65% buffer A (15 mM sodium phosphate, 10 mM tetrabutylammonium hydrogen sulfate, and pH 7.0) and 35% buffer B (35 mM sodium phosphate, 10 mM tetrabutylammonium hydrogen sulfate, 30% acetonitrile, and pH 7.0) to 100% buffer B over 4 min. UDP–ManNAc and UDP elution peaks were monitored at 262 nm and integrated to determine turnover rate.

3.6.7 Native mass spectrometry

TagA was exchanged into 500 mM ammonium acetate buffer, pH 7.4, and incubated at 4 °C overnight, then diluted to a concentration of 10 μM with varying concentrations of DDM (0–400 μM). Nanoelectrospray ionization native MS analysis was performed on a G2-Si Synapt (Waters Corporation) in negative ion mode with the following parameters: 2000 to 8000 *m/z*, capillary: 1 kV; source temperature: 50 °C; desolvation temperature: 150 °C; sampling cone: 20 V; trap collision energy: 6 to 50 V; and transfer collision energy: 2 V. The collision energies in the "trap" region of the mass spectrometer used to remove DDM adducts from the protein were: 6, 20, 20, 30, and 50 V for 0 μM, 10 μM, 100 μM, 200 μM, and 400 μM DDM, respectively. The *m/z* raw data were centered (15 channels, 80% centroid top) in MassLynx (version 4.1; Waters Corporation) prior to input and deconvolution in UniDec (77). The deconvolution parameters are as follows: peak full-width half-maximum: 50 Th; peak shape function: gaussian; beta: 1000; charge smooth width: 1.0; point smooth width: 100; and maximum number of iterations: 100.

3.7 Figures

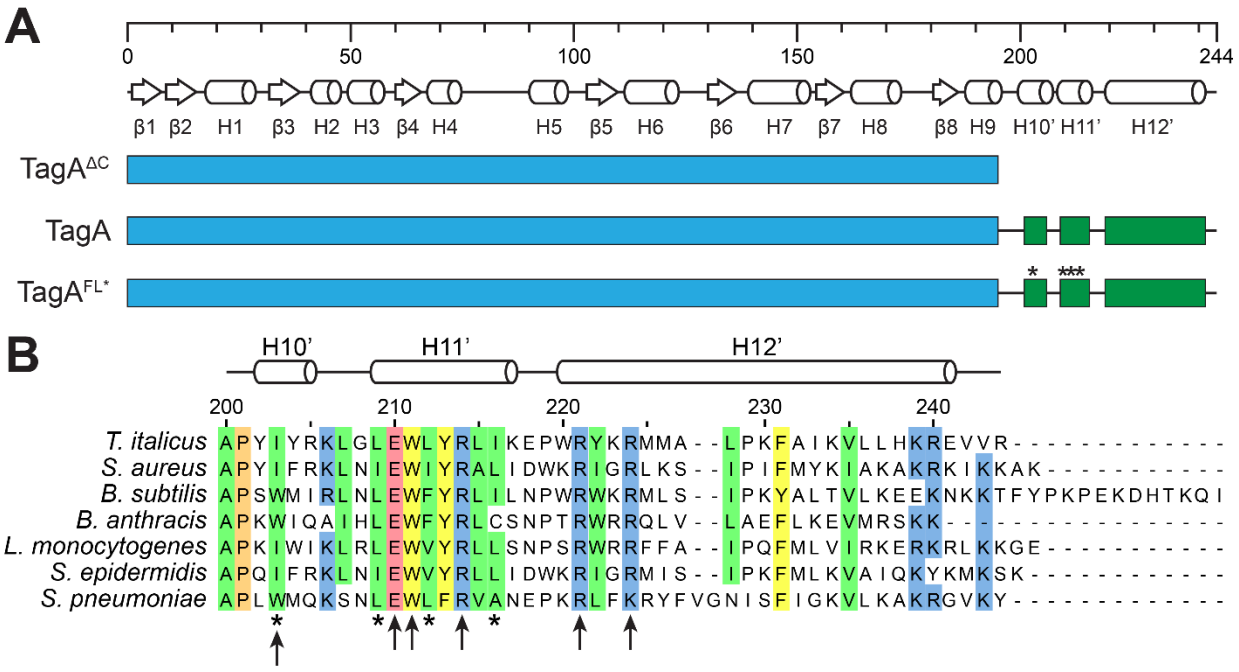

Figure 3.1 TagA protein constructs and sequence alignment. *A*, three *Thermoanaerobacter italicus* TagA proteins were used in this study: TagA$^{\Delta C}$, residues M1–G195; TagA, residues M1–R244; and TagA^{FL}*, residues M1–R244 with four amino-acid substitutions (I203E/L209Q/L212K/I216E). The folded core domain of TagA (*blue*) and computationally predicted helices (*green*) in the CTT are shown as *horizontal bars*. The secondary structure elements are shown with the predicted helices in the tail labeled H10', H11', and H12'. *Asterisks* indicate the location of the amino-acid substitutions in TagA^{FL}*. *B*, primary sequence alignment of select TagA proteins showing the CTT. Residues with high similarity across species are colored as follows: nonpolar (*green*), basic (*blue*), acidic (*red*), aromatic (*yellow*), and proline (*orange*). *Asterisks* indicate the location of the amino-acid substitutions in TagA^{FL}*, and *arrows* indicate conserved residues that were altered in this study. CTT, C-terminal tail.

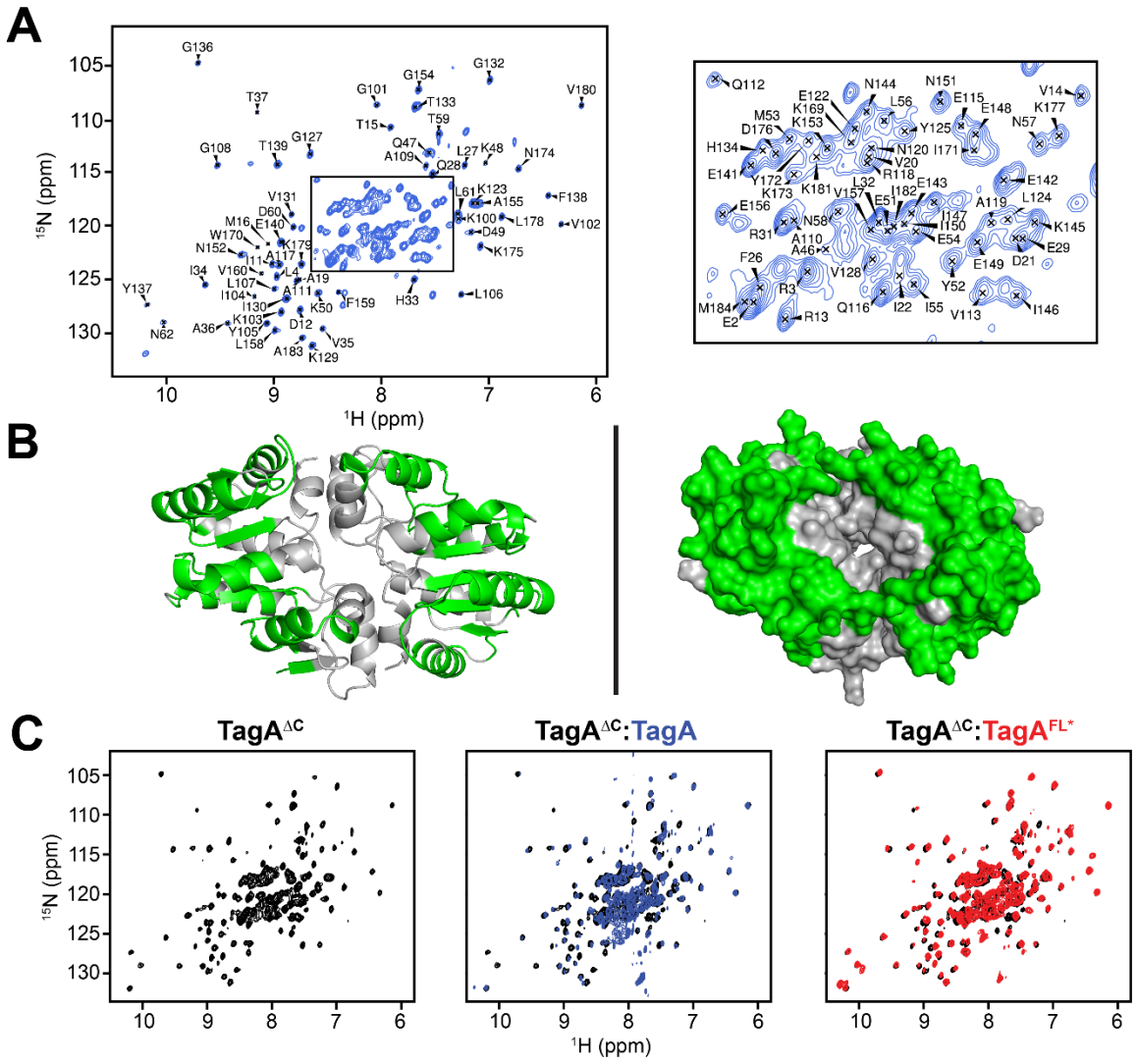

Figure 3.2 NMR studies of TagA. *A*, assigned ^{1}H–^{15}N TROSY–HSQC spectrum of TagA$^{\Delta C}$, with the center of the spectrum expanded and shown on the *right*. About 61% of amide residues were assigned for TagA$^{\Delta C}$. *B*, crystal structure of the TagA$^{\Delta C}$ dimer colored to show amino acids whose chemical shifts were assigned (*shaded green*). The TagA$^{\Delta C}$ dimer is shown in *cartoon* (*left*) and *surface* (*right*) representations (Protein Data Bank code: 5WB4). *C*, overlays of ^{1}H–^{15}N TROSY–HSQC spectra: TagA$^{\Delta C}$ alone (*left*), overlaid with the spectrum of TagA (*center*, *blue*) and with spectrum of TagA^{FL}* (*right*, *red*). HSQC, heteronuclear single quantum coherence; TROSY, transverse relaxation optimized spectroscopy.

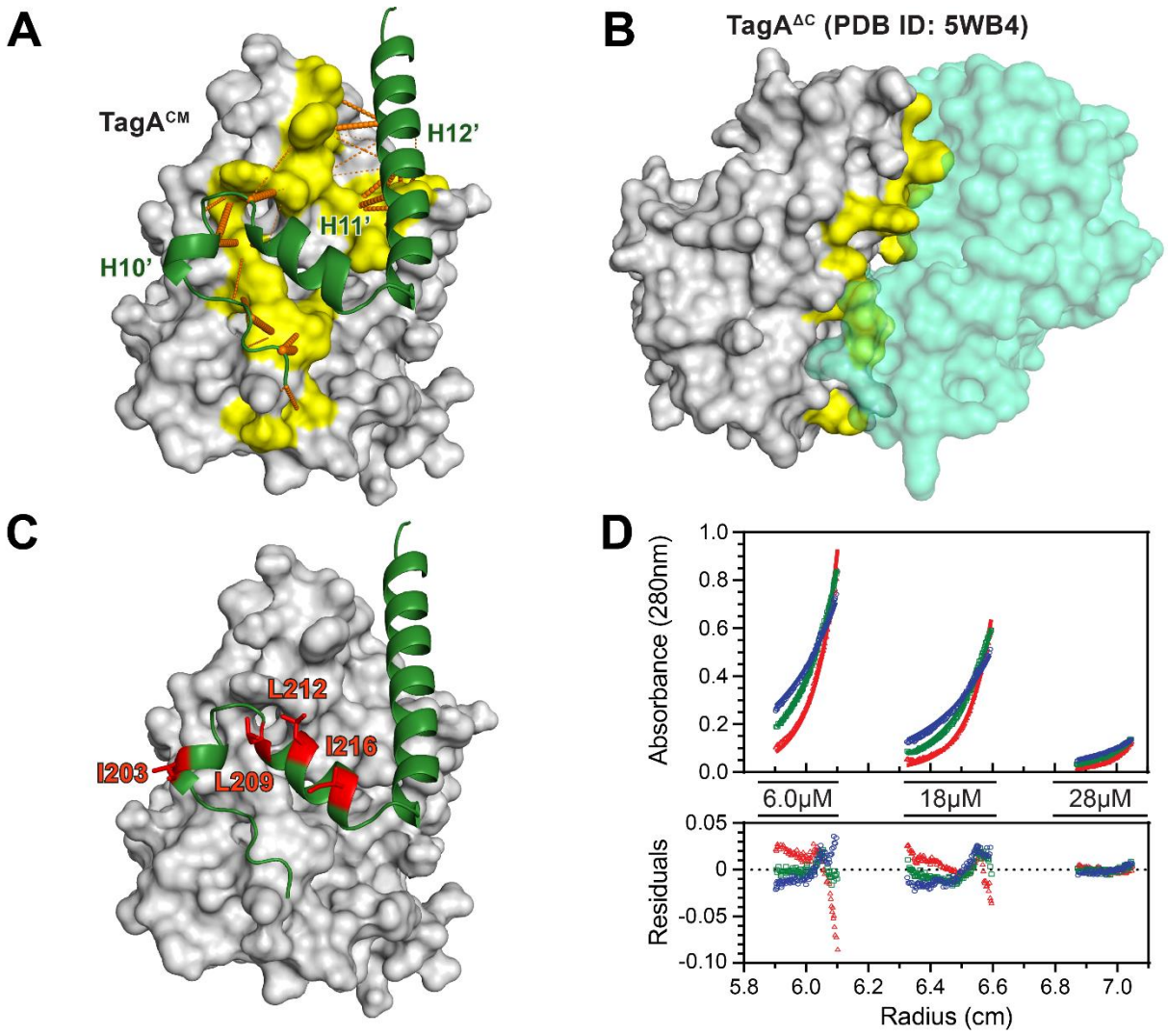

Figure 3.3 A model of the monomeric form of full-length TagA and its use in creating a solubility-enhanced TagA protein. *A*, TagA^{CM}, a computationally derived model of the TagA monomer. *Orange bars* mapped onto the structure connect coevolving amino acids between the core domain (*yellow surface*) and the CTT (*green cartoon*) that were identified using the program RaptorX (41). The thickness of the bars indicates the probability of residues being within 8 Å (50%—thin, 100%—thick). *B*, *surface* representation of the crystal structure of the TagA$^{\Delta C}$ dimer showing coevolving core domain residues in *yellow*. The predicted CTT binding surface on the monomeric form of the enzyme conflicts with the dimer interface. *C*, TagA^{CM} model showing the

location of the nonpolar amino acids (*red*) that were altered to create the solubility-enhanced protein ($TagA^{FL}*$, I203E/L209Q/L212K/I216E). *D*, SE–AUC experiments of $TagA^{FL}*$. Data were collected using three sample concentrations at three rotor speeds: 15,000 (*blue*), 19,000 (*green*), and 24,000 (*red*) rpm. Residuals after fitting the data to a monomer–dimer equilibrium are shown. CTT, C-terminal tail; SE-AUC, sedimentation equilibrium–analytical ultracentrifugation.

	TagA^{FL*}:UDP-GlcNAc	TagA^{FL*}:UDP-ManNAc
Data collection		
PDB Code	7MPK	7N41
Space group	P 3_1 2 1	P 3_1 2 1
Cell dimensions		
a, *b*, *c* (Å)	113.24, 113.24, 114.45	113.24, 113.24, 118.55
α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Resolution (Å)	74.5-3.0 (3.1-3.0)[a]	98.4-3.3 (3.4-3.3)[a]
Wavelength (Å)	0.9797	0.9791
R_{merge} (%)	12.0 (88.1)	8.0 (128.8)
I / σI	16.2 (2.9)	14.7 (1.6)
$CC_{1/2}$	99.8 (83.0)	99.8 (66.6)
Completeness (%)	99.7 (97.3)	99.9 (99.8)
Redundancy	9.8 (9.1)	7.4 (7.2)
Wilson B-factor (Å^2)	76.05	84.37
Refinement		
Resolution (Å)	74.5-3.0	98.4-3.3
No. reflections	17470	13694
R_{work} / R_{free}(%)[b]	20.6/25.3	19.1/24.7
No. atoms	4755	4612
Protein	4519	4456
Ligand/ion	234	156
Water	2	0
B-factors (Å^2) (all atoms)	71.2	91.8
Protein	71.4	91.6
Ligand/ion	66.4	97.5
Water	54.1	N/A
R.m.s. deviations		
Bond lengths (Å)	0.013	0.010
Bond angles (°)	1.55	1.26
Ramachandran favored (%)	98.1	96.4
Ramachandran allowed (%)	1.9	3.6
Ramachandran outliers (%)	0.0	0.0

[a]Values in parentheses are for the highest-resolution shell.

[b]R_{merge} calculated using 5% of collected experimental data.

Table 3.1 Crystal data collection and structure refinement statistics.

