2021

Brochemical Studies of Peptidoglycan Hydrolases from Commensal and Pathogenic Bacteria

Juliel Espinosa

BIOCHEMICAL STUDIES OF PEPTIDOGLYCAN HYDROLASES FROM COMMENSAL AND PATHOGENIC BACTERIA

A Thesis Presented to the Faculty of
The Rockefeller University
in Partial Fulfillment of the Requirements for
the degree of Doctor of Philosophy

by
Juliel Espinosa
June 2021

BIOCHEMICAL STUDIES OF PEPTIDOGLYCAN HYDROLASES FROM COMMENSAL AND PATHOGENIC BACTERIA

Juliel Espinosa, Ph.D.
The Rockefeller University 2021

The intestinal microbiota consists of diverse bacterial species and their effectors that play key roles in regulating human health. Interestingly, cell wall, or peptidoglycan, fragments from commensal and pathogenic bacteria can activate host immunity. The mechanism(s) by which immunologically active peptidoglycan fragments are generated, however, are not well-understood. In this regard, peptidoglycan hydrolases are ubiquitous in bacteria and possess diverse activities to remodel the cell wall during cell growth and division. These peptidoglycan hydrolases can also generate cell wall fragments in this process that are shed or recycled and available for triggering host immunity. In this thesis, we describe methodologies for the biochemical characterization of the NlpC/p60 family of peptidoglycan hydrolases to harness their catalytic activity in multiple therapeutic applications involving commensal bacteria, and to target these proteins in pathogenic bacteria.

In Chapter 1, we introduce the NlpC/p60 protein family and review the functions of NlpC/p60 hydrolases in bacterial cell division, their diverse biochemical activities, structure-function relationships, and novel functions beyond bacterial cell division.

In Chapter 2, we describe methods for the biochemical characterization of NlpC/p60 hydrolase activity. These methods encompass bacterial expression and purification of recombinant NlpC/p60 proteins, large scale isolation of peptidoglycan, and *in vitro* activity assays of NlpC/p60 hydrolase activity. Our optimized methodologies were employed to determine the peptidoglycan substrate specificity of the NlpC/p60 hydrolase SagA from the commensal bacterium *Enterococcus faecium*. These methods also enabled the evaluation of features identified in the SagA-NlpC/p60 domain structure that may be important for binding peptidoglycan substrates. Moreover, SagA-generated peptidoglycan products can activate a pattern recognition receptor in mammalian cells, which directly links SagA NlpC/p60 hydrolase activity to the activation of host immune pathways, as observed with SagA+ bacteria in mouse models of enteric infection. Our results emphasize the utility of the biochemical analysis of NlpC/p60 hydrolase activity.

In Chapter 3, we uncover the localization of SagA in *Enterococcus* and describe structure-function studies of the SagA-NlpC/p60 domain. Imaging studies of SagA+ *Enterococcus* species and fluorescently-tagged SagA constructs implicated SagA in peptidoglycan remodeling in actively dividing *Enterococcus* cells. Comparative analysis of full-length SagA and SagA-NlpC/p60 indicated that the coiled-coil SagA N-terminus may play a scaffolding or targeting function during peptidoglycan remodeling as opposed to an autoinhibitory role. A peptidoglycan-bound model of the SagA-NlpC/p60 structure along with alanine screening revealed key ligand-protein interactions that may govern *Enterococcus* peptidoglycan turnover by SagA-NlpC/p60. From this analysis, we define the catalytic dyad in the NlpC/p60 domain of SagA from multiple commensal enterococci. Together, these results highlight the possible role of SagA in *E. faecium* viability and provide mechanistic insight into how SagA NlpC/p60 hydrolase activity processes peptidoglycan and generates immunostimulatory cell wall fragments.

In Chapter 4, we explore SagA-mediated activation of host immunity and improved responsiveness to anticancer therapy in mouse tumor models. *Enterococcus* species and strains were recovered from cancer patients demonstrating responsiveness to cancer immunotherapy, but the molecular mechanism(s) behind this association were not understood. Here, we show that *Enterococcus* species and strains possessing SagA orthologs with highly conserved NlpC/p60 domains specifically enhanced efficacy of cancer immunotherapy by checkpoint blockade. Our optimized biochemical methods were applied to show that the active enterococci share conserved SagA expression, cell wall composition, and *in vitro* NlpC/p60 hydrolase activity of the respective SagA orthologs. In this analysis, I characterized the peptidoglycan hydrolase activity of a SalA from the non-protective *Enterococcus faecalis*. Heterologous expression of SagA in *E. faecalis* showed that SagA is sufficient for improving the antitumor activity of several antibodies. Moreover, using engineered strains of the probiotic *Lactococcus lactis* that heterologously expressed SagA constructs, we validated that SagA-mediated antitumor activity is dependent on secretion of catalytically active SagA. These results provide further evidence of the therapeutic applicability of SagA NlpC/p60 hydrolase activity.