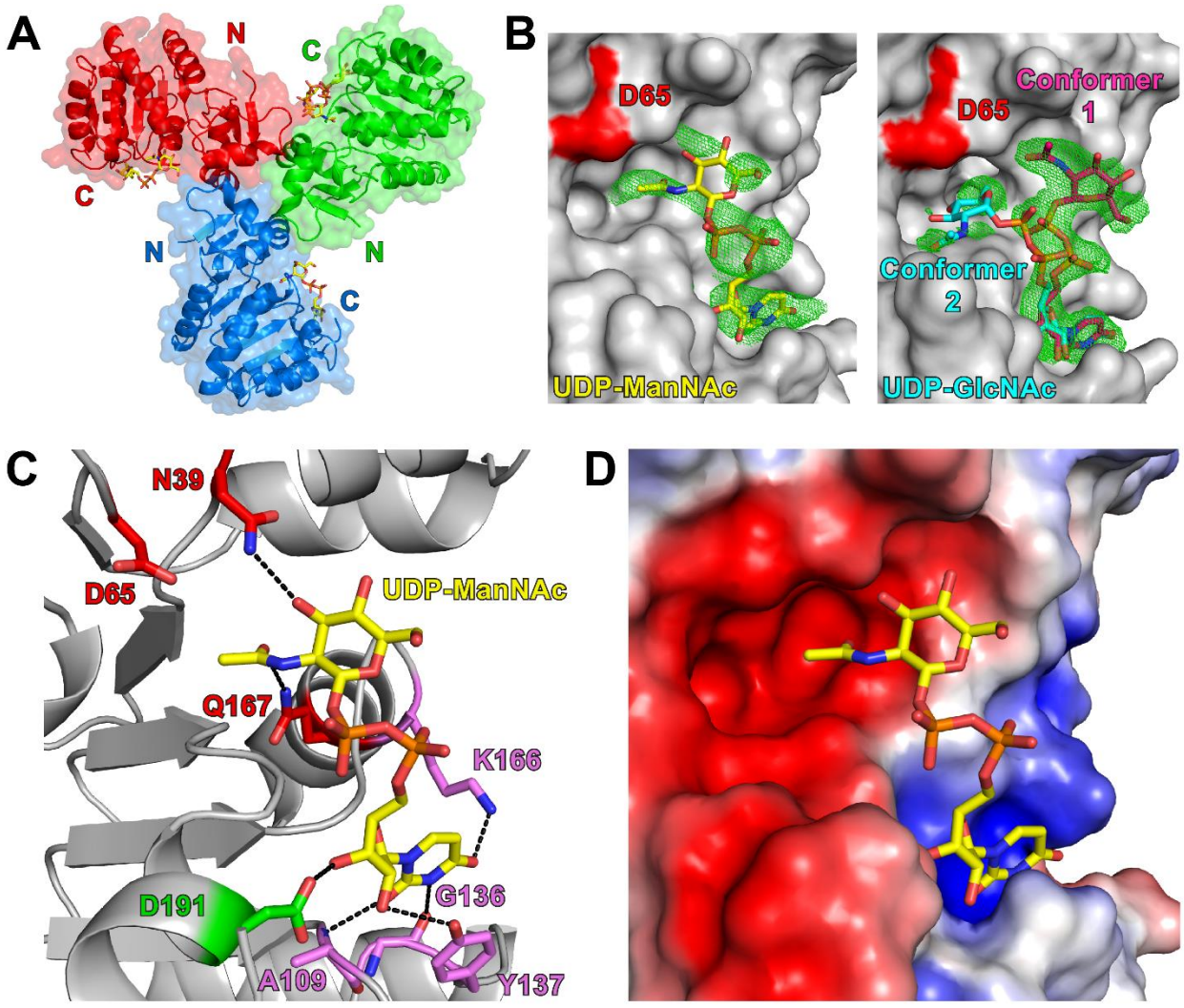

Figure 3.4 Structures of the TagA^{FL}∗:UDP–ManNAc and TagA^{FL}∗:UDP–GlcNAc complexes. *A*, the asymmetric unit of the TagA^{FL}∗:UDP–ManNAc complex (Protein Data Bank code: 7N41). Protein subunits are shown in *red*, *blue*, and *green*. The protein in the TagA^{FL}∗:UDP–GlcNAc complex (Protein Data Bank: 7MPK) adopts a very similar structure. *B*, iterative-build 2mF_o–DF_c composite omit maps showing the location of the UDP–ManNAc (*left*, *yellow*) and UDP–GlcNAc (*right*, *cyan* and *purple*) ligands in the TagA^{FL}∗ complexes (contoured at 1.0 σ) (45). UDP–GlcNAc adopts two conformations, 1 and 2, where the sugar is oriented away from or toward the catalytic pocket containing D65. Additional simulated annealing omit maps for the structures are

presented in Fig. S3.3. *C*, image showing enzyme–substrate interactions in the $TagA^{FL}*$:UDP–ManNAc complex. A *cartoon* representation of the protein is shown in *gray*. UDP–ManNAc and amino acid side chains are shown in *stick* format. Amino acid side chains are colored based on the part of the substrate they contact: uracil base (*purple*), ribose (*green*), and ManNAc moiety (*red*). *D*, electrostatic surface of the substrate binding site in the $TagA^{FL}*$:UDP–ManNAc complex. Coloring is as follows: anionic (*red*), neutral (*white*), and cationic (*blue*). GlcNAc, *N*-acetyl-D-glucosamine; ManNAc, *N*-acetyl-D-mannosamine.

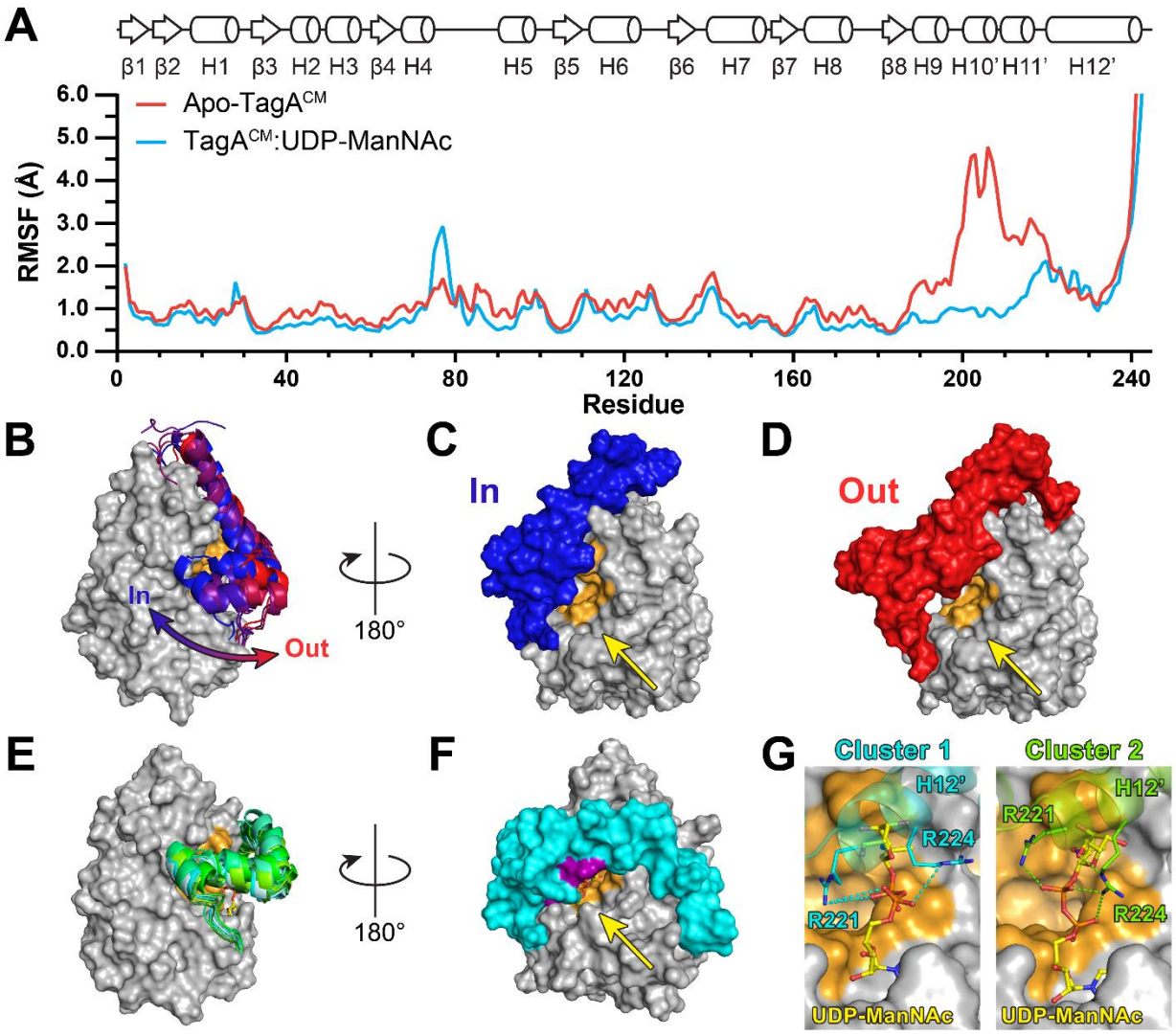

Figure 3.5 MD simulations of apo-TagA^{CM} and the TagA^{CM}:UDP–ManNAc complex. *A*, a plot showing the root mean square fluctuation (RMSF) differences of TagA backbone coordinates during apo and complex MD simulations. *B*, representative clusters of the apo-TagA^{CM} simulation showing the fluctuations of the C-terminal tail (CTT). The surface of the core domain is colored *gray*, residues within the catalytic pocket are colored *orange*, and helices H10'–H12' in the CTT are colored based on their positioning ("in state" [*blue*, *C*] or "out state" [*red*, *D*]). The UDP–ManNAc binding surface is indicated by *yellow arrows*. *E*, representative clusters of the TagA^{CM}:UDP–ManNAc complex simulation. The H10' and H11' helices are stable and pack

against the core in a conformation that resembles the “in state” observed in the simulation of apo-TagA^{CM}. The CTT for each cluster is represented in *cartoon* format and shaded from *light blue* to *green*. *F*, the surface representation of the primary cluster (*cyan*) from the complex simulation. Coloring as in panels (*B*–*D*). Conserved CTT arginine residues (R214, R221, and R224) are colored *purple*. *G*, enzyme–substrate contacts observed in the MD simulation of the TagA^{CM}:UDP–ManNAc complex. In the two most populated clusters in the trajectory (cluster 1, 51% of the trajectory, *cyan*; cluster 2, 25% of the trajectory, *green*), the diphosphate group in UDP–ManNAc interacts with two highly conserved arginine residues, R221 and R224. ManNAc, *N*-acetyl-D-mannosamine; MD, molecular dynamics.

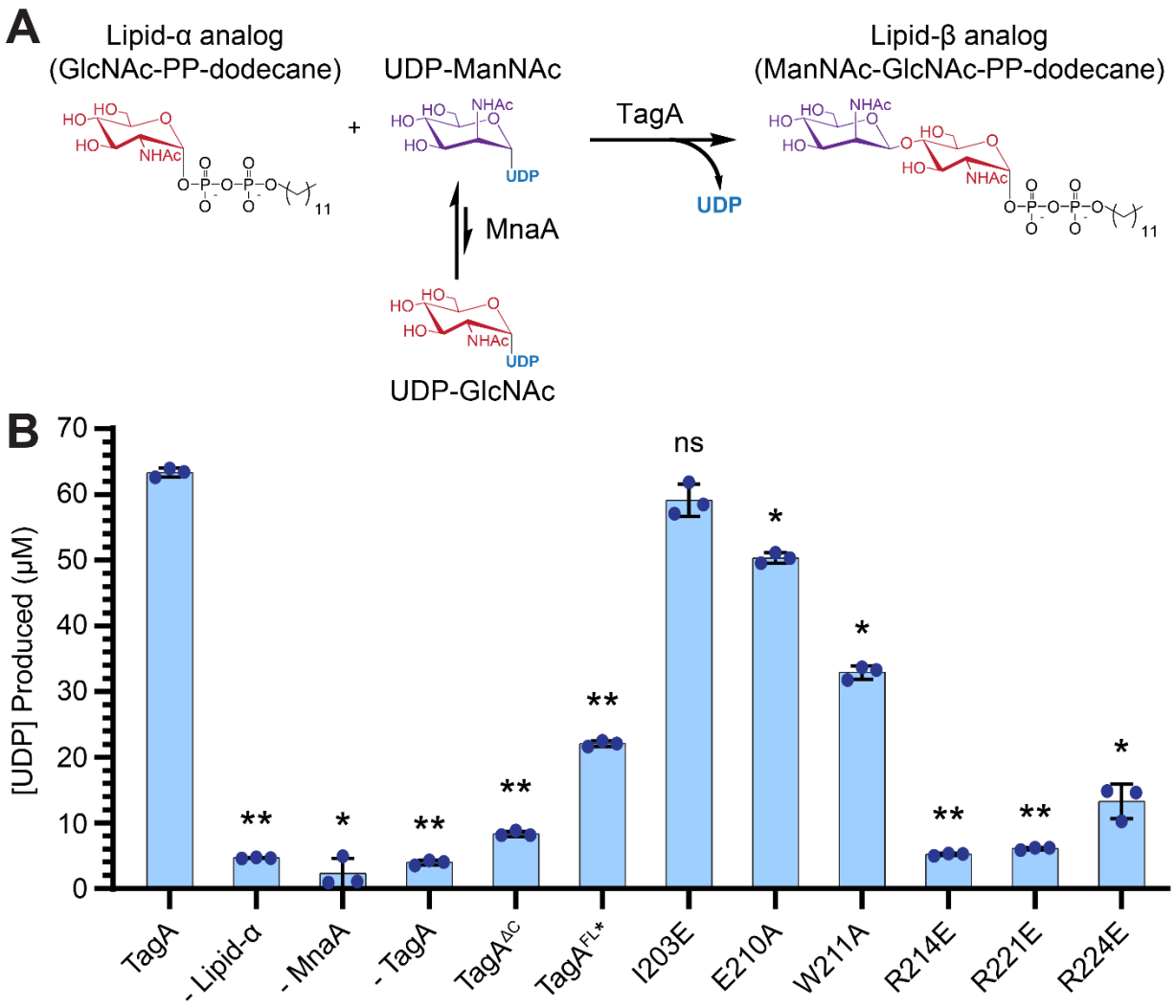

Figure 3.6 Enzyme activities of TagA variants. *A*, schematic of the *in vitro* enzyme-coupled TagA activity assay. Synthetic lipid-α analog is coincubated with UDP–GlcNAc and an MnaA epimerase before addition of the TagA enzyme. UDP product accumulation is quantified by absorbance at 262 nm. *B*, a chart of UDP product formation for a series of variants of the *Thermoanaerobacter italicus* TagA enzyme following an end-point activity assay. Bars labeled "-lipid-α," "-MnaA," and "-TagA" indicate assays in which these components were not present. All single amino-acid substitutions were introduced in the native full-length TagA enzyme. Each experiment was performed in technical triplicate, and error bars represent the standard deviation of measurements. The statistical significance (*p* value) between native TagA and variant activity

datasets was determined using the Analysis ToolPak in Microsoft Excel. *Asterisks* indicate the statistical significance between the activity of native TagA and variant datasets ($*p < 0.005$, $**p < 0.0001$, and ns). GlcNAc, *N*-acetyl-D-glucosamine; ns, not significant.

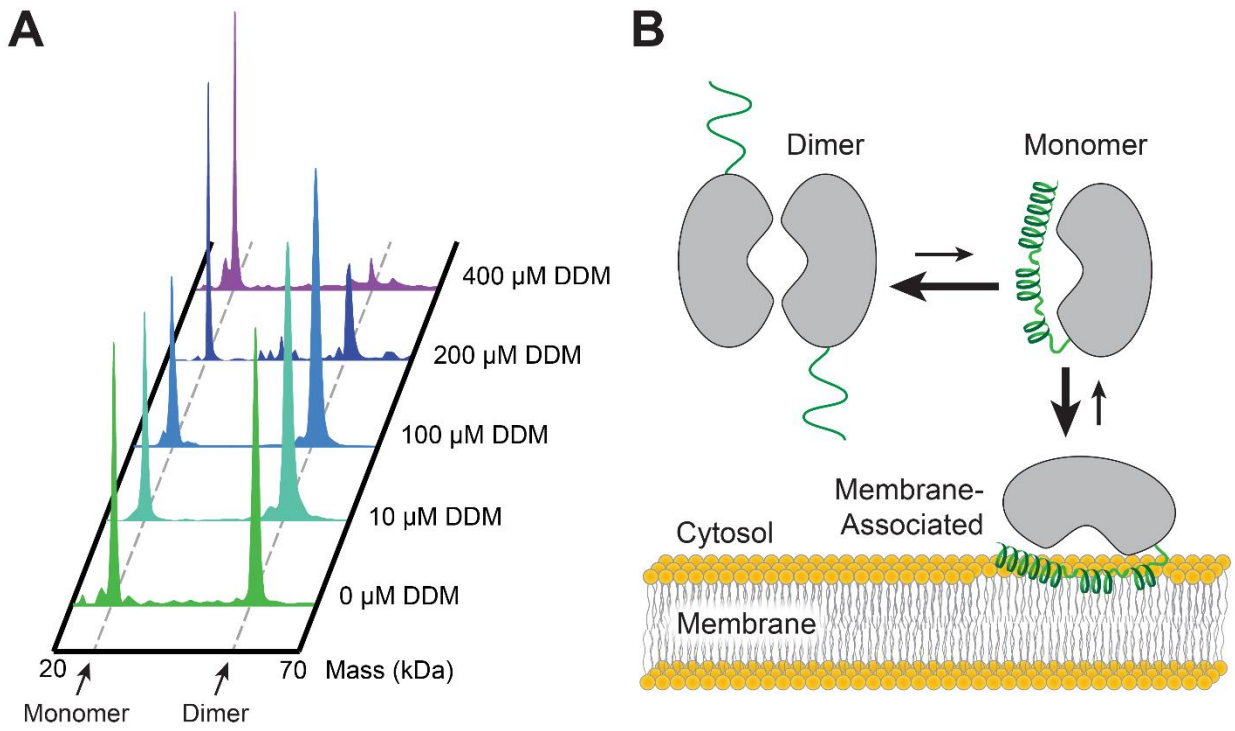

Figure 3.7 Native MS studies of micelle binding by TagA. *A,* negative ion mode MS of 10 μM TagA. Deconvoluted mass spectra of TagA with varying concentrations of DDM detergent (critical micelle concentration [CMC]: 170 μM) are shown with monomeric and dimeric peaks identified. *B,* proposed model of membrane association by TagA based on the results of experimental and computational studies. In solution, TagA exists in a monomer–dimer equilibrium. When a membrane is present, a hydrophobic patch on the CTT that exists only in the monomeric form of the enzyme favorably interacts with the membrane and is stabilizing. As a monomer on the membrane, TagA is poised to bind to its lipid-α substrate. CTT, C-terminal tail; DDM, *n*-dodecyl-β-D-maltoside.

3.8 Supporting Information

3.8.1 Supporting figures

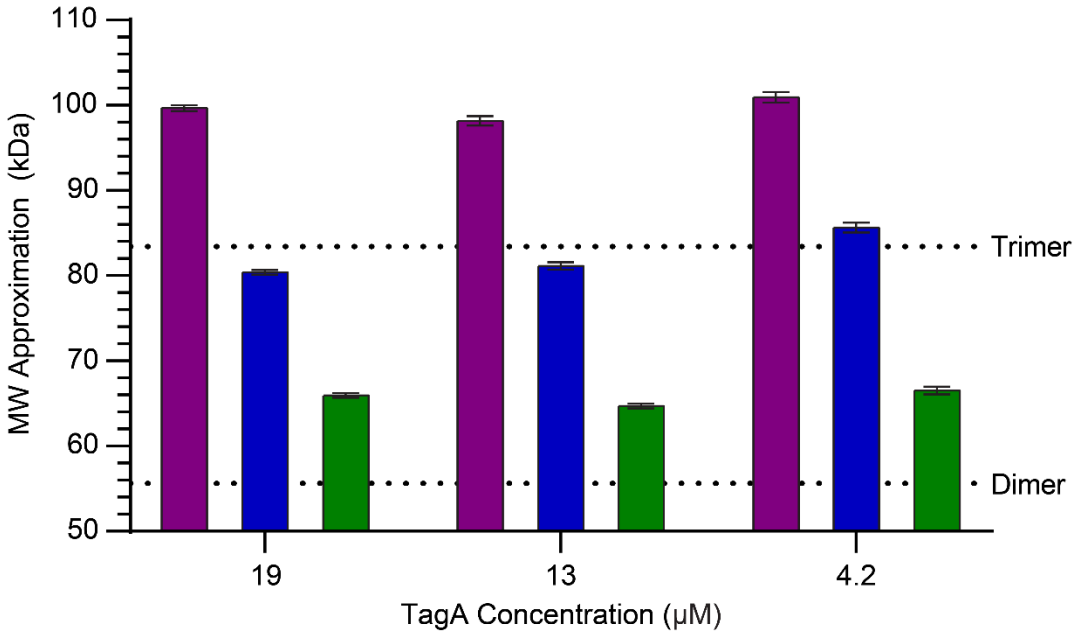

Figure S3.1 Solution state SE-AUC experiments on TagA. Single molecular weight approximations of TagA SE-AUC experiments at three concentrations (19 μM, 13 μM, and 4.2 μM) and three rotor speeds (12k rpm, purple, 15k rpm, blue, and 19k rpm, green). Averaged molecular weight species are concentration-independent and speed-dependent between experiments, indicating an aggregation of oligomers of indeterminable composition and quantity.

TagA residue pair [a]	Secondary structure elements between paired residues	Likelihood of residue pair C_β within 8 Å [b]
191-197	H9-(H9-H10 loop)	0.992
67-208	H4-(H10-H11 loop)	0.899
85-205	(H4-H5 loop)-H10	0.893
192-199	H9-(H9-H10 loop)	0.885
83-205	(H4-H5 loop)-H10	0.867
83-208	(H4-H5 loop)-(H10-H11 loop)	0.837
72-235	H4-H12	0.811
40-228	H2-H12	0.801
188-197	L17-(H9-H10 loop)	0.791
87-199	H5-(H9-H10 loop)	0.787
40-231	H2-H12	0.784
44-227	H2-H12	0.773
40-227	H2-H12	0.726
44-231	H2-H12	0.718
43-231	H2-H12	0.717
113-196	H6-(H9-H10 loop)	0.715
85-210	(H4-H5 loop)-H11	0.711
75-235	H4-H12	0.697
41-227	H2-H12	0.695
44-230	H2-H12	0.655
83-206	(H4-H5 loop)-(H10-H11 loop)	0.65
85-200	(H4-H5 loop)-(H9-H10 loop)	0.603
72-231	H4-H12	0.577
192-198	H9-(H9-H10 loop)	0.567
71-208	H4-(H10-H11 loop)	0.565
116-196	H6-(H9-H10 loop)	0.56
68-232	H4-H12	0.557
47-234	H2-H12	0.545
68-231	H4-H12	0.525
67-210	H4-H11	0.522
68-208	H4-(H10-H11 loop)	0.508
72-232	H4-H12	0.5

[a] Residue numbering according to *T. italicus* TagA. Only residue pairs between the core domain (M1-G195) and CTT (K196-R244) are listed.

[b] Likelihood generated by the RaptorX server, cutoff at 0.5.

Table S3.1 Coevolutionary-based probability of paired core domain and CTT residues within 8 Å between Cβ coordinates.

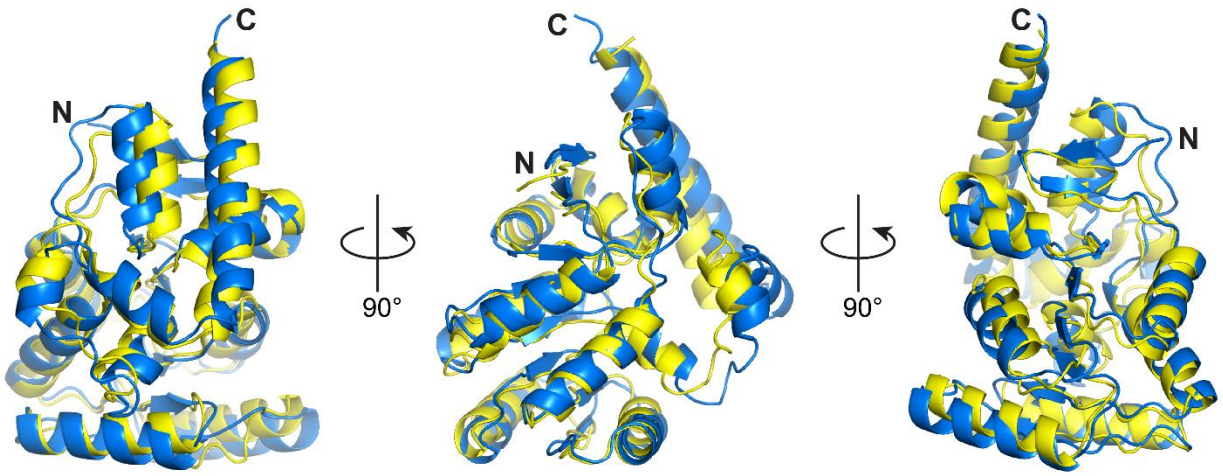

Figure S3.2 Computational models of full-length *T. italicus* TagA are structurally similar. The structure of full-length *T. italicus* TagA was modeled using either the GREMLIN (yellow) or RaptorX (blue) programs that leverage coevolution for structure prediction. Three views are shown that differ by 90^{o} rotations. Both programs predict similar structures that can be superimposed with a complete backbone RMSD of 2.05 Å. The core domains alone can be superimposed with a backbone RMSD of 1.71 Å, indicating slight variations in the positioning of the CTT in the modeling programs. The procedures used to model the structures are described in the text.

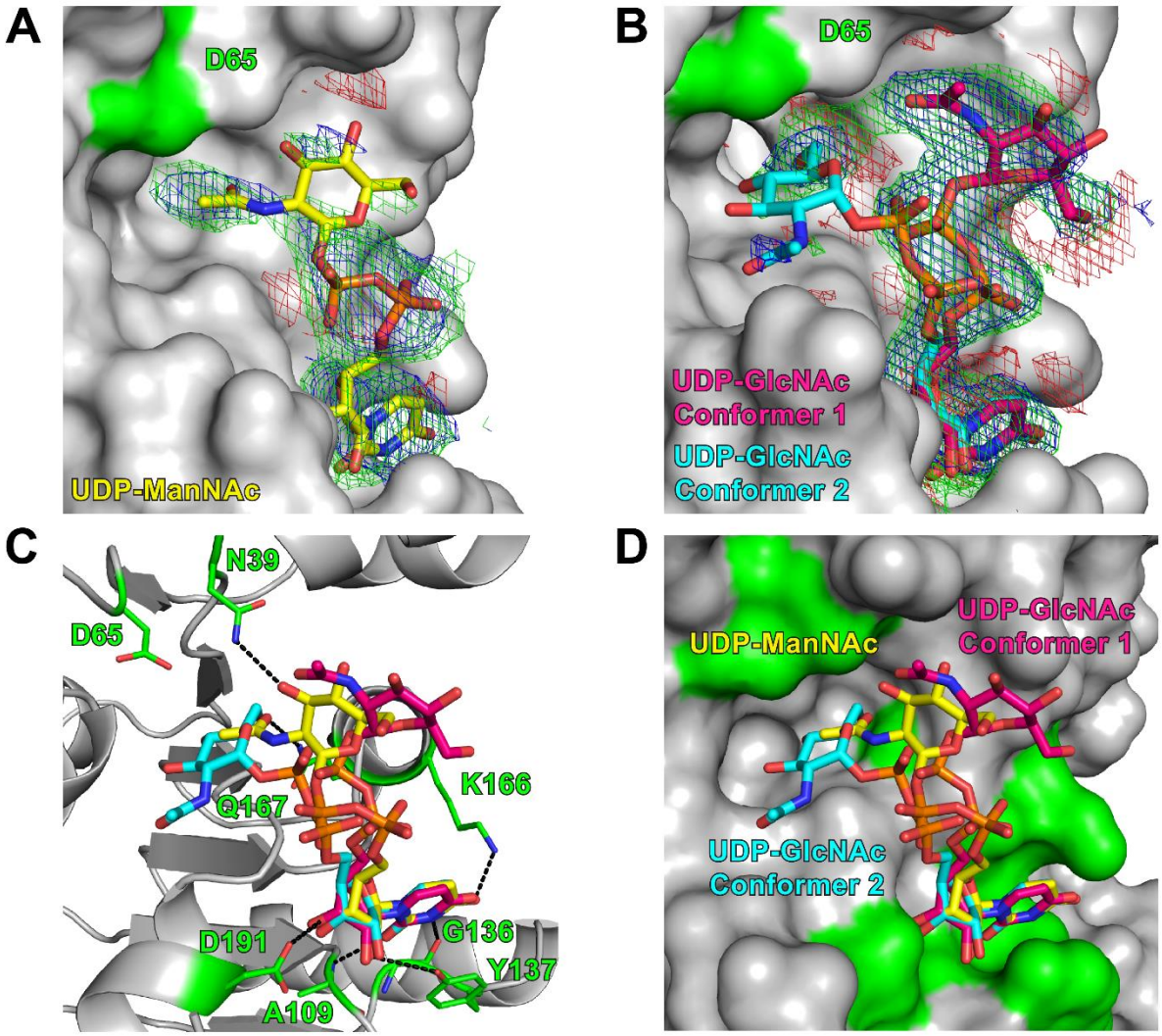

Figure S3.3 Comparison of crystallographic contacts between TagA and bound nucleotide sugars. *A-B*, $2mF_o$-DF_c and mF_o-DF_c simulated annealing omit maps for the ligands in the $TagA^{FL*}$:UDP-ManNAc (A) and the $TagA^{FL*}$:UDP-GlcNAc (B) structures. The $2mF_o$-DF_c omit maps (blue) are contoured at 1.0 σ and the mF_o-DF_c omit maps (green: positive, red: negative) are contoured at 2.0 σ. *C*, overlay of the crystallographic coordinates of UDP-ManNAc (yellow) and UDP-GlcNAc (pink, cyan) in their respective TagA complexes. Observed hydrogen bond contacts between core domain residues shown in stick format (green) and the UDP-ManNAc ligand are identified by black dashed lines. *D*, as in panel *C*, but the protein is shown in surface representation.

UDP-ManNAc adopts a single conformation (yellow) and UDP-GlcNAc adopts a major conformation (conformer 1, 60% abundance, pink) and a minor conformation (conformer 2, 40% abundance, cyan).

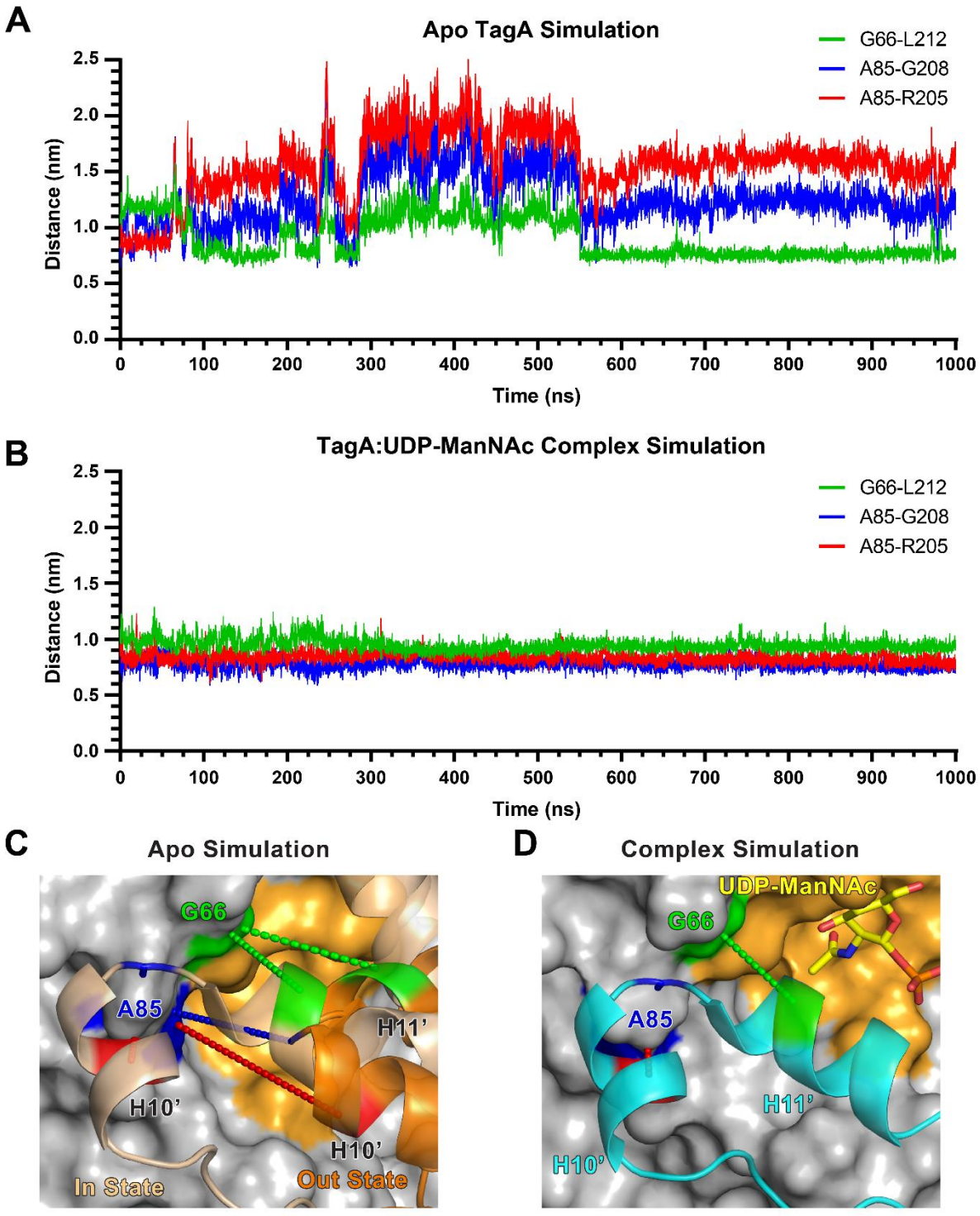

Figure S3.4 CTT and core domain interactions through apo and complex MD simulations. The distance between C_α atoms in three residues pairs were tracked through the MD simulations after equilibration. *A*, measurements between residues G66-L212 (green), A85-G208 (blue), and

V192-R205 (red) showed variable distances and highlighted the flexibility of H10' and H11' in the apo simulation. *B*, in the complex simulation, the same three residue pairs were tracked and displayed steady distance distributions through the simulation. Residue pairs were chosen for proximity between the CTT and core domain that would have a pronounced difference between the in and out states observed in the apo simulation. Representative C_α distances are displayed for the in and out states for the apo simulation (C) and for the most representative cluster of the ligand simulation (D). Colors of the residues and measurements are represented as in (A) and (B).

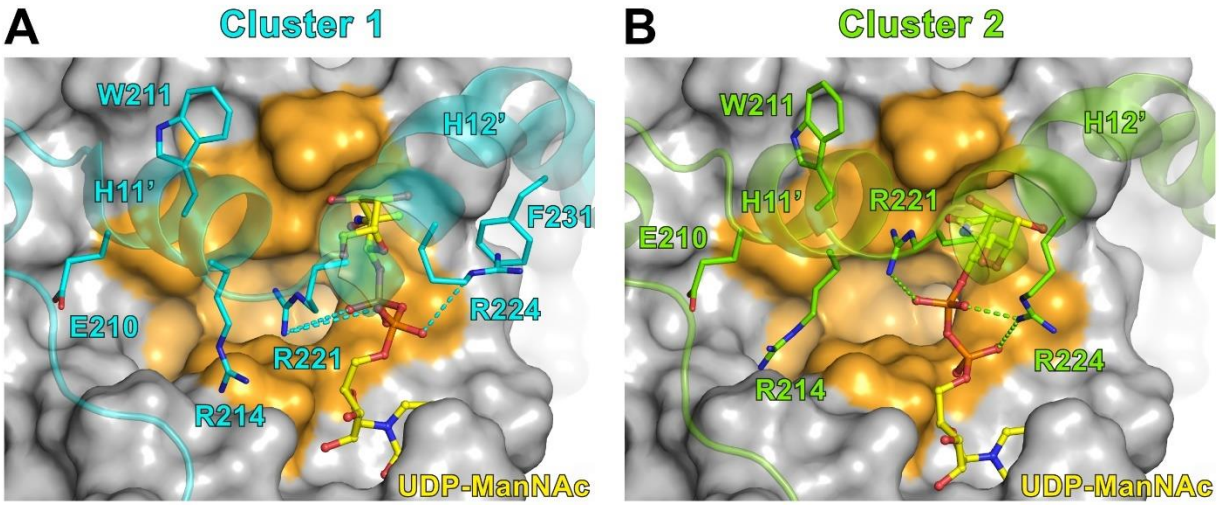

Figure S3.5 Complex MD simulations shows conserved CTT residues that project side chains into the catalytic pocket. *A*, cluster 1 (51% of the trajectory) from the $TagA^{CM}$:UDP-ManNAc complex simulation shows three conserved residues (E210, W211, and R214) orienting their side chains near the catalytic pocket (orange) and do not make any direct substrate. R221 and R224 are also highly conserved but directly coordinate the substrate pyrophosphate groups. A terminal amine group of R221 forms a stabilizing N-O bridge to the α- and β-phosphate oxygen atoms (4.7 Å). R224 appears to participate in a cation-π interaction with F231 and donate a weak hydrogen bond via its ε-amine to the α-phosphate oxygen (3.9 Å). *B*, Cluster 2 (25% of the trajectory) shows relatively similar positions for E210, W211 and R214, but R221 and R224 make stronger stabilizing contacts with the substrate. R221 makes a salt bridge only with a β-phosphate oxygen (3.1 Å) and R224 is close enough to make a salt bridge with α- and β-phosphate oxygen atoms (3.9 Å and 2.9 Å, respectively).