In Chapter 5, we explore targeting NlpC/p60 hydrolase activity in the context of multi-drug resistant *Enterococcus* therapy. Strains of *E. faecium* Com12 displaying resistance to a particular phage encoded *sagA* mutations localized at the NlpC/p60 domain. Phage resistance upon *sagA* mutation coincided with an antibiotic fitness tradeoff that was exploited to discover the synergistic effect of phage-antibiotic combination therapy on these strains. We show that while the phage resistant strains share similar SagA secretion and cell wall profiles compared to wild type, the respective mutations abrogate NlpC/p60 hydrolase activity *in vitro*. These results imply that the phage resistant *E. faecium* Com12 strains produce catalytically impaired SagA mutants, which manifests as enhanced antibiotic susceptibility and raises the possibility of targeting NlpC/p60 hydrolase activity to treat multi-drug resistant strains of *E. faecium*.

In Chapter 6, we identify and biochemically characterize two NlpC/p60 proteins, 501F and 51B9 CwlT, as candidate multi-drug resistant *E. faecium* targets. Using our established biochemical methods, we confirm that the proteins are functional peptidoglycan hydrolases *in vitro*, thus laying the groundwork for future functional studies of these enzymes in pathogenic *E. faecium*. Our results collectively indicate that NlpC/p60 hydrolase activity may serve as a marker for the viability of pathogenic bacteria.

In Chapter 7, we summarize our work on NlpC/p60 hydrolases from commensal and pathogenic bacteria. We also discuss our ongoing efforts to develop high throughput activity assays for NlpC/p60 hydrolases and describe the next frontier of these enzymes as multi-drug resistant *Enterococcus* targets.

Together, the work described in this thesis underscore the importance of NlpC/p60 peptidoglycan hydrolase activity in mediating host-microbe interactions.

For my family

ACKNOWLEDGMENTS

I have been fortunate in my life to have strong support systems, and my time as a graduate student at the Rockefeller University has made that more apparent. I thank my thesis advisor Dr. Howard Hang for giving me the opportunity to join his lab and for his continued mentorship. Howard always challenged me to be better and ensured his door was open – physically and virtually – if I needed any guidance. I am grateful for Dr. Sebastian Klinge and Dr. Vincent Fischetti for serving in my thesis committee and demonstrating a great investment in my development as a scientist. I also thank Dr. Catherine Grimes for agreeing to be the external member of my thesis committee – I appreciate your participation and insightful questions.

I want to thank all the Hang lab members for their support, friendship, lab lunches/dinners, and memes. Dr. Byungchul Kim graciously mentored me at the beginning of my tenure in the Hang lab and for that I am forever grateful. Emma Garst initially motivated me to rotate in the lab and I was lucky to join along with Tandrila Das, Taku Tsukidate, Dr. Charles Hespen, and later Dr. Matthew Griffin and Victor Chen. Together, we formed a special bond and helped contribute to the family environment of the Hang lab. The additions of Dr. Xiaohui Zhao, Dr. Xinglin Yang, Dr. Kathryn Stein, Dr. Kyong Fam and Steven Klupt have made the Hang lab a fun and positive environment in which to work. Of course, things would not go smoothly if it were not for Dr. Angela Kong – thank you for your cheerfulness and dedication. I am also thankful to past members of the lab, Dr. Avital Percher, Dr. Virginia Pedicord, Dr. Nathan Westcott, Dr. Yen-Chih Wang, Dr. Zhenrun J. Zhang, Dr. Qiang Li, and Dr. Ti-Yu Lin, for being great colleagues.

I would not be at the Rockefeller University if it were not for the Dean's Office. I first walked through the gates at York Avenue and East 66th Street as a summer research intern in 2015 and immediately noticed how passionate everyone was about working at RU. My interactions with the Dean's Office then and throughout my graduate studies reinforced that I made the right decision to be here. I greatly appreciate Dr. Sidney Strickland, Dr. Emily Harms, Dr. Andrea Morris, Kristen Cullen, Marta Delgado, Cristian Rosario, and Stephanie Fernandez of the Dean's Office for welcoming me and helping me feel at home at RU.

The Rockefeller University has exceptional resource facilities and I have had the privilege to interact with the tremendous staff. I thank Dr. Joseph Fernandez, Caitlin Steckler, and Susan Powell of the Proteomics Resource Center for their assistance with mass spectrometry. I am indebted to Dr. Henrik Molina of the PRC and Dr. Deena Oren of the Structural Biology Resource Center who helped shape my project with their scientific guidance and technical support.

I also thank my wonderful family and friends. I have been blessed to grow up in a large Dominican family from whom I learned about the importance of family, teamwork, and dedication. I am grateful for their continued support, along with that of my in-laws and friends both local and far – thank you for your encouragement. I owe everything to my parents, Belkis and Julio, for their endless love and support. Thank you to my big little brother Julian for always cheering me on and motivating me to be the best possible version of myself. Lastly, I want to thank my amazing wife, Tiffany, for her love and companionship; I could not have asked for a better partner in life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1. NlpC/p60 hydrolases: function, activity, and structure.