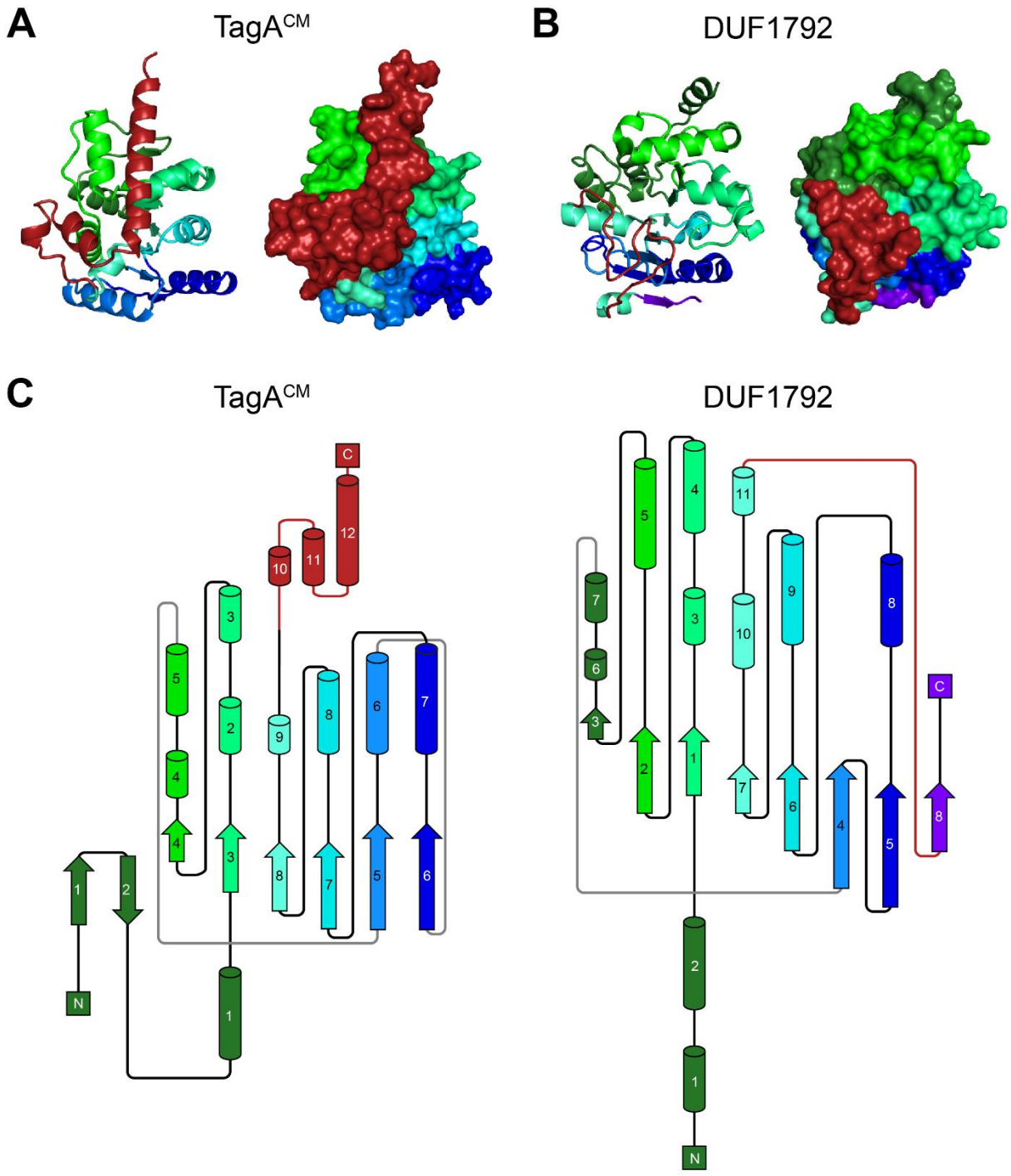

Figure S3.6 Structural comparison of known GT-E and GT-D folds. The GT-E fold (TagA^{CM}, A) and the GT-D fold (DUF1792, B) contain structural similarity in secondary structure topology (C) about the central β-sheet in each structure. The Rossmann-like domains are identically colored

shades of green, cyan, and blue based on topology between the two folds. TagA contains a flexible CTT composed of 3 helices to encapsulate the catalytic pocket whereas DUF 1792 has an extended loop region between H11-β8 to close its substrate binding site (colored maroon in both structures).

3.9 References

1. Formstone, A., Carballido-Lopez, R., Noirot, P., Errington, J., and Scheffers, D. J. (2008) Localization and interactions of teichoic acid synthetic enzymes in *Bacillus subtilis*. *J. Bacteriol*. **190**, 1812–1821
2. Brown, S., Santa Maria, J. P., and Walker, S. (2013) Wall teichoic acids of gram-positive bacteria. *Annu. Rev. Microbiol*. **67**, 313–336
3. Sewell, E. W. C., and Brown, E. D. (2014) Taking aim at wall teichoic acid synthesis: New biology and new leads for antibiotics. *J. Antibiot*. **67**, 43–51
4. Schade, J., and Weidenmaier, C. (2016) Cell wall glycopolymers of Firmicutes and their role as nonprotein adhesins. *FEBS Lett*. **590**, 3758–3771
5. Swoboda, J. G., Campbell, J., Meredith, T. C., and Walker, S. (2010) Wall teichoic acid function, biosynthesis, and inhibition. *Chembiochem* **11**, 35–45
6. Brown, S., Xia, G., Luhachack, L. G., Campbell, J., Meredith, T. C., Chen, C., Winstel, V., Gekeler, C., Irazoqui, J. E., Peschel, A., and Walker, S. (2012) Methicillin resistance in *Staphylococcus aureus* requires glycosylated wall teichoic acids. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 18909–18914
7. Farha, M. A., Leung, A., Sewell, E. W., D'Elia, M. A., Allison, S. E., Ejim, L., Pereira, P. M., Pinho, M. G., Wright, G. D., and Brown, E. D. (2013) Inhibition of WTA synthesis blocks the cooperative action of PBPs and sensitizes MRSA to β-lactams. *ACS Chem. Biol*. **8**, 226–233
8. Heckels, J. E., Lambert, P. A., and Baddiley, J. (1977) Binding of magnesium ions to cell walls of *Bacillus subtilis* W23 containing teichoic acid or teichuronic acid. *Biochem. J.* **162**, 359–365

9. Holland, L. M., Conlon, B., and O'Gara, J. P. (2011) Mutation of tagO reveals an essential role for wall teichoic acids in *Staphylococcus epidermidis* biofilm development. *Microbiology* **157**, 408–418

10. Misawa, Y., Kelley, K. A., Wang, X., Wang, L., Park, W. B., Birtel, J., Saslowsky, D., and Lee, J. C. (2015) *Staphylococcus aureus* colonization of the mouse gastrointestinal tract is modulated by wall teichoic acid, capsule, and surface proteins. *PLoS Pathog*. **11**, e1005061

11. Peschel, A., Vuong, C., Otto, M., and GöTz, F. (2000) The d-alanine residues of *Staphylococcus aureus* teichoic acids alter the susceptibility to vancomycin and the activity of autolytic enzymes. *Antimicrob. Agents Chemother*. **44**, 2845–2847

12. Schirner, K., Marles-Wright, J., Lewis, R. J., and Errington, J. (2009) Distinct and essential morphogenic functions for wall- and lipo-teichoic acids *in Bacillus subtilis*. *EMBO J*. **28**, 830–842

13. Wanner, S., Schade, J., Keinhörster, D., Weller, N., George, S. E., Kull, L., Bauer, J., Grau, T., Winstel, V., Stoy, H., Kretschmer, D., Kolata, J., Wolz, C., Bröker, B. M., and Weidenmaier, C. (2017) Wall teichoic acids mediate increased virulence in *Staphylococcus aureus*. *Nat. Microbiol*. **2**, 16257

14. Kojima, N., Araki, Y., and Ito, E. (1985) Structure of the linkage units between ribitol teichoic acids and peptidoglycan. *J. Bacteriol*. **161**, 299–306

15. Neuhaus, F. C., and Baddiley, J. (2003) A continuum of anionic charge: Structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiol. Mol. Biol. Rev*. **67**, 686–723

16. Yokoyama, K., Miyashita, T., Araki, Y., and Ito, E. (1986) Structure and functions of linkage unit intermediates in the biosynthesis of ribitol teichoic acids in *Staphylococcus aureus* H and *Bacillus subtilis* W23. *Eur. J. Biochem*. **161**, 479–489

17. Soldo, B., Karamata, D., and Lazarevic, V. (2002) tagO is involved in the synthesis of all anionic cell-wall polymers in *Bacillus subtilis* 168. *Microbiology* **148**, 2079–2087

18. Ginsberg, C., Zhang, Y.-H., Yuan, Y., and Walker, S. (2006) In vitro reconstitution of two essential steps in wall teichoic acid biosynthesis. *ACS Chem. Biol*. **1**, 25–28

19. Zhang, Y. H., Ginsberg, C., Yuan, Y., and Walker, S. (2006) Acceptor substrate selectivity and kinetic mechanism of *Bacillus subtilis* TagA. *Biochemistry* **45**, 10895–10904

20. D'Elia, M. A., Henderson, J. A., Beveridge, T. J., Heinrichs, D. E., and Brown, E. D. (2009) The N-acetylmannosamine transferase catalyzes the first committed step of teichoic acid assembly in *Bacillus subtilis* and *Staphylococcus aureus*. *J. Bacteriol*. **191**, 4030–4034

21. Bhavsar, A. P., Truant, R., and Brown, E. D. (2005) The TagB protein in *Bacillus subtilis* 168 is an intracellular peripheral membrane protein that can incorporate glycerol phosphate onto a membrane-bound acceptor in vitro. *J. Biol. Chem*. **280**, 36691–36700

22. Endl, J., Seidl, H. P., Fiedler, F., and Schleider, K. H. (1983) Chemical composition and structure of cell wall teichoic acids of staphylococci. *Arch. Microbiol*. **135**, 215–223

23. Naumova, I. B., Shashkov, A. S., Tul'Skaya, E. M., Streshinskaya, G. M., Kozlova, Y. I., Potekhina, N. V., Evtushenko, L. I., and Stackebrandt, E. (2001) Cell wall teichoic acids: Structural diversity, species specificity in the genus *Nocardiopsis*, and chemotaxonomic perspective. *FEMS Microbiol. Rev*. **25**, 269–283

24. Xia, G., Maier, L., Sanchez-Carballo, P., Li, M., Otto, M., Holst, O., and Peschel, A. (2010) Glycosylation of wall teichoic acid in *Staphylococcus aureus* by TarM. *J. Biol. Chem.* **285**, 13405–13415

25. Chen, L., Hou, W.-T., Fan, T., Liu, B., Pan, T., Li, Y.-H., Jiang, Y.-L., Wen, W., Chen, Z.-P., Sun, L., Zhou, C.-Z., and Chen, Y. (2020) Cryo-electron microscopy structure and transport mechanism of a wall teichoic acid ABC transporter. *mBio* **11**, e02749-19

26. Schirner, K., Stone, L. K., and Walker, S. (2011) ABC transporters required for export of wall teichoic acids do not discriminate between different main chain polymers. *ACS Chem. Biol*. **6**, 407–412

27. Gale, R. T., Li, F. K. K., Sun, T., Strynadka, N. C. J., and Brown, E. D. (2017) *B. subtilis* LytR-CpsA-Psr enzymes transfer wall teichoic acids from authentic lipid-linked substrates to mature peptidoglycan *in vitro*. *Cell Chem. Biol*. **24**, 1537–1546.e4

28. Campbell, J., Singh, A. K., Santa Maria, J. P., Kim, Y., Brown, S., Swoboda, J. G., Mylonakis, E., Wilkinson, B. J., and Walker, S. (2011) Synthetic lethal compound combinations reveal a fundamental connection between wall teichoic acid and peptidoglycan biosyntheses in *Staphylococcus aureus*. *ACS Chem. Biol*. **6**, 106–116

29. Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., and Bateman, A. (2016) The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res*. **44**, D279–D285

30. Haft, R. F., Wessels, M. R., Mebane, M. F., Conaty, N., and Rubens, C. E. (1996) Characterization of cpsF and its product CMP-N-acetylneuraminic acid synthetase, a group

B streptococcal enzyme that can function in K1 capsular polysaccharide biosynthesis in *Escherichia coli*. *Mol. Microbiol*. **19**, 555–563

31. Rahman, A., Barr, K., and Rick, P. D. (2001) Identification of the structural gene for the TDP-Fuc4NAc:Lipid II Fuc4NAc transferase involved in synthesis of enterobacterial common antigen in *Escherichia coli* K-12. *J. Bacteriol*. **183**, 6509–6516

32. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The carbohydrate-active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res*. **37**, D233–D238

33. Liu, J., and Mushegian, A. (2003) Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci*. **12**, 1418–1431

34. Campbell, J. A., Davies, G. J., Bulone, V., and Henrissat, B. (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J*. **326**, 929–939

35. Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B. (2003) An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol*. **328**, 307–317

36. Kattke, M. D., Gosschalk, J. E., Martinez, O. E., Kumar, G., Gale, R. T., Cascio, D., Sawaya, M. R., Philips, M., Brown, E. D., and Clubb, R. T. (2019) Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria. *PLoS Pathog*. **15**, e1007723

37. [dataset] Kattke, M. D., Gosschalk, J. E., Martinez, O. E., Kumar, G., Gale, R. T., Cascio, D., Sawaya, M. R., Philips, M., Brown, E. D., and Clubb, R. T. (2019) Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria. *Protein Data Bank*, 5WB4 and 5WFG

38. McRorie, D. K., and Voelker, P. J. (2004) *Self-Associating Systems in the Analytical Ultracentrifuge*. Beckman Instruments, Inc, Brea, California

39. Ovchinnikov, S., Kamisetty, H., and Baker, D. (2014) Robust and accurate prediction of residue–residue interactions across protein interfaces using evolutionary information. *Elife* **3**, e02030

40. Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738

41. Wang, S., Sun, S., Li, Z., Zhang, R., and Xu, J. (2017) Accurate de novo prediction of protein contact map by ultra-deep learning model. *PLoS Comput. Biol.* **13**, e1005324

42. Xu, J. (2019) Distance-based protein folding powered by deep learning. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 16856–16865

43. Strong, M., Sawaya, M. R., Wang, S., Phillips, M., Cascio, D., and Eisenberg, D. (2006) Toward the structural genomics of complexes: Crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 8060–8065

44. Duarte, J. M., Srebniak, A., Schärer, M. A., and Capitani, G. (2012) Protein interface classification by evolutionary analysis. *BMC Bioinformatics* **13**, 334

45. Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Adams, P. D., Read, R. J., Zwart, P. H., and Hung, L.-W. (2008) Iterative-build OMIT maps: Map improvement by iterative model building and refinement without model bias. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 515–524

46. Webb, B., and Sali, A. (2016) Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics* **54**, 5.6.1–5.6.37

47. Daura, X., Gademann, K., Jaun, B., Seebach, D., Gunsteren, W. F.v., and Mark, A. E. (2004) Peptide folding: When simulation meets experiment. *Angew. Chem. Int. Ed. Engl*. **38**, 236–240

48. Laganowsky, A., Reading, E., Hopper, J. T. S., and Robinson, C. V. (2013) Mass spectrometry of intact membrane protein complexes. *Nat. Protoc*. **8**, 639–651

49. Le Maire, M., Champeil, P., and Møller, J. V. (2000) Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim. Biophys. Acta Biomembr*. **1508**, 86–111

50. Privé, G. G. (2009) Lipopeptide detergents for membrane protein studies. *Curr. Opin. Struct. Biol*. **19**, 379–385

51. Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem*. **77**, 521–555

52. Zhang, H., Zhu, F., Yang, T., Ding, L., Zhou, M., Li, J., Haslam, S. M., Dell, A., Erlandsen, H., and Wu, H. (2014) The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold. *Nat. Commun*. **5**, 4339

53. Zhang, H., Zhou, M., Yang, T., Haslam, S. M., Dell, A., and Wu, H. (2016) New helical binding domain mediates a glycosyltransferase activity of a bifunctional protein. *J. Biol. Chem*. **291**, 22106–22117

54. Holm, L. (2020) DALI and the persistence of protein shape. *Protein Sci*. **29**, 128–140

55. [dataset] Zhang, H., Zhu, F., Yang, T., Ding, L., Zhou, M., Li, J., Haslam, S. M., Dell, A., Erlandsen, H., and Wu, H. (2014) The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold. *Protein Data Bank*, 4PHR

56. Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* **20**, 71–75

57. Friedhoff, P., Gimadutdinow, O., Ruter, T., Wende, W., Urbanke, C., Thole, H., and Pingoud, A. (1994) A procedure for renaturation and purification of the extracellular *Serratia marcescens* nuclease from genetically engineered *Escherichia coli*. *Protein Expr. Purif.* **5**, 37–43

58. Salzmann, M., Wider, G., Pervushin, K., Senn, H., and Wüthrich, K. (1999) TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. *J. Am. Chem. Soc*. **121**, 844–848

59. Eletsky, A., Kienhöfer, A., and Pervushin, K. (2001) TROSY NMR with partially deuterated proteins. *J. Biomol. NMR* **20**, 177–180

60. Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293

61. Keller, R. L. J. (2004) *The Computer Aided Resonance Assignment Tutorial*, 1st ed, CANTINA Verlag, Goldau, Switzerland

62. Shen, Y., and Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241

63. Lee, D., Hilty, C., Wider, G., and Wüthrich, K. (2006) Effective rotational correlation times of proteins from NMR relaxation interference. *J. Magn. Reson*. **178**, 72–76

64. Robson, S. A., Dag, Ç., Wu, H., and Ziarek, J. J. (2021) TRACT revisited: An algebraic solution for determining overall rotational correlation times from cross-correlated relaxation rates. *J. Biomol. NMR* **75**, 293–302

65. Hoogland, C., Gattiker, A., Duvaud, S. E., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) *Protein Identification and Analysis Tools on the ExPASy Server*. Humana Press, Inc, Totowa, NJ: 571–607

66. Becerra, S. P., Kumar, A., Lewis, M. S., Widen, S. G., Abbotts, J., Karawya, E. M., Hughes, S. H., Shiloach, J., and Wilson, S. H. (1991) Protein-protein interactions of HIV-1 reverse transcriptase: Implication of central and C-terminal regions in subunit binding. *Biochemistry* **30**, 11707–11719

67. Ross, P. D., Howard, F. B., and Lewis, M. S. (1991) Thermodynamics of antiparallel hairpin-double helix equilibria in DNA oligonucleotides from equilibrium ultracentrifugation. *Biochemistry* **30**, 6269–6275

68. Muthana, M. M., Qu, J., Li, Y., Zhang, L., Yu, H., Ding, L., Malekan, H., and Chen, X. (2012) Efficient one-pot multienzyme synthesis of UDP-sugars using a promiscuous UDP-sugar pyrophosphorylase from *Bifidobacterium longum* (BLUSP). *Chem. Commun*. **48**, 2728

69. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr*. **66**, 125–132

70. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr*. **40**, 658–674

71. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., et al. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol*. **75**, 861–877

72. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr*. **60**, 2126–2132

73. Smart, O. S., Womack, T. O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., and Bricogne, G. (2012) Exploiting structure similarity in refinement: Automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D Biol. Crystallogr*. **68**, 368–380

74. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25

75. Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., De Groot, B. L., Grubmüller, H., and Mackerell, A. D. (2017) CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73

76. Soldo, B., Lazarevic, V., Pooley, H. M., and Karamata, D. (2002) Characterization of a *Bacillus subtilis* thermosensitive teichoic acid-deficient mutant: Gene mnaA (yvyH) encodes the UDP-N-acetylglucosamine 2-epimerase. *J. Bacteriol*. **184**, 4316–4320

77. Marty, M. T., Baldwin, A. J., Marklund, E. G., Hochberg, G. K. A., Benesch, J. L. P., and Robinson, C. V. (2015) Bayesian deconvolution of mass and ion mobility spectra: From binary interactions to polydisperse ensembles. *Anal. Chem*. **87**, 4370–4376

Chapter 4

Progress towards understanding how monomeric TagA controls WTA display and cell morphology

4.1 Overview

Staphylococcus aureus, *Bacillus subtilis*, and other Gram-positive bacteria in the *Firmicutes* phylum heavily decorate their peptidoglycan with wall teichoic acid (WTA) glycopolymers. The first committed step in WTA biosynthesis is catalyzed by the conserved glycosyltransferase, TagA. TagA is a promising drug target because its genetic elimination attenuates bacterial pathogenicity. Prior chapters have described structural and mechanistic studies of the TagA glycosyltransferase from *Thermoanaerobacter italicus*, a close homolog to *S. aureus* and *B. subtilis* TagA enzymes. In these structural studies, the C-terminal tail of TagA was not visualized and its role in catalysis was only inferred from computational modeling and molecular dynamics simulations. In this chapter, I describe unpublished progress towards elucidating the structure of the full-length monomeric form of the TagA protein. In addition, complementation studies of the *S. aureus tagA* gene in a *B. subtilis* 168 Δ*tagA* background (a strain called *pr*Xyl) are presented that conclusively demonstrate the role of TagA in controlling cell morphology as observed by negative-stain transmission electron microscopy.