Abstract

Most bacteria have a rigid cell wall as a protective layer consisting of peptidoglycan (PG), a disaccharide polymer of N-acetylglucosamine (GlcNAc) and N-acetyl-muramic acid (MurNAc) with a three to five amino acid stem attached to MurNAc. This peptide stem is characterized by either meso-diaminopimelic acid (mDAP) or L-Lysine (L-Lys) at the third peptide, which functions as the cross-linking site between PG monomers that renders the cell wall its characteristic rigidity. This highly crosslinked structure provides bacteria mechanical resistance from environmental stress and also confers their shape and morphology. During bacterial cell growth and division, PG fragments must undergo dynamic insertion and removal from the cell wall to account for changes in size. Cell wall hydrolases help control PG size and are ubiquitous in bacteria and diverse in mechanism. In particular, NlpC/p60 hydrolases are broadly distributed among bacteria and play crucial roles in cell wall turnover and cell division by cleaving bonds throughout the peptidoglycan structure. Since their discovery in the 1990s, NlpC/p60 hydrolases have been identified and studied throughout most bacterial species. Of note, NlpC/p60 hydrolases are involved in host invasion or immune activation in addition to cell wall degradation. The available functional and structural data, including the putative catalytic triad active site, provide insight into catalysis and substrate recognition. The functional, biochemical, and structural characterization of NlpC/p60 hydrolases, as presented here, is key to understanding the impact of NlpC/p60 hydrolases in evolution, cell wall recycling, and novel functions. In this chapter, we delineate factors that may be important for characterizing NlpC/p60 hydrolases, including genomics, cell division, biochemical activity, structure-function, and novel functions.

1.1 Introduction

NlpC/p60 (New lipoprotein C from *Escherichia coli* and protein of 60 kDa of *Listeria monocytogenes*) is a large family of cysteine peptidases that are broadly distributed in bacteria, viruses, archaea and eukaryotes (*1*, *2*). This family belongs to the CHAP (cysteine, histidine-dependent amidohydrolases / peptidases) superfamily, playing important roles in cell division and cell-wall recycling (*2*). Four major families have been identified, namely (i) the p60-like family, (ii) the AcmB/LytN-like family, (iii) the YaeF/Poxvirus G6R family, and (iv) the lecithin retinol acyltransferase-like (LRAT-like) family (*1*). YaeF/Poxvirus G6R / LRAT-like families show a circularly permuted catalytic domain where the relative positions of the cysteine and histidine / polar residue of the conserved putative catalytic triad are swapped in the primary sequence (*1*).

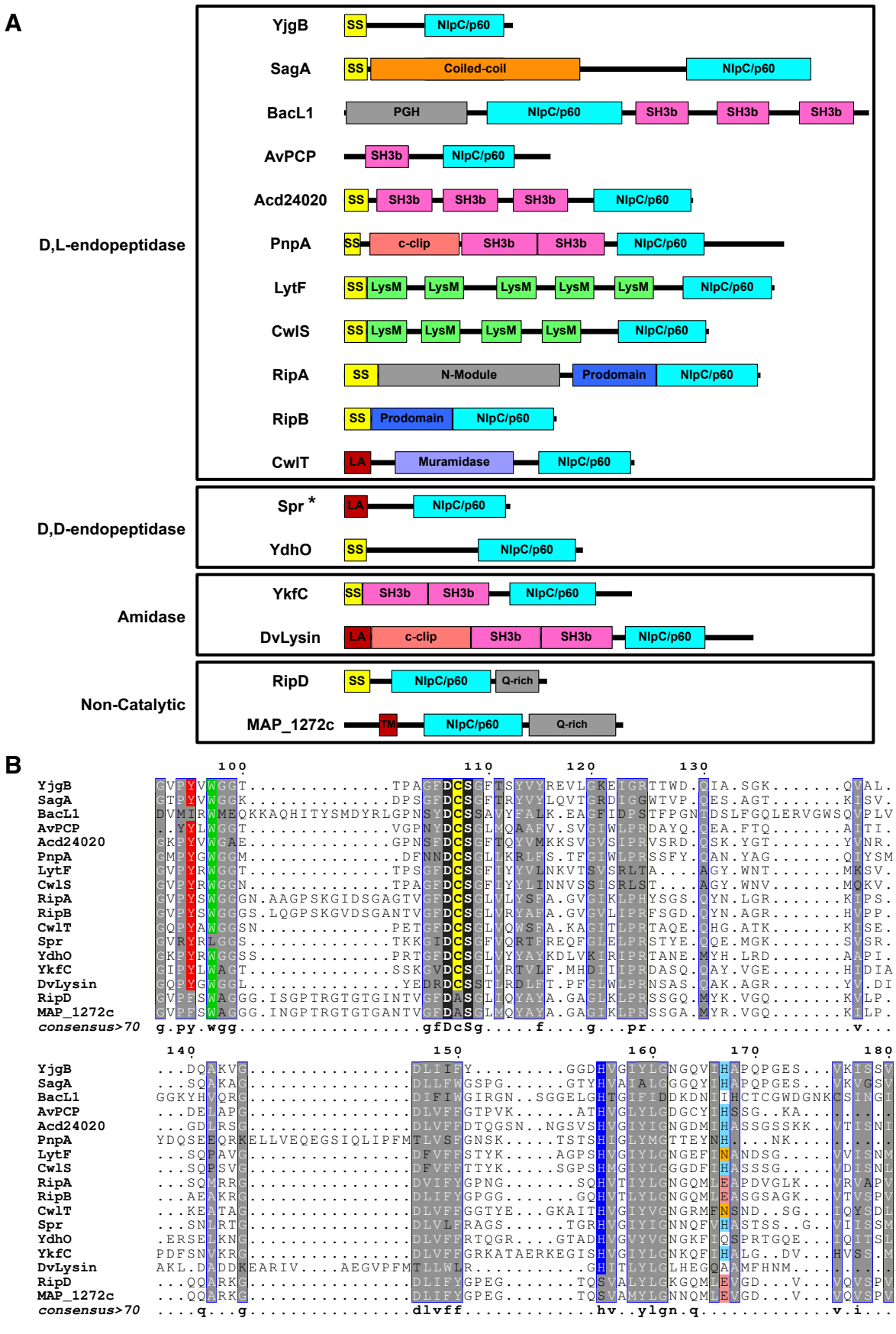
The first member of the NlpC/p60 (p60-like) family from the genus *Bacillus subtilis*, CwlE (LytF) (**Table 1.1**, **Fig. 1.1A,B**), has been shown to have D,L-endopeptidase activity that hydrolyze the D-glutamyl-meso-diaminopimelate linkage in the cell-wall extracted peptidoglycan (PG) (**Fig. 1.2A**) (*3*). Until now, multiple paralogous proteins with NlpC/p60 domains are found to be present in most bacteria (**Fig. 1.1A,B**), such as the *Bacillus subtilis* autolysins LytE / LytF (*4*, *5*), the *Listeria monocytogenes* proteins p60 / p45 (*6*), and the *Mycobacterium tuberculosis* RipA / RipB (*7*). These findings suggest that the NlpC/p60 family is a conserved family with a widespread role in the dynamics of the bacterial cell wall. Structural determination and identification of conserved motifs containing invariant Cys and His residues has led to the identification of a putative papain-like fold with a catalytic Cys in the active site of these proteins (**Fig. 1.1A,B**). The active sites of NlpC/p60s superimpose well, but the potential third triad residue can be either the usual His or Glu, or the more infrequent Gln or Asn (**Fig. 1.1B**).

In this chapter, we summarize many of the various aspects of the NlpC/p60 family; genomics, function in bacterial division, comparative analysis of activity / structure, and novel functions. As this chapter is comprehensive, it can serve as a reference point for researchers in the field.

Table 1.1. Biochemically characterized bacterial NlpC/p60 proteins. Abbreviations for Table 1 are as follows: EP, endopeptidase; CP, carboxypeptidase; NC, non-catalytic; γ-D-Gln-L-Lys, γ-D-Glutaminyl-L-Lysine; γ-D-Glu-L-Lys, γ-D-Glutamyl-L-Lysine; γ-D-Glu-mDAP, γ-D-Glutamyl-meso-Diaminopimelic acid; D-Ala-mDAP, D-Alanine-meso-Diaminopimelic acid; mDAP-D-Ala, meso-Diaminopimelic acid-D-Alanine; MurNAc-L-Ala, N-Acetylylmuramoyl-L-Alanine.