This chapter includes unpublished work that will contribute to future publications from the Clubb laboratory. My contributions to this work included: testing and analyzing the effect of the UDP-GlcNAc titration into $^{Ti}\text{TagA}^{FL*}$ by NMR, performing and analyzing SE-AUC experiments, characterizing the effect of xylose on the growth of the genetically-engineered *pr*Xyl *B. subtilis* 168 strain, and examining the cell morphology defects of *pr*Xyl in response to xylose supplementation by transmission electron microscopy. Other contributors to this work include Jason Gosschalk, who constructed the *pr*Xyl *B. subtilis* 168 strain, and Dr. Brendan Mahoney, who sequence specifically assigned the chemical shifts of $^{Ti}\text{TagA}^{FL*}$ using solution-state NMR spectroscopy.

4.2 Introduction

Many pathogenic Gram-positive bacteria from the *Firmicutes* phylum, like *Staphylococcus aureus*, have a thick peptidoglycan (PG) sacculus that encapsulates the cell. This PG layer is heavily functionalized with additional proteins and glycopolymers that contribute to the bacterium's physiological properties, including maintaining the cell's morphology, mediating host-pathogen interactions, and enabling antibiotic resistance (1-5). Among the most abundant components of the Gram-positive bacterial cell wall are wall teichoic acids (WTAs): peptidoglycan-anchored alditol-phosphate polymers that have essential functions in cell morphology regulation, host immune system evasion, cationic antimicrobial peptide resistance, and biofilm formation (2,3,6-8). The WTA biosynthetic pathway has drawn significant interest in the development of antivirulence compounds because the genetic elimination of its early-stage synthetic enzymes re-sensitizes methicillin-resistant *S. aureus* (MRSA) to β-lactam antibiotics and attenuates virulence while preserving cell viability (9-12).

WTA glycopolymers consist of poly-ribitol or poly-glycerol phosphate repeat units that can be decorated with accessory sugars or amino acids (4,9,13). They are anchored to the cell wall by a conserved linkage unit that is composed of one to three glycerol-phosphate (Gro-P) units linked to an N-acetyl-D-mannosamine(ManNAc)-(β1,4)-N-acetyl-D-glucosamine(GlcNAc) disaccharide monophosphate moiety which is produced by the TagO and TagA enzymes (4,14,15). Mature WTA glycopolymers play several important roles in virulence, division, and morphology for the bacterial cell. The elimination of WTAs through the mutation or deletion of *tagO* or *tagA* genes results in reduced pathogenicity and abrogated biofilm development in mammalian hosts for *Staphylococcus* genus bacteria (2,3,8). Additionally, the same genetic deletions have major impacts on the architecture of the cell wall. WTA inhibition in methicillin-resistant *S. aureus*

(MRSA) mislocalizes penicillin binding proteins—responsible for PG crosslinking—and results in cells with an increased susceptibility to lysozyme and β-lactam antibiotics (10,16). Thus, disruption of the biosynthesis of WTA polymers may lead to novel therapeutic strategies to combat antibiotic resistant bacteria.

TagA catalyzes a metal ion-independent inverting GT reaction that directly transfers ManNAc to the C4 hydroxyl of the terminal GlcNAc in the lipid-α substrate via an S_N2-like direct displacement mechanism (17,18). We recently reported the construction of a solubility-enhanced variant of TagA from *Thermoanaerobacter italicus* ($^{Ti}TagA^{FL*}$) and its co-crystal structure with its native sugar donor substrate, UDP-ManNAc (19). The structure provides insight into the first binding step in the Bi-Bi mechanism (19). We also determined a second crystal structure of $^{Ti}TagA^{FL*}$ with UDP-GlcNAc, a C2 epimer of UDP-ManNAc which TagA cannot utilize in its GT reaction (19). However, both of these co-crystal structures lacked density for the C-terminal tail (CTT) of TagA, which is important for catalytic activity *in vitro* and membrane association *in vivo* (18,19). As a result, the role of the CTT could only be inferred from the results of molecular dynamics simulations and activity measurements.

A major problem that limited the study of the full-length protein is the propensity of the CTT to be proteolytically degraded. Additional major complications involve the formation of dimeric and trimeric forms of the enzyme in crystals via inter-subunit surfaces that coincided with the CTT binding site on the enzyme. In this chapter, I describe research designed to capture the monomeric form of TagA in which the active site is fully formed by CTT interactions with the surface of the core domain. I also detail the development of a xylose-inducible TagA expression system in a *B. subtilis* 168 strain (called *pr*Xyl) in which the endogenous $^{Bs}tagA$ gene and its native promoter have been eliminated and replaced with a xylose-inducible $^{Sa}tagA$ gene. I demonstrate

that the absence of native ^{Bs}TagA leads to severe defects in cell morphology and growth, which can be recovered by the expression of the ^{Sa}TagA enzyme. To my knowledge, this is the first time that the enzymatic activities of ^{Sa}TagA and ^{Bs}TagA have shown to be interchangeable in WTA biosynthesis. Additionally, this ^{Sa}TagA-producing *B. subtilis* 168 strain represents the first example in which the effects of ^{Sa}TagA activity on cell morphology can be directly studied and it serves as a platform for ongoing cell-based screening for small molecule ^{Sa}TagA inhibitors. Collectively, the research described in this chapter provides a foundation for the discovery of TagA-specific small molecule inhibitors that could be further developed into antibacterial therapeutics.

4.3 Results and Discussion

4.3.1 Efforts to capture a stable monomeric form of TagA

Using ligand binding to promote the monomeric form of ^{Ti}TagA.

Significant progress has been made towards understanding the structure and mechanism of TagA from *T. italicus*. The structure of the ^{Ti}TagA core domain (^{Ti}TagA$^{\Delta C}$) has been solved in its apo, UDP-ManNAc-bound, and UDP-GlcNAc-bound states, but the coordinates of the crucial CTT have yet to be visualized experimentally (18,19). The CTT is important for membrane association and has been predicted to project nonpolar residues that enable the enzyme to interact with the cell membrane. Previous studies demonstrated that the replacement of four of these nonpolar residues with polar amino acids in the *T. italicus* TagA enzyme (^{Ti}TagA^{FL*}) reduces the propensity to dimerize and increases protein solubility in solution. The improved solubility of ^{Ti}TagA^{FL*} enabled backbone amide signals for 75 of 244 residues (~31%) of the primary sequence

of ^{Ti}TagA^{FL*} to be sequence specifically assigned using triple-resonance solution-state NMR methods (Fig. 4.1A). However, similar to the native protein, NMR analysis of ^{Ti}TagA^{FL*} revealed that the enzyme likely exchanges between monomeric and dimeric states at protein concentrations greater than 200 μM (19). Moreover, NMR signals for the 49 amino acids in the CTT could not be assigned, presumably because the CTT is detached from the enzyme and structurally disordered.

We reasoned that substrate binding could stabilize the monomeric form of ^{Ti}TagA^{FL*} by facilitating noncovalent interactions with the CTT that position it on the enzyme surface over the active site. We therefore titrated UDP-GlcNAc into a sample of ^{15}N-^{Ti}TagA^{FL*} and achieved a molar excess of 32:1. However, the spectrum of the complex does not contain additional NMR signals that correspond to an ordered CTT nor a global reduction in the linewidths of observed signals that should accompany an increased population of monomeric protein (Fig. 4.1A). Small chemical shift perturbations (CSPs) in response to adding UDP-GlcNAc were observed, but mapping the CSPs onto the ^{Ti}TagA^{FL*}:UDP-GlcNAc crystal structure suggest that the ligand binding is nonspecific (Fig. 4.1B). Overall, UDP-GlcNAc binding to ^{Ti}TagA^{FL*} was not observed by NMR spectroscopy at concentrations greater than 200 μM. I conclude that when the protein is at a concentration of 200 μM, binding of UDP-GlcNAc to the catalytic pocket is prevented because the enzyme is primarily dimeric, such that protein-protein contacts at the dimeric interface occlude nucleotide sugar binding even when the ligand is in excess.

To investigate protein oligomerization at concentrations below 200 μM, we performed sedimentation equilibrium by analytical ultracentrifugation (SE-AUC) experiments using ^{Ti}TagA and ^{Ti}TagA^{FL*} in the presence of 5-fold molar excess UDP-GlcNAc. Data acquired from native ^{Ti}TagA with UDP-GlcNAc indicated that it existed as a heterogeneous mixture of aggregated states because the averaged molecular weight approximations were concentration-independent and

uniformly decreased as rotor speed increased (Fig. 4.1C). Results were consistent with native ^{Ti}TagA SE-AUC data in its apo-form, so the addition of UDP-GlcNAc did not stabilize the monomeric form of the wild type protein (19). Data collected for ^{Ti}TagA^{FL*} with UDP-GlcNAc was best represented by a monomer-dimer equilibrium with an apparent K_D of 160 ± 45 μM (Fig. 4.1D). This dimer dissociation constant is within the error of measurement of the K_D of apo-^{Ti}TagA^{FL*} (210 ± 50 μM), indicating again that binding of UDP-GlcNAc does not sufficiently stabilize the monomeric form of the protein (19).

The finding that UDP-GlcNAc does not increase the propensity of monomeric TagA is consistent with its binding not affecting CTT association with the core domain. The CTT element was absent in crystal structures of ^{Ti}TagA^{FL*} bound to UDP-GlcNAc (19). Also, in the structure of the complex, UDP-GlcNAc adopted two conformations in the substrate pocket. The ligand conformation with the highest occupancy (~60% abundance) was one in which the sugar moiety was oriented away from the catalytic pocket (19). Interestingly, modeling suggests that this conformation produces steric clashes that would prevent CTT association with the core domain, and thus occlude formation of the monomer. On the other hand, crystals of the enzyme bound to UDP-ManNAc show that the sugar occupies a single conformation that is expected to favor CTT association with the core domain through electrostatic interactions which bridge the two subdomains. Thus, further experiments that probe the effects of UDP-ManNAc on TagA oligomerization are warranted.

Inter-domain fusion approach to promote monomeric ^{Ti}TagA.

As an alternative route to stabilize the monomeric form of TagA, we constructed a novel ^{Ti}TagA fusion protein. Computational modeling and molecular dynamics simulations suggest that

the TagA CTT associates with the core domain at the same interface that is used to promote protein oligomerization (19). To stabilize the positioning of the CTT on the core domain, we integrated native ^{Ti}TagA within the β2-β3 loop of enhanced green fluorescent protein (EGFP) (Fig. 4.2A). This inter-domain fusion method is expected to hold the N- and C-termini close to one another—a conformation in which the CTT may be positioned over the presumed active site. The termini of TagA are held together by noncovalent interactions between EGFP subdomains that form the properly folded fluorescent protein (20). With the integration of the ^{Ti}TagA polypeptide in EGFP, the CTT should be restricted in its range of motion and held near the active site—assuming EGFP forms its canonical β-barrel. Thus, we should be able to tune the length of the linkers connecting the termini in ^{Ti}TagA to the subdomains of EGFP to stabilize the monomeric form of the enzyme.

We constructed fusion proteins containing soluble linkers composed of repeats with 2x glycine and 1x serine (Fig. 4.2B-C) (21). Fusion proteins containing longer linkers (≥9 amino acids) were unstable and degraded during overexpression in *Escherichia coli* cells. The shorter linker (3 amino acids) stably overexpresses the EGFP-^{Ti}TagA fusion when it contains a small ubiquitin-like modifier (SUMO) solubility tag. This protein could be extracted from the cell lysate in the presence of CHAPS micelles. However, a significant amount of overexpressed protein remained in the cell lysate, potentially due to the interactions between exposed nonpolar surface residues in the CTT and the insoluble cell membrane. Purification of the protein was challenging, as the N-terminal 6xHis tag did not bind the polypeptide to Co^{2+}-NTA resin and had to be purified by anion exchange chromatography. Additionally, site-specific proteolysis by ULP1 was inefficient. Future studies will need to further purify this protein to determine if it retains catalytic activity and EGFP fluorescence. Solution-state studies to probe its oligomerization must also be performed and followed by structural characterization if warranted.

Targeted mutagenesis to promote the monomeric form of ^{Sa}TagA.

^{Ti}TagA has high primary sequence similarity to ^{Sa}TagA. Both proteins are predicted to contain a 3-helix CTT that is required for enzyme activity and membrane targeting. Toward the goal of increasing its solubility in solution to facilitate structural studies of the active form of the TagA enzyme, we generated a computational model of full-length ^{Sa}TagA (^{Sa}TagA^{CM}) using AlphaFold2 that includes its CTT (C-terminal helices H10-H12) (Fig. 4.2D) (22). As in models of full-length ^{Ti}TagA, several surface exposed side chains in the CTT are nonpolar and presumably limit its solubility (Fig. 4.2E) (18). These nonpolar surfaces may also limit the solubility of the protein in solution, so we implemented series of variants to alter the hydrophobic properties of the CTT and improve the solubility of ^{Sa}TagA. Guided by the ^{Ti}TagA^{FL*} model, we generated ^{Sa}TagA^{FL*}, which contains four amino acid substitutions that replace nonpolar surface exposed residues with hydrophilic residues (I211E/I217Q/I220K/L224E). ^{Sa}TagA^{FL*} was recombinantly overexpressed in *E. coli* cells and purified as described previously for ^{Ti}TagA^{FL*} (19). Purified ^{Sa}TagA^{FL*} exhibited superior solubility in aqueous buffer as compared to native ^{Sa}TagA and was primarily monomeric when assessed by SEC (only a small proportion of dimeric species was observed). However, the incorporation of the four amino acid substitutions in ^{Sa}TagA^{FL*} compromised its stability, as the protein formed insoluble aggregates at salt concentrations below 0.5 M and it degrades soon after purification (half-life of approximately 4 days).

In addition to ^{Sa}TagA^{FL*}, other variants of the enzyme were produced in which nonpolar CTT residues were replaced with hydrophilic amino acids. Here, two pairs of residues were modified. One pair alters the nonpolar surface located at the carboxyl ends of helices H10 and H11 (L215Q and L224Q, ^{Sa}TagA^{CTT1}) and the other pair changes the nonpolar surface at the carboxyl end of H11 and the amino end of H12 (L224Q and I230E, ^{Sa}TagA^{CTT2}). ^{Sa}TagA^{CTT2} could not be

overexpressed in *E. coli* cells, suggesting that the alterations destabilize the stability of the protein. In contrast, ^{Sa}TagA^{CTT1} could be stably overexpressed in *E. coli* cells and is a promising candidate for further study. The strategies mentioned here cover multiple methods to improve the interactions of the CTT with solvent and may inform on constructs to build upon and generate a solubility-enhanced variant for the structural analysis of full-length ^{Sa}TagA. These results also represent a step towards structure-guided development of ^{Sa}TagA-specific inhibitors that disrupt the synthesis of WTA virulence factors in pathogenic *S. aureus*, which may have been unattainable in the ^{Ti}TagA system.

4.3.2 ^{Sa}TagA and ^{Bs}TagA enzymes are functionally interchangeable in *B. subtilis* 168

Deleting the *tagO* and *tagA* genes in the *B. subtilis* 168 and *S. aureus* genomes slows their bacterial growth (11,12). Interestingly, the effect of these deletions on the ultrastructure of each type of bacterium is distinct. *S. aureus*, naturally cocci bacteria, do not have a significant morphological defect when the *tagA* gene is deleted (11). In *B. subtilis* 168, however, the deletion of the *tagO* or *tagA* genes causes cells to assume clumped, spherical shapes rather than their native rod-shaped morphology (11,12). This stark phenotypic shift in *B. subtilis* permits cytological profiling in response to WTA production.

We aimed to leverage the morphology change in *B. subtilis* to develop a cell-based assay to screen for small molecule inhibitors of the *S. aureus* TagA enzyme (^{Sa}TagA), which is a potential target for antivirulence therapeutic development. Toward this goal, we sought to determine if the ^{Sa}TagA and ^{Bs}TagA enzymes are functionally interchangeable in *B. subtilis* 168, such that the

morphological changes in Δ^{Bs}*tagA B. subtilis* strains complemented with ^{Sa}*tagA* could be used for inhibitor screening.

Construction of the xylose-inducible ^{Sa}*tagA* complementation strain of *B. subtilis* 168 (*pr*Xyl).

In wild type *B. subtilis* 168, the *tagA*, *tagB*, and *tagC* genes are under the control of a single promoter, *pr*^{Bs}*tagABC*. We constructed a strain in which the *S. aureus tagA* (^{Sa}*tagA*) gene under a xylose-inducible promoter (*pr*Xylose) replaces the native *tagA* (^{Bs}*tagA*) and the *pr*^{Bs}*tagABC* promoter (Fig. 4.3, Table 4.2). The strain is called *pr*Xyl and in it the *tagB* and *tagC* genes are unaffected, but their expression is also xylose-inducible. The strain was constructed by Jason Gosschalk using the pSWEET vector, which contains a promoter-operator system which regulates gene transcription downstream of the xylose-inducible P_{XylA} promoter (23). The plasmid was integrated into the genome by a double recombination event at regions flanking the *tagB* gene and the sequence prior to the native *pr*^{BS}*tagABC* promoter. Wild type ^{Sa}*tagA* was incorporated downstream of the Shine-Dalgarno sequence that follows P_{XylA} to enable xylose-inducible transcription. Chloramphenicol and spectinomycin resistance cassettes were both included in the plasmid used to engineer *pr*Xyl. Chloramphenicol resistance was incorporated within the integration regions, along with ^{Sa}*tagA* and the machinery required for the xylose-inducible system, to enable antibiotic selection for successful double recombination events. Spectinomycin resistance was included outside of the integration sequences in the plasmid to counter select against single Campbell-like crossover events. Successful incorporation of ^{Sa}*tagA* under the xylose-inducible promoter in *B. subtilis* 168 was verified by sequencing.

Growth characterization of the *pr*Xyl *B. subtilis* 168 complementation strain.

We compared *B. subtilis* growth rates in response to xylose in the growth media to validate complementation of ^{Bs}*tagA* with ^{Sa}*tagA* in *pr*Xyl. We presumed that if the activity of ^{Sa}TagA could recapitulate WTA biosynthesis, then the growth of *pr*Xyl would be slowed in the absence of xylose because WTAs are not produced. In these growth experiments, pre-cultures were grown to their stationary phase and washed with phosphate buffer solution (PBS) before induction of the experimental cultures at a low cell density (measured optical density at 600 nm (Abs_{600nm}) of ~0.05). Three concentrations of xylose in the growth media were used to investigate the effects of low growth (0% xylose), intermediate growth (0.03% xylose), and wild type growth (0.25% xylose) on cell physiology (Fig. S4.1). Indeed, the growth curve of *pr*Xyl at high xylose concentrations (0.25%) is comparable to *B. subtilis* 168—albeit with a slightly lower cell density in the stationary phase. This suggests that ^{Sa}TagA is enzymatically active in *B. subtilis* 168 cells and recovers WTA biosynthesis (Fig. S4.1B-C).

To further substantiate that ^{Sa}TagA expression was induced by adding xylose to the growth cultures and not by residual intracellular xylose present in the inoculant cultures, we repeated the above growth experiments using different types of inoculant cultures. In these experiments, pre-cultures were grown in media containing either low (0.05%), high (1%), or no (0%) xylose (Fig. S4.1). Few deviations were observed for growths originating from the 0.05% and 1% xylose pre-cultures, but a significant lag was observed for growths that began with the 0% xylose pre-culture. After 25 hours, however, cell densities in the stationary phase were consistent with growths using other inoculants. The only exception was that the intermediate cultures with 0.03% xylose had two distinct growth characteristics with high variability between them. This duality reflected a growth

or no growth phenotype where some *B. subtilis* cells recover from lacking xylose in the inoculum and grow at a rate similar to cells that have abundant access to xylose, albeit with a longer lag phase. To investigate this phenomenon at a cellular level, we subjected these cell cultures to transmission electron microscopy so that we could observe the phenotypic effect of xylose incorporation on overall cell morphology.

4.3.3 ^{Sa}*tagA* in *pr*Xyl recovers cell morphology in response to xylose

Cell morphology characterization of the *pr*Xyl *B. subtilis* 168 complementation strain.

B. subtilis 168 strains in which the *tagO* or *tagA* genes are deleted show marked morphological effects in response to WTA depletion (11,12). Therefore, the *pr*Xyl strain should exhibit large xylose-dependent morphological changes as xylose controls the amount of functioning WTA that is produced when ^{Sa}TagA is expressed. To investigate this relation, we first used transmission electron microscopy (TEM) with negative-stained cells to study the morphologies of wild type *B. subtilis* 168 and Δ^{Bs}*tagA* strains. As expected, *B. subtilis* 168 cells generally formed long end-to-end strings of rod-like bacteria, with several separated pairs of cells (Fig. 4.4A, Fig. S4.2A). To quantify the distribution of cell morphologies observed by TEM, the length and width of cells were measured, and the length/width (L/W) ratios were determined (Fig. 4.4B,D, Fig. S4.2H). Wild type cells were generally about 3.5 μm in length and 1.0 μm in width, in agreement with previous literature values for measurements of *B. subtilis* cells (24-26). Δ^{Bs}*tagA* bacteria displayed abnormal morphologies through a loss of their rod-shaped architecture, aberrantly divided cells, and asymmetric cell wall ultrastructures (Fig. 4.4A, Fig. S4.2B). Additionally, cells were often found in clumps of closely associated and mis-divided cells despite rigorous mechanical disruption of the cell solutions before their application to the electron

microscopy grid. Due to the nature of these cells and lack of a clear division septum, we were unable to select and quantitatively measure individual Δ^{Bs}*tagA* cells.

The *pr*Xyl cells exhibited xylose-dependent morphological changes and became more rod-shaped when increasing amounts of xylose was present in the growth media. *pr*Xyl cells divided unequally when no xylose was included, which resulted in several large, ellipsoid daughter cells that lack typical rod-shaped morphologies (Fig. 4.4C, Fig. S4.2C,H). At high xylose concentrations (≥0.25% xylose), the cells generally regained their rod shapes, divided into equally sized daughter cells, and were symmetric in their overall architecture (Fig. 4.4C, Fig. S4.2E-F). Interestingly, at the intermediate xylose concentration (0.03% xylose), a distribution of morphologies was observed. Cells appeared to adopt rod-like, curved, or spherical cell morphologies (Fig. 4.4C, Fig. S4.2D). We expected a relatively uniform distribution of cells shaped like short, stunted rods based on growth curve data from the 0.05% and 1% pre-cultures. However, we observed a wide assortment of cell shapes that suggest an inconsistency in xylose-induced transcription of ^{Sa}*tagA*, where some cells appear to produce TagA (and WTAs) at high enough quantities to completely recover their rod-shaped morphology and others fail to produce WTAs and maintain their mis-divided, spherical cells. As a negative control for WTA-devoid cells, we explored the effects of adding tunicamycin, a TagO-specific inhibitor, to a *pr*Xyl culture containing 0.25% xylose. The resulting cell morphologies were similar to Δ^{Bs}*tagA*, showing clusters of round and asymmetric cells (Fig. 4.4A,C, Fig. S4.2G). The morphology-based analysis was consistent with the trends observed in the growth curves of *pr*Xyl with xylose and further support that the xylose-induced expression of the nonnative ^{Sa}TagA is capable of producing WTA glycopolymers in *B. subtilis*.

4.4 Conclusion

This work begins to establish a structural and cellular framework to search for small molecule inhibitors of the ^{Sa}TagA enzyme that may be useful in treating infections caused by *S. aureus*. In particular, promising variants of ^{Sa}TagA that alter its CTT may be useful in elucidating the monomeric form of the protein in the presence and absence of small molecule inhibitors. These solubility-enhanced constructs of ^{Sa}TagA may also be used for structure-function studies to gain new insight into the catalytic mechanism of TagA. The construction and characterization of the xylose-inducible ^{Sa}*tagA*-complemented *B. subtilis* 168 strain (*pr*Xyl) also provides a foundation for a cell morphology-based screening of ^{Sa}TagA-specific inhibitors. These small molecule screening efforts are ongoing and will be completed by other members of the Clubb laboratory.

4.5 Materials and Methods

4.5.1 Protein expression and purification

All TagA constructs were cloned and expressed using standard methods as described previously (18,19). Briefly, ^{Ti}TagA and ^{Sa}TagA constructs contained an N-terminal 6xHis-tag and tobacco etch virus (TEV) protease recognition sequence (ENLYFQS) in the pMAPLe4 expression vector. EGFP-TagA fusion constructs contained an N-terminal small ubiquitin-like modifier (SUMO) solubility domain with a 6xHis-tag in the pSUMO expression vector. Protein was expressed in *E. coli* BL21(DE3) cells. Cultures were grown in Miller LB Broth (Thermo Fisher) in the presence of 50 μg/mL kanamycin (Thermo Fisher) at 37 °C in a shaking incubator to the late log growth phase (OD_{600nm} of 0.6-0.8), before induction with 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) (GoldBio). Expression proceeded while shaking at 18 °C for 16

hours. For isotopically labeled samples, the cell pellets were exchanged into M9 media supplemented with $^{15}NH_4Cl$, ^{13}C-glucose, and 70% deuterium oxide (Cambridge Isotope Laboratories) before induction (27). The cells were pelleted by centrifugation, and resuspended in 40 mM CHAPS (Thermo Fisher), 50 mM Tris–HCl, pH 7.5, 500 mM NaCl supplemented with 400 µl of protease inhibitor cocktail (Sigma-Aldrich), and 2 mM phenylmethanesulfonylfluoride, 1 mg egg white lysozyme, and 0.5 mg *Serratia marcescens* nuclease per liter of culture (28). Resuspended cells were lysed using an Emulsiflex high pressure homogenizer (Avestin) and soluble TagA lysate was clarified by centrifugation.

Protein was purified by passing over Co^{2+}-NTA HisPur resin (Thermo Fisher) and washed with 8 mM CHAPS, 50 mM Tris–HCl, pH 7.5, and 500 mM NaCl buffer. Protein was eluted from the resin using the same buffer but supplemented with 200 mM imidazole. Eluted protein was concentrated using Amicon centrifugal filters (Thermo Fisher) at 4 °C. The 6xHis-tags were removed proteolytically using 0.5 mg TEV protease (for pMAPLe4 proteins) or 1.0 mg ULP1 protease (for pSUMO proteins) and buffer exchanged by dialysis into a 10% v/v glycerol, 50 mM Tris–HCl, pH 7.5, and 500 mM NaCl buffer. The protein was passed over Co^{2+}-NTA again and washed off the column with 10% v/v glycerol, 50 mM Tris–HCl, pH 7.5, and 500 mM NaCl buffer. Wash fractions containing protein were concentrated, and the protein was subjected to size-exclusion chromatography using a Superdex 75 preparation grade column (GE). Protein was concentrated prior to storage or use.

4.5.2 NMR spectroscopy

Isotopically labeled samples were dissolved in NMR buffer at pH 6.8 (50 mM sodium phosphate, pH 6.8, and 200 mM NaCl). $^{Ti}TagA^{FL*}$ was uniformly isotopically labeled with ^{15}N or

$^{15}N/^{13}C/^{2}H$. UDP-GlcNAc (Sigma-Aldrich) was dissolved in NMR buffer to 20 mM and iteratively added to a $^{Ti}TagA^{FL*}$ protein sample for NMR experiments. TROSY-enhanced ^{15}N-HSQC experiments were performed for the UDP-GlcNAc titration, with increasing scans in the direct dimensions to account for the loss in signal-to-noise due to protein dilution. Spectra were acquired at 303 K on Bruker Avance III HD 600 MHz (14.1 T) and Bruker Avance NEO 800 MHz (18.8 T) spectrometers equipped with triple resonance cryogenic probes. Backbone assignments of $^{Ti}TagA^{FL*}$ were determined by carrying out TROSY-enhanced variants of the following experiments: ^{15}N-HSQC, HN(CO)CA, HNCA, HNCACB, and HN(CO)CACB (29,30). Data were processed using NMRPipe, and CARA was used to perform sequential assignment (31,32).

4.5.3 Analytical ultracentrifugation

SE–AUC experiments were performed on an Optima XL-A analytical ultracentrifuge (Beckman Coulter). Three concentrations for each sample ($^{Ti}TagA$: 5.3, 18, and 28 μM; $^{Ti}TagA^{FL*}$: 6.3, 20, and 30 μM) with 5-fold molar excess UDP-GlcNAc were subjected to multiple ultracentrifuge speeds ($^{Ti}TagA$: 12,000, 15,000, 19,000 rpm, and 24,000 rpm; $^{Ti}TagA^{FL*}$: 15,000, 19,000, and 24,000 rpm) and allowed to reach SE at 4 °C. To avoid spectral absorbance of UDP-GlcNAc at 280 nm, data were collected at 292 nm. Data regression analysis was performed using the Beckman-Coulter Optima Analytical Ultracentrifuge Origin Data Analysis Package. The data were fit to multiexponential and single-exponential models. $^{Ti}TagA^{FL*}$ was best represented by a monomer–dimer multiexponential model that was calculated using the predicted monomeric molecular weight of 27,862 Da by the ExPASy ProtParam tool (33). The dissociation constant (K_D) was determined to be the inverse of $K_{A(conc)}$ using the following equation (34,35):

$$K_{A(conc)} = K_{A(abs)} ((\varepsilon l)^{n-1}/n)$$

where the molar extinction coefficient (ε) was determined to be 20,873 cm^{-1} M^{-1} at an absorbance wavelength of 292 nm using the ExPASy ProtParam tool, l is the path length of 1.2 cm, n is the order of oligomerization, and $K_{A(abs)}$ is the absorbance association constant.

4.5.4 *Bacillus subtilis* 168 cloning and growth

To construct the pXyl shuttle vector for integration into *B. subtilis* 168, ^{Sa}*tagA* was amplified from genomic *S. aureus* DNA and integrated downstream of the P_{XylA} promoter in pSWEET by Gibson assembly (23,36). Flanking DNA fragments upstream of the ^{Bs}*tagA* ribosome binding site (RBS) and downstream of ^{Bs}*tagA* were then amplified by PCR of genomic *B. subtilis* 168 DNA. The counter selection antibiotic resistance cassettes (spectinomycin and ampicillin) and vector DNA leading up to the *amyE* flanking regions in pDG1662 were amplified by PCR. Then, all fragments were assembled into the pXyl shuttle vector by Gibson assembly. pXyl plasmid assembly was confirmed by sequencing (Laragen Sequencing). Competent *B. subtilis* 168 cells were transformed as previously described with pXyl and double recombination events were selected by resistance to chloramphenicol (5 μg/mL) (Thermo Fisher) (37). Colonies with chloramphenicol resistance were then counter selected using spectinomycin (100 μg/mL) (Thermo Fisher), which would only be present for single crossover events and integration of the entire pXyl vector. Successful incorporation of pXyl into the *B. subtilis* 168 genome was confirmed by sequencing (Laragen Sequencing).

*pr*Xyl cells were streaked onto LB agar plates containing 5 μg/mL chloramphenicol and 1% w/v xylose (Thermo Fisher). Individual bacterial colonies were collected and grown in 25 mL pre-cultures containing 5 μg/mL chloramphenicol and 0%, 0.05%, or 1% w/v xylose for 16 hours at 37 °C while shaking. The cells were then pelleted by centrifugation, washed with phosphate

buffer solution (Thermo Fisher), and pelleted again to remove residual xylose from the pre-culture media. Cells were resuspended in LB and diluted to an Abs_{600nm} of 0.05 into 50 μL growth cultures containing 5 μg/mL chloramphenicol and 0%, 0.03%, or 0.25% w/v xylose in 384-well plate format. Cells were grown at 37 °C for 25 hours.

Several replicate wells were collected in the late log phase of growth for application to electron microscopy grids. Prior to gridding, the cells were pelleted by centrifugation at 5000 x g for 5 minutes at room temperature, resuspended with water, and pelleted again. The water wash was repeated twice and finally diluted to an OD_{600nm} of 0.6 using water. The cells were vigorously mixed before application to electron microscopy grids.

4.5.5 Transmission electron microscopy

Negative-stained grids containing *B. subtilis* cells were prepared for electron microscopy by placing 3 μL of the cell solution onto a glow discharged (90 seconds) 300-mesh Formvar-carbon coated Cu electron microscopy grid (Ted Pella, Inc.). Cells were incubated for 3 minutes, washed 3 times with 3 μL water, negative-stained with 3 μL 2% uranyl acetate, and washed 3 times with 3 μL water. Grids were imaged at room temperature using a Tecnai T12 TEM operating at 120 keV at 1,200x to 8,000x magnifications on a Gatan MSC SI003 (model 794) camera (1,024 x 1,024 pixels). Images were analyzed and cells were measured using ImageJ (38).

4.6 Figures

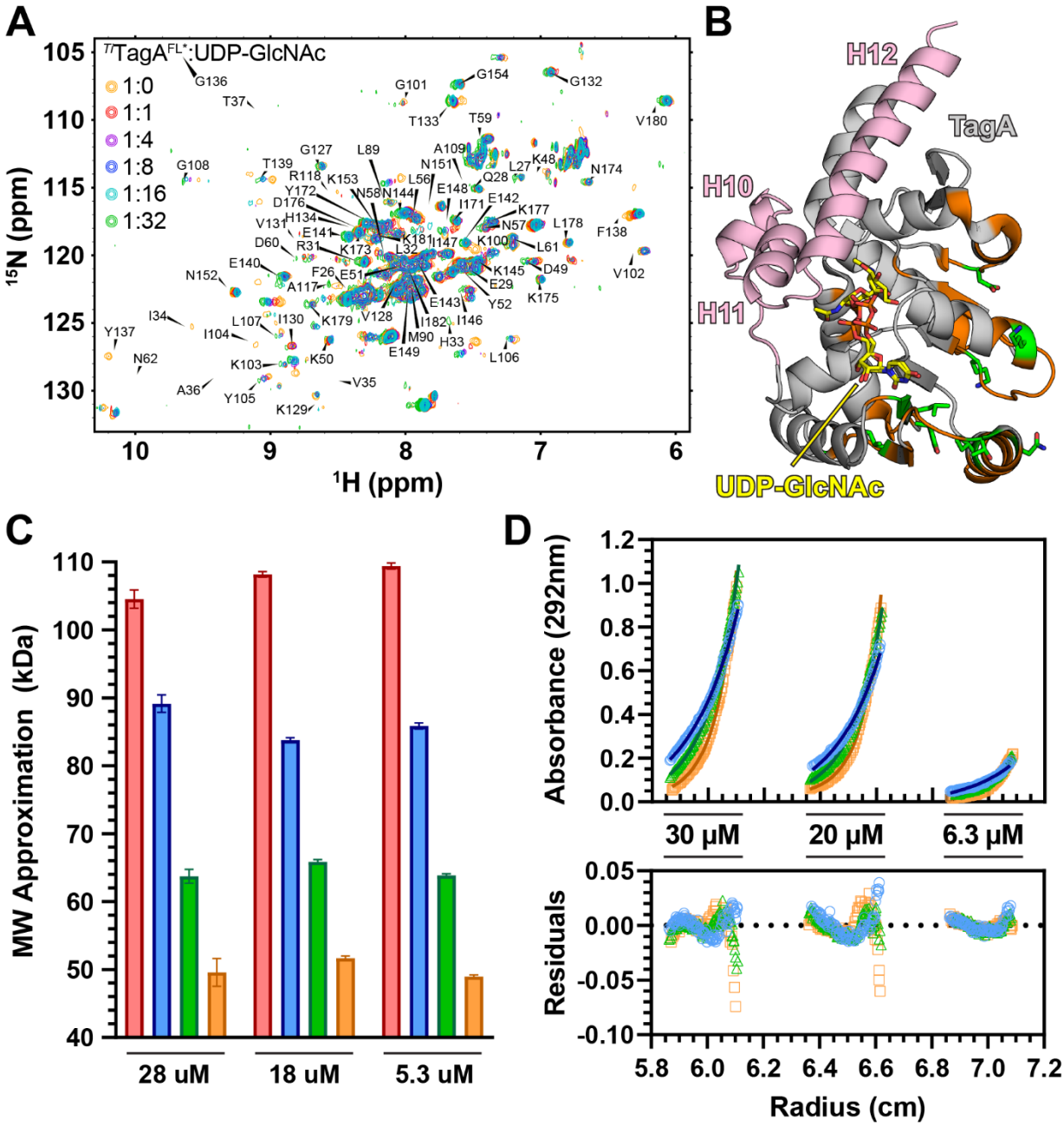

Figure 4.1 Efforts to stabilize monomeric full-length TagA through UDP-GlcNAc binding. (A) ^{1}H-^{15}N transverse relaxation optimized spectroscopy heteronuclear single quantum coherence (TROSY-HSQC) NMR spectra of ^{15}N labeled ^{Ti}TagA^{FL*} with increasing amounts of UDP-GlcNAc. Approximately 31% of amide residues (75 of 244) were assigned for apo-^{Ti}TagA^{FL*}

(labeled). Overlaid spectra are colored based on the concentration of UDP-GlcNAc in the sample (in molar equivalence to ^{Ti}TagA^{FL*} protein): 0x (orange), 1x (red), 4x (purple), 8x (blue), 16x (cyan), and 32x (green). (B) Chemical shift perturbations were observed for some of the assigned peaks and are displayed on the ^{Ti}TagA^{FL*}:UDP-GlcNAc co-crystal structure (gray) (PDB: 7MPK) with the computationally-modeled C-terminal tail (pink). Residues with peak assignments without chemical shift perturbations (orange), assigned residues with chemical shift perturbations in response to UDP-GlcNAc (green, stick format), and crystallographic UDP-GlcNAc conformations (yellow, stick format) are shown on the structure. (C) SE-AUC experiments of ^{Ti}TagA with 5-fold molar excess UDP-GlcNAc. Data were collected at three concentrations (5.3 μM, 18 μM, and 28 μM) and four rotor speeds (12k rpm – red, 15,000 rpm – blue, 19,000 rpm – green, and 24,000 rpm – orange). Single fit averaged molecular weight approximations are displayed and modeled concentration-independent and speed-dependent trends, indicating aggregated oligomers of indeterminable composition and quantity. (D) SE-AUC experiments of ^{Ti}TagA^{FL*} with 5-fold molar excess UDP-GlcNAc. Data were collected at three concentrations (6.3 μM, 20 μM, and 30 μM) and at three rotor speeds (15,000 rpm – blue, 19,000 rpm – green, and 24,000 rpm – orange). The data were best fit by a monomer-dimer equilibrium model. Residuals after fitting the data to a monomer-dimer equilibrium are shown below the plot of experimental data (symbols) and exponential fits (lines).