NlpC/p60	Organism	Catalytic Function	Accessory Domains	Cleavage Bond Specificity	References
YjgB	*Lactococcus lactis*	D,L-EP	-	γ-D-Gln-L-Lys	Redko et al., 2007
SagA	*Enterococcus faecium*	D,L-EP	Coiled-Coil	γ-D-Glu-L-Lys	Kim et al., 2019
BacL_1	*Enterococcus faecalis*	D,L-EP	SH3b (x3)	γ-D-Glu-L-Lys	Kurushima et al., 2013
AvPCP	*Anabaena variabilis*	D,L-EP	SH3b	γ-D-Glu-mDAP	Xu et al., 2009
NlpC_A1 & NlpC_A2	*Trichomonas vaginalis*	D,L-EP	SH3b (x2)	γ-D-Glu-mDAP	Pinheiro et al. 2018
Acd24020	*Clostridioides difficile*	D,L-EP	SH3b (x3)	γ-D-Glu-mDAP	Sekiya et al. 2020
PnpA	*Photobacterium damselae* subsp. *piscicida*	D,L-EP	c-clip & SH3b (x2)	γ-D-Glu-mDAP	Lisboa et al., 2021
LytF	*Bacillus subtilis*	D,L-EP	LysM (x5)	γ-D-Glu-mDAP	Ohnishi et al., 1999
CwlS	*Bacillus subtilis*	D,L-EP	LysM (x4)	γ-D-Glu-mDAP	Fukushima et al., 2006
RipA	*Mycobacterium tuberculosis*	D,L-EP	Prodomain	γ-D-Glu-mDAP & γ-D-Glu-L-Lys	Böth et al., 2011; Ruggiero et al., 2010
RipB	*Mycobacterium tuberculosis*	D,L-EP	-	γ-D-Glu-mDAP	Böth et al., 2011
CwlT	*Bacillus subtilis*	D,L-EP	Muramidase	γ-D-Glu-mDAP	Fukushima et al., 2008
Spr	*Escherichia coli*	D,D-EP & L,D-CP	-	D-Ala-mDAP & mDAP-D-Ala	Singh et al., 2012
YdhO	*Escherichia coli*	D,D-EP	-	D-Ala-mDAP	Singh et al., 2012
YkfC	*Bacteroides thetaiotaomicron & Bacteroides ovatus*	Amidase	SH3b (x2)	L-Ala-peptide	Xu et al., 2015
DvLysin	*Desulfovibrio vulgaris* Hildenborough	Amidase	c-clip & SH3b (x2)	MurNAc-L-Ala	Xu et al., 2015
RipD	*Mycobacterium tuberculosis*	NC	Pentapeptide Repeat	-	Böth et al., 2014
MAP_1272c	*Mycobacterium avium* subsp. *paratuberculosis*	NC	-	-	Bannantine et al., 2016

Figure 1.1. Domain architecture schematics (A) and multiple sequence alignment (B) of biochemically characterized NlpC/p60 proteins. (**A**) Abbreviations are as follows: EP, endopeptidase; NC, non-catalytic; SS, signal sequence; PGH, bacteriophage-type peptidoglycan hydrolase; SH3b, bacterial Src homology 3 domain; LysM, lysine motif; LA, Lipid Anchor; Q-rich, glutamine-rich pentapeptide repeat; TM, transmembrane helix. *Spr also exhibits L,D-carboxypeptidase activity *in vitro*. (**B**) NlpC/p60 domains were identified using InterPro, and multiple sequence alignment was completed with Clustal Omega. Regions with >70% similarity are highlighted in grey, and identical residues are highlighted in black. The conserved active site tyrosine (red), tryptophan (green), cysteine (yellow), first histidine (blue), second histidine (light blue), or polar residue (asparagine, orange; glutamate, pink) are highlighted.


A
D,L-endopeptidase
D,D-endopeptidase
Amidase
Non-Catalytic
YjgB
SagA
BacL1
AvPCP
Acd24020
PnpA
LytF
CwlS
RipA
RipB
CwlT
Spr *
YdhO
YkfC
DvLysin
RipD
MAP_1272c
SS
NlpC/p60
Coiled-coil
PGH
SH3b
LysM
c-clip
N-Module
Prodomain
LA
Muramidase
Q-rich
TM
B
100
110
120
130
140
150
160
170
180
consensus>70

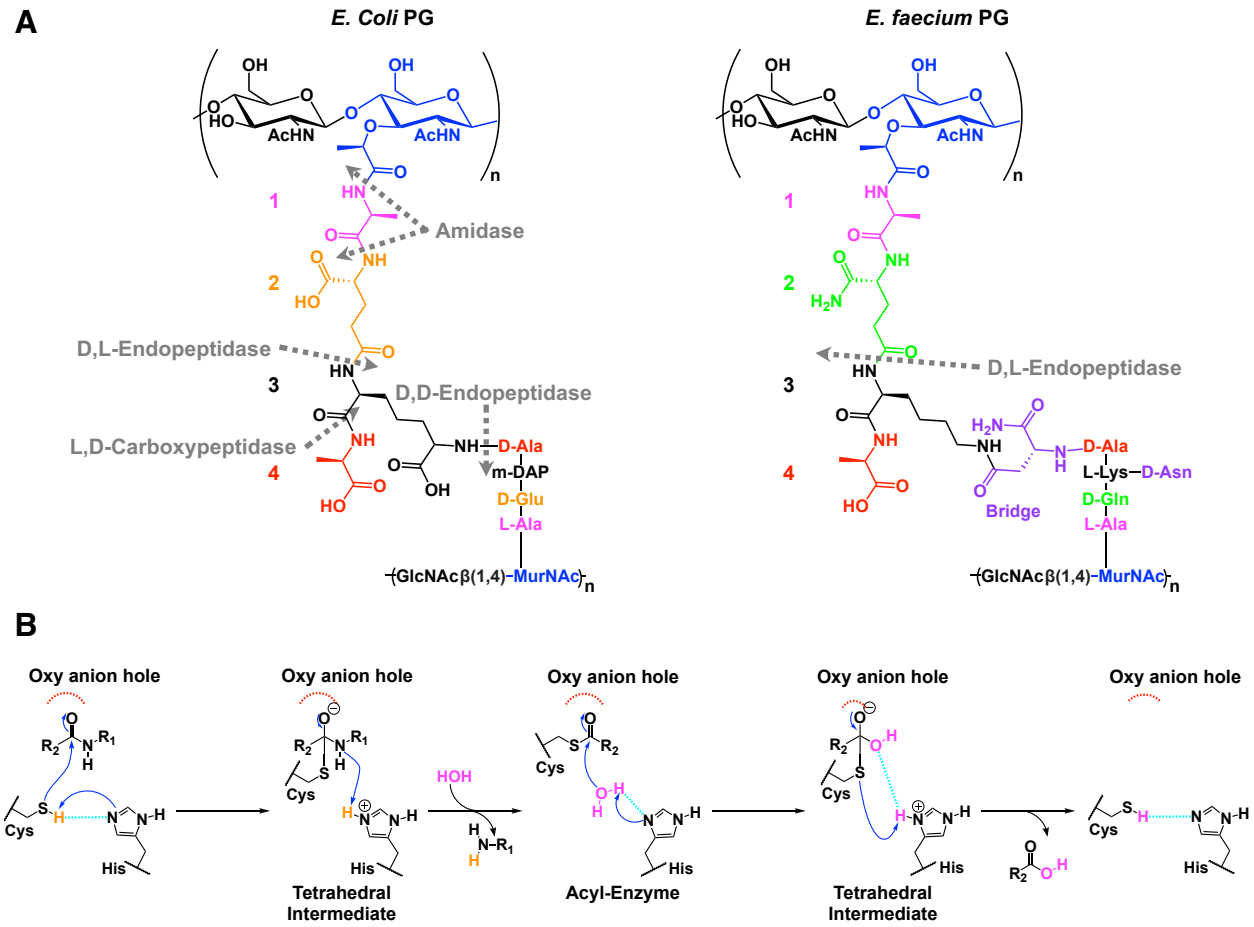


Figure 1.2. Hydrolytic specificity (A) and putative catalytic mechanism (B) of NlpC/p60 PG hydrolases. (**A**) Cross-linked *E. coli* and *E. faecium* PG fragments shown to represent m-DAP and L-lys- type PG, respectively. Position of amino acid in peptide stem numbered on the left. Cleavage sites of different NlpC/p60 peptidoglycan hydrolases marked with dashed arrows. (**B**) Putative catalytic mechanism of NlpC/p60 peptidoglycan hydrolases (adapted from Elsässer et al. *ACS Catal.* 2017).

1.2 Genomics

Proteins belonging to the NlpC/p60 superfamily can be found in all kingdoms of life, including virus, bacteria, archaea, and eukaryotes (*1*). The NlpC/p60 superfamily is diverse at sequence level and members of this superfamily contain a papain-like fold and are known to be peptidases, amidases, and acetyltransferases. Most proteins containing the NlpC/p60 domain are bacterial autolysins that participate in cell wall hydrolysis and play an important role in cell division, but can also have other biological functions, as discussed below. Sequence alignments and phylogenetic tree construction reveal broad sequence distribution of peptidoglycan active NlpC/p60 enzymes (**Fig. 1.3**). There is no defined primary sequence motif to predict NlpC/p60 substrate specificity or activity. Several eukaryotic relatives NlpC/p60 proteins (e.g., *Trichomonas*

vaginalis NlpC/p60 proteins) have closest relatives to bacterial NlpC/p60 proteins indicating horizontal gene transfer from bacteria to eukaryotes (*8*).