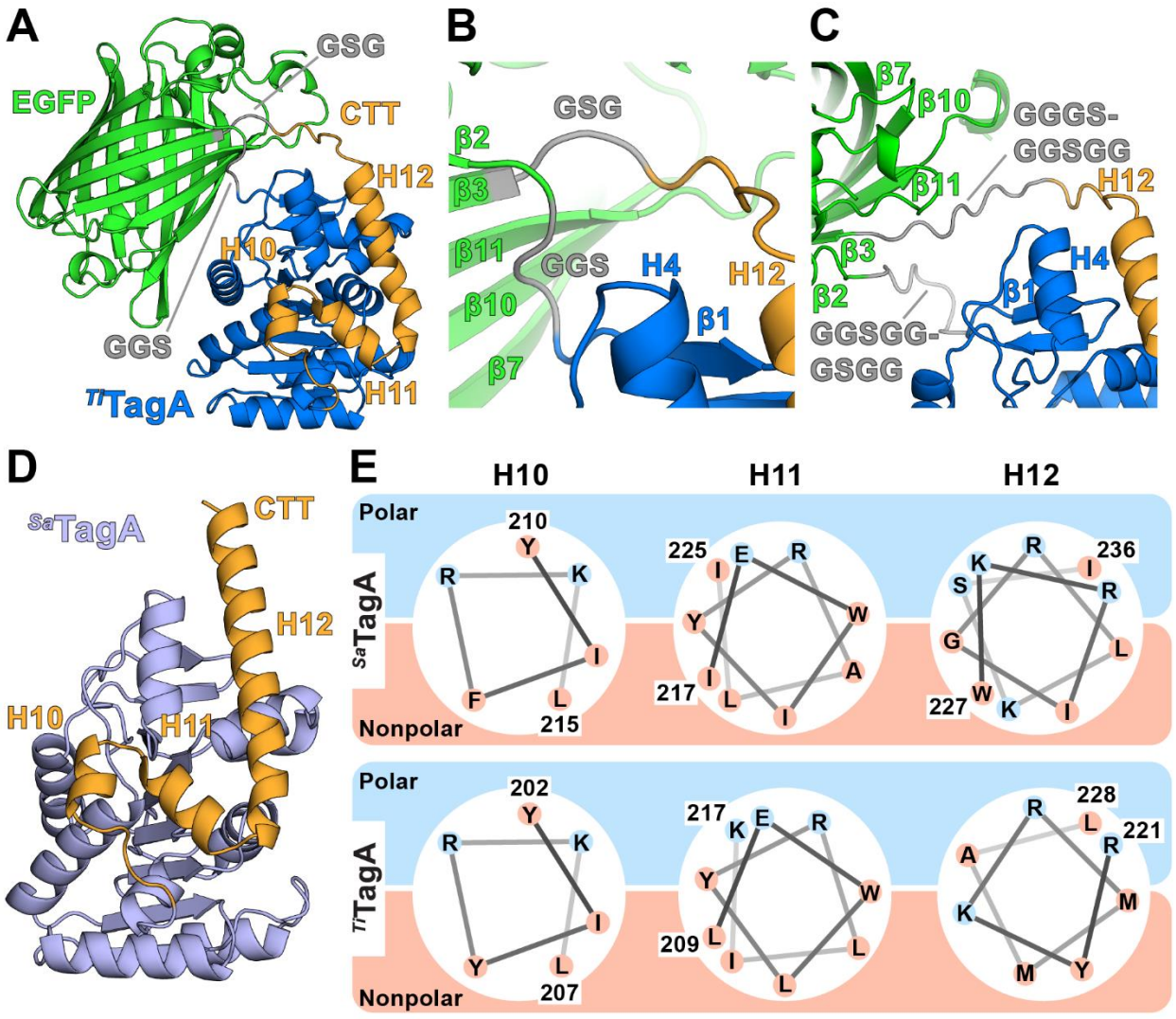

Figure 4.2 Efforts to capture the capture the monomeric form of TagA. (A) The hypothetical insertion of the *Ti*TagA domain into the β2-β3 loop region of enhanced green fluorescent protein (EGFP). The two domains fold independently of one another. EGFP forms a classical β-barrel encapsulating its chromophore (green) and TagA adopts its GT-E glycosyltransferase fold (blue) with its dynamic C-terminal tail (CTT, orange) fused to EGFP. Three amino acid linkers (gray) connect the N- and C-termini of TagA to the EGFP β2 and β3 sheets, respectively. The structure is predicted by AlphaFold2 (22). (B-C) Two different glycine-serine flexible linker lengths were selected for the TagA-EGFP fusion construct: three amino acids (GGS/GSG, B) and nine amino

acids (GGSGGGSGG/GGGSGGSGG, C). Colors are assigned as in (A). (D) The computational model of full-length ^{Sa}TagA predicted by AlphaFold2 (22). The TagA core domain (light blue) forms a classical Rossmann-like fold and the CTT (orange) adopts three amphipathic helices. (E) Helical wheel projections of the CTT helices (H10-H12) from computational models of ^{Ti}TagA and ^{Sa}TagA. The helices display a general amphipathicity, where polar and charged residues (blue) face towards the TagA core domain and nonpolar residues (red) are surface exposed. Helix H12 is projected to the conserved proline residue in both models.

Protein Construct	Organism of origin	Mutations	Location of mutations	Maximum observed solubility	Approximate Half Life
^{Ti}TagA	*Thermoanaerobacter italicus* Ab9	C111A	Core domain	497 μM	4 weeks
^{Ti}TagA$^{\Delta C}$	*Thermoanaerobacter italicus* Ab9	C111A, ΔG195	Core domain	2236 μM	1 year
^{Ti}TagA E218	*Thermoanaerobacter italicus* Ab9	C111A, ΔE218	Core domain	216 μM	Not monitored
^{Ti}TagA^{FL*}	*Thermoanaerobacter italicus* Ab9	C111A/I203E/L209Q /L212K/I216E	H10-H11	1060 μM	3 weeks
^{Ti}TagA D65A	*Thermoanaerobacter italicus* Ab9	C111A/D65A	Core domain (for activity)	326 μM	Not monitored
^{Ti}TagA E41A	*Thermoanaerobacter italicus* Ab9	C111A/E41A	Core domain (for activity)	231 μM	Not monitored
^{Ti}TagA N39A	*Thermoanaerobacter italicus* Ab9	C111A/N39A	Core domain (for activity)	96 μM	Not monitored
^{Ti}TagA G68C	*Thermoanaerobacter italicus* Ab9	C111A/G68C	Core domain (FRET)	773 μM	Not monitored
^{Ti}TagA^{FL*} G68C	*Thermoanaerobacter italicus* Ab9	C111A/G68C/I203E/ L209Q/L212K/I216E	Core domain and H10-H11 (FRET)	393 μM	Not monitored
^{Ti}TagA N30C	*Thermoanaerobacter italicus* Ab9	C111A/N30C	Core domain (FRET)	196 μM	Not monitored
^{Ti}TagA G127C	*Thermoanaerobacter italicus* Ab9	C111A/G127C	Core domain (FRET)	752 μM	Not monitored
^{Ti}TagA ROSIE/ GREMLIN (SW4E)	*Thermoanaerobacter italicus* Ab9	C111A/I69S/F71W/F 87E/V192E/L209E/L 212E	Core domain and H10-H11	435 μM	4 days
^{Ti}TagA Srf7	*Thermoanaerobacter italicus* Ab9	C111A/I203E/L207N /L209Q/L212K/L215 S/I216E/I233Q	H10-H12	194 μM	3 days
^{Ana}TagA	*Anaerostipes* sp. 992a	Native	N/A	9000 μM	4 weeks
^{Hao}TagA	*Halothermothrix orenii*	Native	N/A	Could not purify	N/A
^{Lba}TagA	*Lachnospiraceae bacterium* KH1T2	Native	N/A	Could not purify	N/A
^{Ros}TagA	*Roseburia* sp. CAG:197	Native	N/A	Could not purify	N/A
^{Tpa}TagA	*Treponema pallidum* (strain Nichols)	Native	N/A	Could not purify	N/A
^{Dip}TagA	*Diploscapter pachys*	Native	N/A	Could not purify	N/A
^{Sa}TagA	*Staphylococcus aureus*	Native	N/A	19 μM	Not monitored
^{Sa}TagA$^{\Delta C}$	*Staphylococcus aureus*	ΔA204	Core domain	450 μM	Not monitored
^{Sa}TagA^{FL*}	*Staphylococcus aureus*	I211E/I217Q/I220K/ L224E	H10-H11	773 μM	4 days
^{Bs}TagA^{FL*}	*Bacillus subtilis* 168	W202E/L208Q/F211 K/I215E	H10-H11	Could not purify	N/A

Table 4.1 TagA protein constructs and homologs used in these studies.

Strain	Genotype/Description	Source
Bs168	*B. subtilis* wild type *trpC2*	BGSC
Δ^{Bs}*tagA*	Bs168, Δ*tagA*	(11)
*pr*Xyl	Bs168, Δ*tagA*, Δ*prtagABC*, P_{XylA}:^{Sa}*tagA*:Cam	Current study

Plasmid	Genotype/Description	Source
pSWEET	pDG364 with *xyl* expression system, integrated at the *amyE* locus in the *B. subtilis* 168 genome	(23)
pXyl	pSWEET with ^{Sa}*tagA*, *cam* resistance cassette (for selection), integrated 3' of ^{Bs}*tagA* and 5' of *prtagABC*, *spc* and *bla* cassettes outside of the flanking regions for counter selection	Current study
pDG1662	*B. subtilis* vector, integrating at the *amyE* locus, *cam* resistance cassette (for selection), *spc* and *bla* cassettes outside of the flanking regions for counter selection	BGSC

Table 4.2 *Bacillus subtilis* strains and plasmids used in these studies.

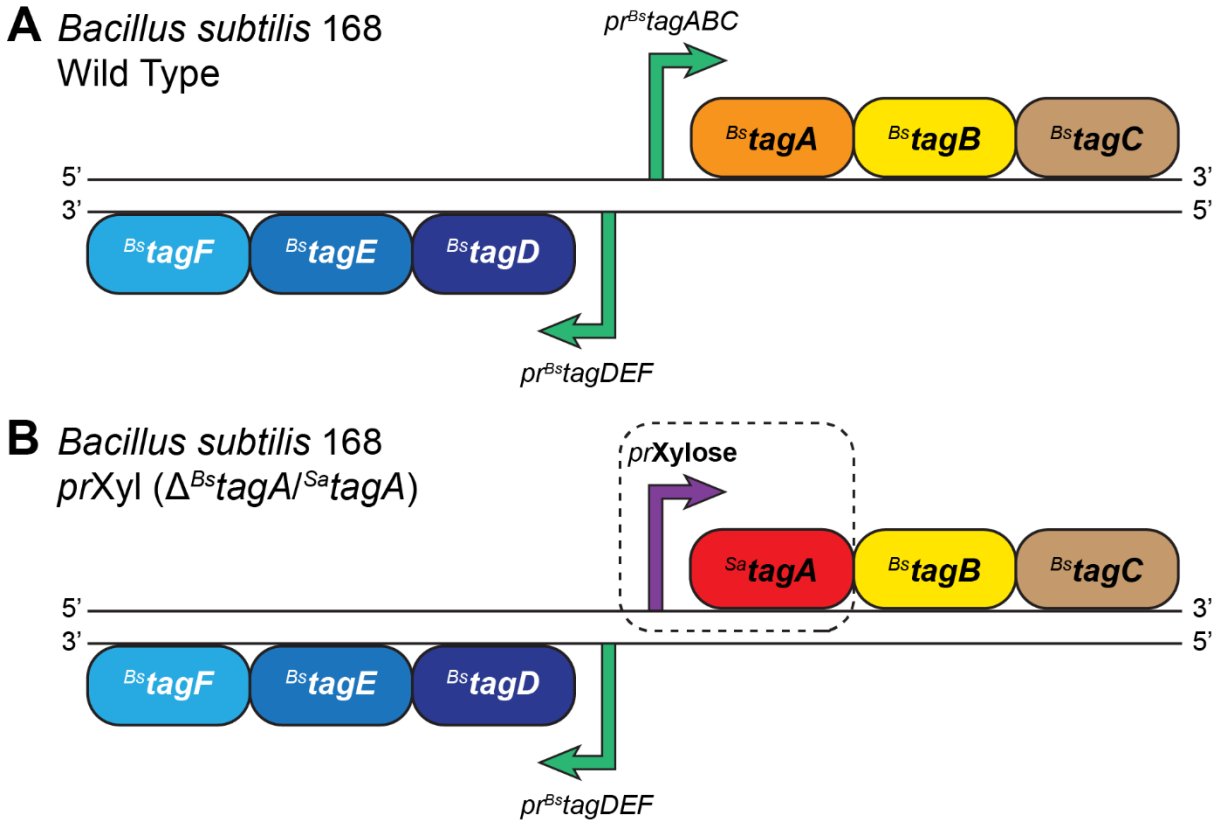

Figure 4.3 Assembly of a xylose-inducible ^{Sa}TagA strain of *B. subtilis* 168. (A) ^{Bs}*tagA* is natively controlled by the *pr*^{Bs}*tagABC* promoter, which also controls transcription of ^{Bs}*tagB* and ^{Bs}*tagC*. ^{Bs}*tagD*, ^{Bs}*tagE*, and ^{Bs}*tagF* are separately controlled by the *pr*^{Bs}*tagDEF* promoter on the antisense strand 3' of *pr*^{Bs}*tagABC*. (B) Double homologous recombination of pXyl with the ^{Bs}*tagA* locus of the *B. subtilis* 168 genome yields non-native ^{Sa}*tagA* from *Staphylococcus aureus* controlled by a xylose-inducible promoter (*pr*Xylose) from pSWEET in place of ^{Bs}*tagA* and the *pr*^{Bs}*tagABC* promoter, respectively.

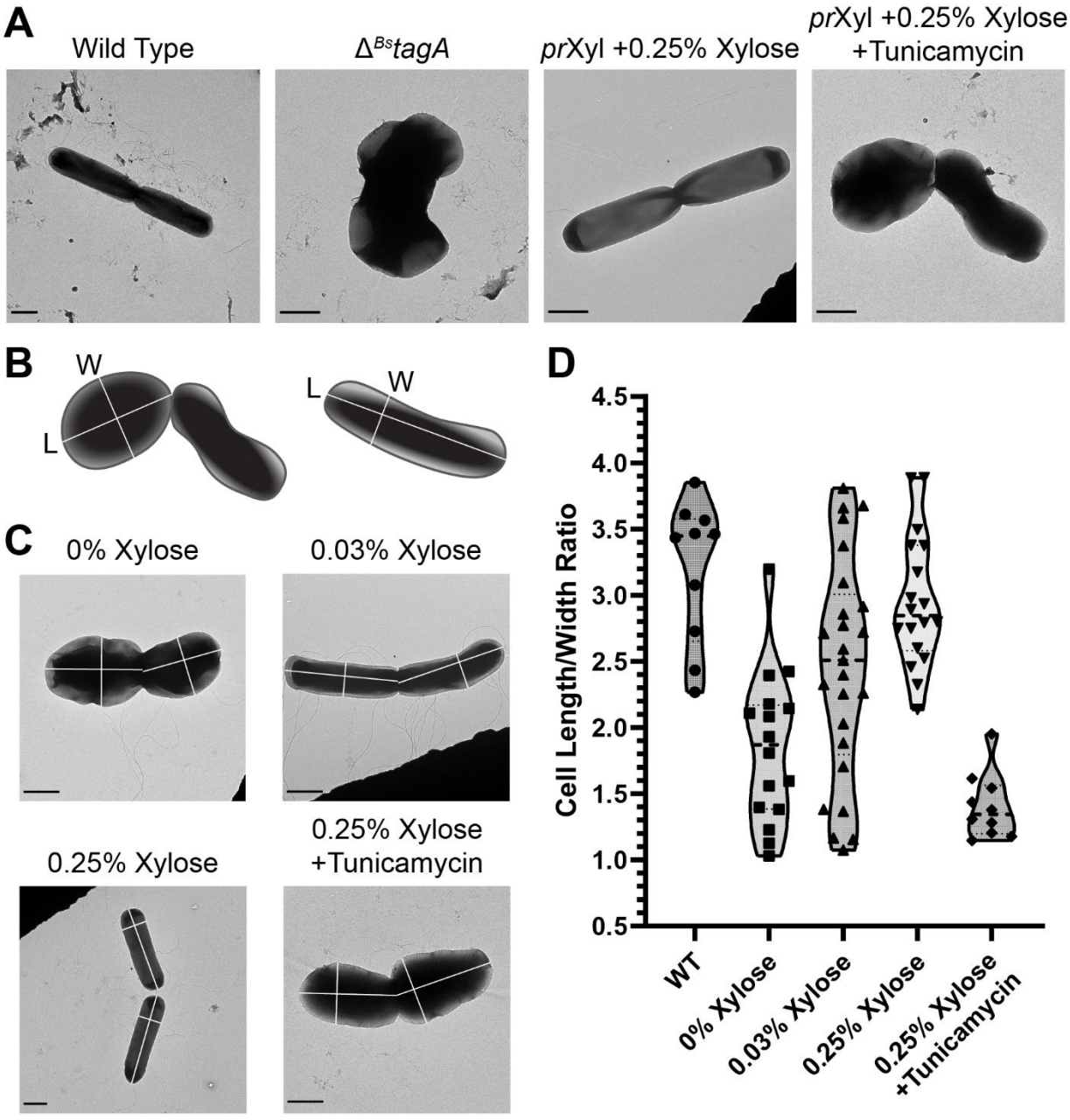

Figure 4.4 Ultrastructure of *B. subtilis* 168 changes in response to TagA expression. Bacteria were harvested at the late log phase of growth and applied to copper electron microscopy grids before negative staining. (A) Representative TEM micrographs of *Bacillus subtilis* 168 wild type, Δ^{Bs}*tagA*, *pr*Xyl grown with 0.25% xylose, and *pr*Xyl grown with 0.25% xylose and 2 μg/mL tunicamycin, a TagO-specific inhibitor. Scale bars represent 1 μm. (B) Examples of length and

width measurements of individual cells for quantitative analysis of cell morphology for individual bacteria or dividing cells with a complete division septum. Length was defined as the longest axis of cell growth. Width was defined as the longest axis of the cell perpendicular to the length axis. (C) Representative micrographs of *pr*Xyl grown with increasing amounts of xylose in the culture media. White lines represent the measured length and width axes of the cells. Scale bars represent 1 μm. (D) Individual cell length/width ratios for wild type *B. subtilis* 168 and *pr*Xyl cells. Violin plots are shown with individual data points. The median of each dataset is represented by a thick dashed line and quartiles are represented by thin dotted lines.