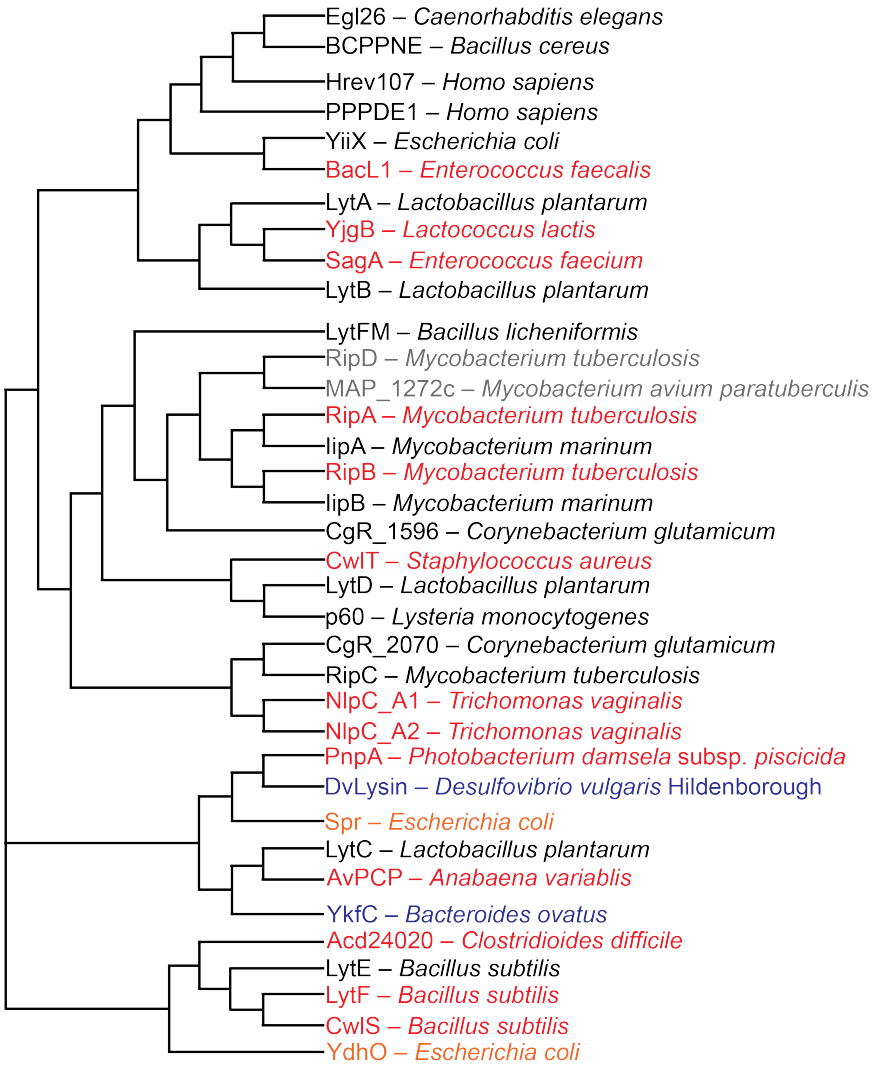


Figure 1.3. Phylogenetic cladogram of known NlpC/p60 proteins. Cladogram was constructed using Clustal Omega. Biochemically characterized NlpC/p60 enzymes are colored by activity: amidase in blue; D,L-endopeptidase in red; D,D-endopeptidase in orange, and non-catalytic in gray.

1.3 Functions in bacterial cell division

Peptidoglycan (PG) is a biopolymer surrounding the bacteria and protecting them against osmotic pressure. PG is composed of glycan strands of the alternating disaccharide N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) crosslinked by peptide bonds between the peptide stems attached to MurNAc (**Fig. 1.2A**). The stem peptides typically contain four or five amino acid residues in which the third amino acid is often meso-diaminopimelic acid (meso-DAP, mDAP) or L-lysine (L-Lys), depending on the species (**Fig. 1.2A**) (*9*). During the bacterial cell cycle, the balance between PG synthesis and hydrolysis is tightly regulated and PG remodeling is essential for cell growth and division (*10*). Enzymes involved in PG disassembly

are PG hydrolases, including glycosidases that digest the glycan strands and peptidases that degrade the connection between peptide stems.

Several PG hydrolases with NlpC/p60 domains have been identified in bacteria. Some of these enzymes are endopeptidases that cleave the peptide bond between the second and third amino acids and have shown to play an important role in cell division during which PG is synthesized for septum formation and is hydrolyzed for cell separation (**Fig. 1.2A**). For example, protein p60 encoded by the *iap* (invasion-associated protein) gene in the food-borne pathogen *L. monocytogenes* has been proposed to possess PG hydrolase activity (*11*). Cells depleted of the p60 protein form short cell filaments with defective septal structures between individual cells (*12*). This Δ*iap* mutant has reduced virulence in mice and decreased actin-based intracellular movement (*12*).

In *Corynebacterium glutamicum* R, CgR_1596 and CgR_2070 have been predicted to have an N-terminal signal peptide for secretion and a C-terminal NlpC/p60 domain responsible for PG hydrolysis (*13*). CgR_1596 is mainly localized at the septum and deletion of the *CgR_1596* gene caused cell elongation and delayed cell growth (*13*). Although single deletion of the *CgR_2070* gene did not cause any abnormalities, double deletion of both the *CgR_1596* and *CgR_2070* genes resulted in more severe phenotypes seen in the Δ*CgR_1596* strain (*13*). This mutant has a defect in cell separation and reveals longer cells having multiple septa with curved shape. They were also shown to be more susceptible to β-lactam antibiotics (*13*).

In *Lactobacillus plantarum*, two NlpC/p60 endopeptidases, LytA and LytB, take part in the positioning of the cell division site during standard growth conditions, while LytC and LytD do not (*14*). Deletion of the *lytA* gene created round cells and caused a misplacement of divisional FtsZ rings (*14*). While remaining as rods, the Δ*lytB* mutant undergoes asymmetrical division and generates mini cells. Co-inactivation of LytA and LytB led to a severe defect of cell growth, suggesting the two PG hydrolases as potential antibiotic targets (*14*).

In *Bacillus subtilis*, three NlpC/p60 endopeptidases have been implicated in the division process. LytE and LytF localize at cell separation sites and cells lacking both proteins display extra-long chains, suggesting their roles in cell separation (*3*, *15*, *16*). Although a *cwlS* mutant exhibits a wild-type shape phenotype, it colocalizes with LytE and LytF and a *lytE*, *lytF*, *cwlS* triple mutant appears as microfibers, suggesting that the three enzymes act together during cell separation (*17*). LytE has also been shown to interact with MreBH and localize at the cell periphery, implicating its dual functions in cell elongation and division (*18*).