4.7 Supplemental Figures

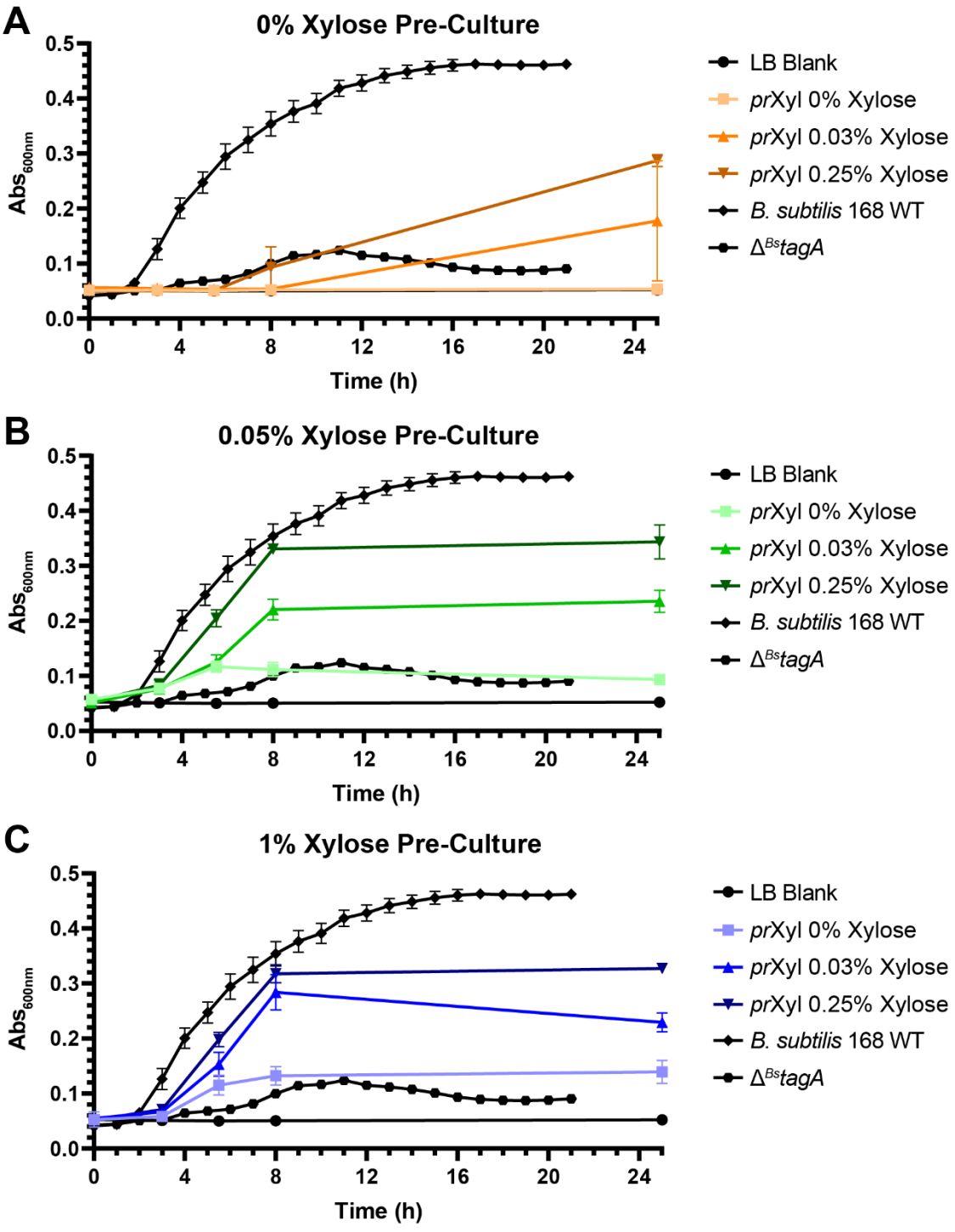

Figure S4.1 *pr*Xyl cell growth is responsive to xylose in culture media. *pr*Xyl *Bacillus subtilis* 168 was cultured in variable amounts of xylose using pre-cultures with different xylose

concentrations. (A) When no xylose was included in the pre-culture, the growth curves of *pr*Xyl stalled, but the 0.03% and 0.25% xylose cultures reached similar optical densities to other pre-culture experiments. The 0.03% growth condition had a high variability at 25h because all cultures had Abs_{600nm} readings >0.24 (62%) or <0.08 (38%), suggesting a xylose-dependent cellular switch that regulates growth. For 0% xylose in the growth media, the cells did not grow over 25h. (B) When 0.05% xylose was included in the pre-culture, cells had distinct growth patterns depending on the xylose concentration in the growth media. The cell density at stationary phase for 0%, 0.03%, and 0.25% xylose in the media was easily differentiated. (C) When 1% xylose was included in the pre-culture, cells grew similarly to (B). All growths were completed with ≥6 replicates.

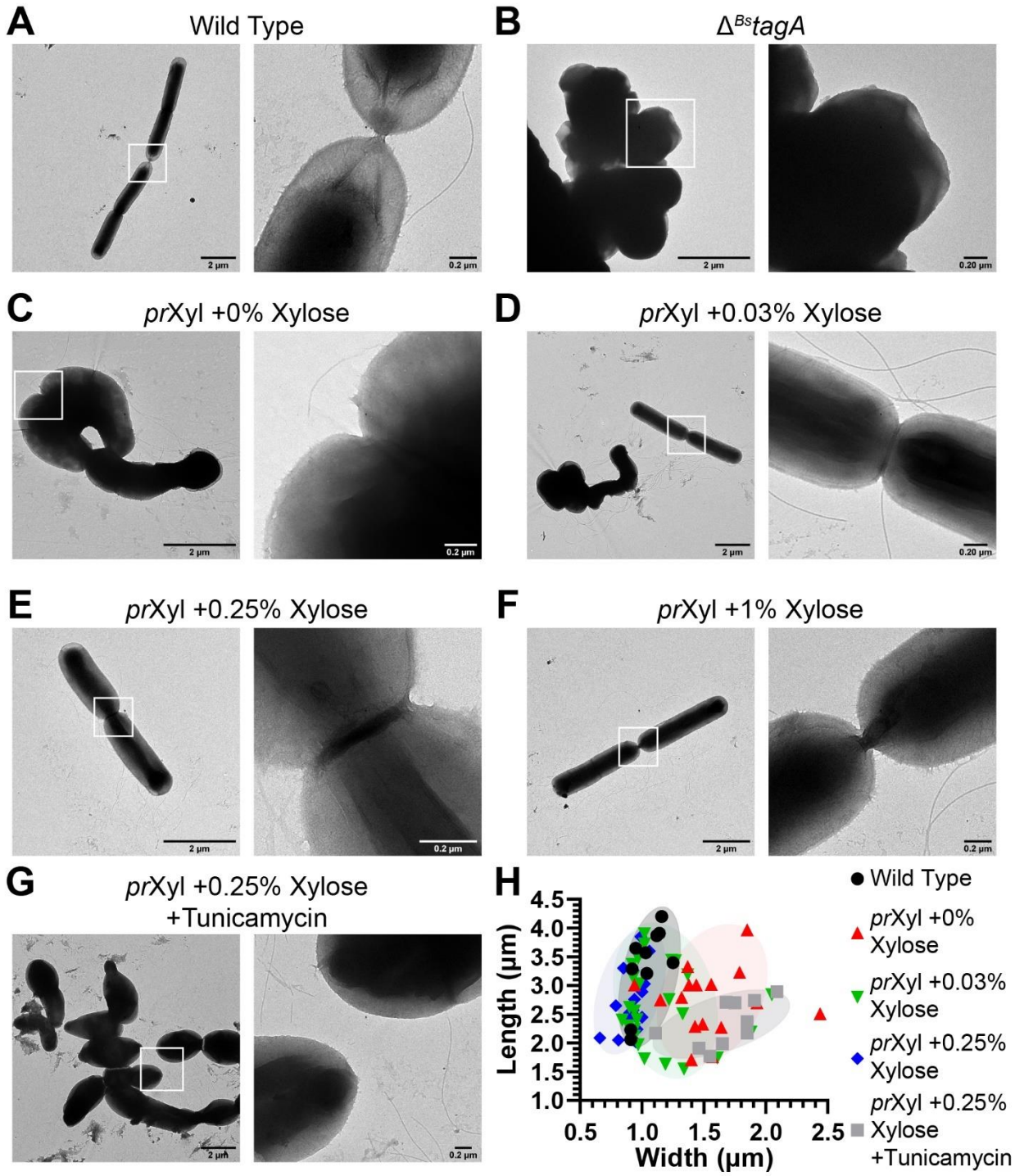

Figure S4.2 Representative micrographs of *B. subtilis* 168 strains. (A-G) All strains of *B. subtilis* 168 were grown and harvested in the late log phase before cells were applied to TEM grids and negative-stained. *pr*Xyl *B. subtilis* cells were grown in increasing concentrations of xylose

(0%, 0.03%, 0.25%, and 1%) in the growth media. Higher amounts of xylose recovered the rod-shaped morphology displayed by wild type *B. subtilis* 168. Inclusion of tunicamycin, a TagO-specific inhibitor, with a high concentration of xylose in the growth media exhibited morphological defects similar to *pr*Xyl with 0% xylose and Δ^{Bs}*tagA* cells. White boxes indicated in the images on the left represent the location of the images on the right for each panel. (H) Cell lengths and widths were measured for individual cells that could be identified by separation on a grid or the presence of a complete division septum indicating two mature daughter cells. The lengths were defined as the longest axis of the cell and the widths perpendicular to the length axis at the widest point of the cell.

4.8 References

1. Schade, J., and Weidenmaier, C. (2016) Cell wall glycopolymers of Firmicutes and their role as nonprotein adhesins. *FEBS Letters* **590**, 3758-3771
2. Holland, L. M., Conlon, B., and O'Gara, J. P. (2011) Mutation of tagO reveals an essential role for wall teichoic acids in Staphylococcus epidermidis biofilm development. *Microbiology* **157**, 408-418
3. Misawa, Y., Kelley, K. A., Wang, X., Wang, L., Park, W. B., Birtel, J., Saslowsky, D., and Lee, J. C. (2015) Staphylococcus aureus Colonization of the Mouse Gastrointestinal Tract Is Modulated by Wall Teichoic Acid, Capsule, and Surface Proteins. *PLOS Pathogens* **11**, e1005061
4. Brown, S., Santa Maria, J. P., and Walker, S. (2013) Wall Teichoic Acids of Gram-Positive Bacteria. *Annual Review of Microbiology* **67**, 313-336
5. Percy, M. G., and Gründling, A. (2014) Lipoteichoic Acid Synthesis and Function in Gram-Positive Bacteria. *Annual Review of Microbiology* **68**, 81-100
6. Heckels, J. E., Lambert, P. A., and Baddiley, J. (1977) Binding of magnesium ions to cell walls of Bacillus subtilis W23 containing teichoic acid or teichuronic acid. *Biochemical Journal* **162**, 359-365
7. Wanner, S., Schade, J., Keinhörster, D., Weller, N., George, S. E., Kull, L., Bauer, J., Grau, T., Winstel, V., Stoy, H., Kretschmer, D., Kolata, J., Wolz, C., Bröker, B. M., and Weidenmaier, C. (2017) Wall teichoic acids mediate increased virulence in Staphylococcus aureus. *Nature Microbiology* **2**, 16257-16257
8. Weidenmaier, C., Kokai-Kun, J. F., Kristian, S. A., Chanturiya, T., Kalbacher, H., Gross, M., Nicholson, G., Neumeister, B., Mond, J. J., and Peschel, A. (2004) Role of teichoic

acids in Staphylococcus aureus nasal colonization, a major risk factor in nosocomial infections. *Nature Medicine* **10**, 243-245

9. Brown, S., Xia, G., Luhachack, L. G., Campbell, J., Meredith, T. C., Chen, C., Winstel, V., Gekeler, C., Irazoqui, J. E., Peschel, A., and Walker, S. (2012) Methicillin resistance in Staphylococcus aureus requires glycosylated wall teichoic acids. *Proc Natl Acad Sci U S A* **109**, 18909-18914
10. Farha, M. A., Leung, A., Sewell, E. W., D'Elia, M. A., Allison, S. E., Ejim, L., Pereira, P. M., Pinho, M. G., Wright, G. D., and Brown, E. D. (2013) Inhibition of WTA Synthesis Blocks the Cooperative Action of PBPs and Sensitizes MRSA to β-Lactams. *ACS Chemical Biology* **8**, 226-233
11. D'Elia, M. A., Henderson, J. A., Beveridge, T. J., Heinrichs, D. E., and Brown, E. D. (2009) The N-acetylmannosamine transferase catalyzes the first committed step of teichoic acid assembly in Bacillus subtilis and Staphylococcus aureus. *J Bacteriol* **191**, 4030-4034
12. D'Elia, M. A., Millar, K. E., Beveridge, T. J., and Brown, E. D. (2006) Wall teichoic acid polymers are dispensable for cell viability in Bacillus subtilis. *J Bacteriol* **188**, 8313-8316
13. Neuhaus, F. C., and Baddiley, J. (2003) A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiology and molecular biology reviews : MMBR* **67**, 686-723
14. Kojima, N., Araki, Y., and Ito, E. (1985) Structure of the linkage units between ribitol teichoic acids and peptidoglycan. *Journal of bacteriology* **161**, 299-306

15. Yokoyama, K., Miyashita, T., Araki, Y., and Ito, E. (1986) Structure and functions of linkage unit intermediates in the biosynthesis of ribitol teichoic acids in Staphylococcus aureus H and Bacillus subtilis W23. *European journal of biochemistry* **161**, 479-489

16. Atilano, M. L., Pereira, P. M., Yates, J., Reed, P., Veiga, H., Pinho, M. G., and Filipe, S. R. (2010) Teichoic acids are temporal and spatial regulators of peptidoglycan cross-linking in Staphylococcus aureus. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 18991-18996

17. Zhang, Y. H., Ginsberg, C., Yuan, Y., and Walker, S. (2006) Acceptor substrate selectivity and kinetic mechanism of Bacillus subtilis TagA. *Biochemistry* **45**, 10895-10904

18. Kattke, M. D., Gosschalk, J. E., Martinez, O. E., Kumar, G., Gale, R. T., Cascio, D., Sawaya, M. R., Philips, M., Brown, E. D., and Clubb, R. T. (2019) Structure and mechanism of TagA, a novel membrane-associated glycosyltransferase that produces wall teichoic acids in pathogenic bacteria. *PLOS Pathogens* **15**, e1007723

19. Martinez, O. E., Mahoney, B. J., Goring, A. K., Yi, S.-W., Tran, D. P., Cascio, D., Phillips, M. L., Muthana, M. M., Chen, X., Jung, M. E., Loo, J. A., and Clubb, R. T. (2021) Insight into the molecular basis of substrate recognition by the wall teichoic acid glycosyltransferase TagA. *Journal of Biological Chemistry*, 101464

20. Arpino, J. A. J., Czapinska, H., Piasecka, A., Edwards, W. R., Barker, P., Gajda, M. J., Bochtler, M., and Jones, D. D. (2012) Structural Basis for Efficient Chromophore Communication and Energy Transfer in a Constructed Didomain Protein Scaffold. *Journal of the American Chemical Society* **134**, 13632-13640

21. Evers, T. H., van Dongen, E. M. W. M., Faesen, A. C., Meijer, E. W., and Merkx, M. (2006) Quantitative Understanding of the Energy Transfer between Fluorescent Proteins Connected via Flexible Peptide Linkers. *Biochemistry* **45**, 13183-13192

22. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589

23. Bhavsar, A. P., Zhao, X., and Brown, E. D. (2001) Development and Characterization of a Xylose-Dependent System for Expression of Cloned Genes in Bacillus subtilis: Conditional Complementation of a Teichoic Acid Mutant. *Applied and Environmental Microbiology* **67**, 403-410

24. Sauls, J. T., Cox, S. E., Do, Q., Castillo, V., Ghulam-Jelani, Z., Jun, S., and Hughes, K. T. (2019) Control of Bacillus subtilis Replication Initiation during Physiological Transitions and Perturbations. *mBio* **10**, e02205-02219

25. Ojkic, N., Serbanescu, D., and Banerjee, S. (2019) Surface-to-volume scaling and aspect ratio preservation in rod-shaped bacteria. *eLife* **8**, e47033

26. Nordholt, N., van Heerden, J. H., and Bruggeman, F. J. (2020) Biphasic Cell-Size and Growth-Rate Homeostasis by Single Bacillus subtilis Cells. *Current Biology* **30**, 2238-2247.e2235

27. Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins *Journal of Biomolecular NMR* **20**, 71-75

28. Friedhoff, P., Gimadutdinow, O., Ruter, T., Wende, W., Urbanke, C., Thole, H., and Pingoud, A. (1994) A Procedure for Renaturation and Purification of the Extracellular Serratia marcescens Nuclease from Genetically Engineered Escherichia coli. *Protein Expression and Purification* **5**, 37-43

29. Salzmann, M., Wider, G., Pervushin, K., Senn, H., and Wüthrich, K. (1999) TROSY-type Triple-Resonance Experiments for Sequential NMR Assignments of Large Proteins. *Journal of the American Chemical Society* **121**, 844-848

30. Eletsky, A., Kienhöfer, A., and Pervushin, K. (2001) TROSY NMR with partially deuterated proteins. *Journal of Biomolecular NMR* **20**, 177-180

31. Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR* **6**

32. Keller, R. L. J. (2004) *The Computer Aided Resonance Assignment Tutorial*, 1st ed., CANTINA Verlag

33. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. E., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. Humana Press. pp 571-607

34. Becerra, S. P., Kumar, A., Lewis, M. S., Widen, S. G., Abbotts, J., Karawya, E. M., Hughes, S. H., Shiloach, J., and Wilson, S. H. (1991) Protein-protein interactions of HIV-1 reverse transcriptase: implication of central and C-terminal regions in subunit binding. *Biochemistry* **30**, 11707-11719

35. Ross, P. D., Howard, F. B., and Lewis, M. S. (1991) Thermodynamics of antiparallel hairpin-double helix equilibria in DNA oligonucleotides from equilibrium ultracentrifugation. *Biochemistry* **30**, 6269-6275

36. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343-345

37. Jarmer, H., Berka, R., Knudsen, S., and Saxild, H. H. (2002) Transcriptome analysis documents induced competence ofBacillus subtilisduring nitrogen limiting conditions. *FEMS Microbiology Letters* **206**, 197-200

38. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671-675