Mechanisms of cell division in mycobacteria have been discussed in detail (*19–21*), and several studies have described the involvement of NlpC/p60-type Rip proteins. In the nonpathogenic *Mycobacterium smegmatis* (*22*) and the pathogenic *Mycobacterium tuberculosis* (*23*), allelic switching showed that the D,L-endopeptidases RipA and RipB (**Table 1.1**) are individually nonessential, but expression of at least one is necessary for viability. RipA localizes at the septum in *M. tuberculosis* (*24*) and its depletion caused a dramatic growth defect and branching and chaining of cells, suggesting its requirement for cell separation (*25*). Deletion of *ripA* in *M. tuberculosis* led to severe defects in cell division in culture; the cells displayed multiseptated chains, polar bulging, and a higher frequency of branching, which was recovered after complementation with *ripA*. This impaired cell division phenotype was not observed after deletion of *ripB*, however, indicating that *M. tuberculosis* primarily depends on the PG hydrolase activity of RipA for proper cell division (*23*). Conversely, overexpression of RipA caused a change in cell shape, from rod to sphere, and cell lysis (*26*). These data suggest that RipA requires different levels of regulation to ensure its cellular activity. For example, RipA interacts with RpfB (a lytic

transglycosylase) and the interaction regulates its hydrolytic activity (*24*). RipA contains a prodomain located before its C-terminal NlpC/p60 domain, and it has been suggested that RipA needs to undergo a proteolytic processing by MarP to release its prodomain to activate its PG hydrolase activity (**Table 1.1**) (*7*, *27*). Another NlpC/p60 endopeptidase, RipC, may also be a component of mycobacterial division machinery. Activation of RipC relies on interaction with FtsEX system (*28*). However, the role of RipC plays in cell division remains unsolved.

1.4 Biochemical activity

Primary sequence and structural analysis of the NlpC/p60 domain suggest it resembles the evolutionary minimal unit of cysteine peptidases, whose catalytic mechanism has been studied extensively (**Fig. 1.2B**) (*1*, *29–35*). The active site of NlpC/p60 hydrolases is highly conserved and consists of the catalytic cysteine, a histidine that functions as a general base to activate the cysteine for nucleophilic attack, a polar residue implicated in orienting the active site residues and substrate for cleavage, and a tyrosine involved in forming an "oxyanion hole" to stabilize the transition state intermediate formed during catalysis (**Fig. 1.1B**, **Fig. 1.2B**) (*1*, *32*, *33*, *36*, *37*). Sequence homology analysis has thus enabled the identification of bacterial NlpC/p60 PG hydrolases (**Fig. 1.1**), and genetic manipulation has demonstrated the importance of these enzymes in bacterial cell growth and division (*1*).

The enzymatic specificity of many NlpC/p60 PG hydrolases has been accomplished *in vitro* by incubating recombinant protein with various PG substrates and analyzing released PG fragment products. For instance, zymograph, spectrophotometric and in-gel profiling analysis has confirmed the hydrolytic activity, or lack thereof, of putative bacterial NlpC/p60 PG hydrolases (*38–43*). Higher resolution techniques, such as reversed-phase high-performance liquid chromatography (RP-HPLC) and mass spectrometric analysis (MS), has also enabled the biochemical characterization of NlpC/p60 PG hydrolases through the identification of precise PG fragment substrates and products. Here, we describe the biochemical activity of various bacterial NlpC/p60 proteins according to their catalytic function (i.e., endopeptidases and amidases) or ability to bind PG *in vitro* (**Table 1.1**).

1.4.1 Endopeptidase NlpC/p60

Bacteria often encode multiple NlpC/p60 PG hydrolases, many of which are secreted, and some containing accessory domains that modulate their cleavage of specific bonds throughout the endogenous PG structure (**Fig. 1.1A**) (*44*, *45*). D,L-endopeptidases specifically cleave between the second and third (2-3) amino acid in peptide stems, which are γ-D-glutamine/glutamate (γ-D-Gln/u), and typically either mDAP or L-Lys depending on the bacterial species (**Fig. 1.2A**) (*45*). Several NlpC/p60 D,L-endopeptidases have been biochemically characterized, such as YjgB from *Lactococcus lactis*, which is heavily used in manufacturing dairy products (**Table 1.1**). It was hypothesized that the PG hydrolytic activity of YjgB contributed to autolysis and release of cellular contents during cheese ripening (*46*). RP-HPLC analysis of the YjgB reaction products confirmed the presence of the expected PG fragment released upon cleavage of the γ-D-glutaminyl-L-lysine acid linkage, GlcNAc-MurNAc-dipeptide (GMDP) (*46*).

Some NlpC/p60 PG hydrolases are fused to domains with PG-binding affinity that may play a role in their catalytic activity. A notable example is the SH3b (bacterial Src homology 3) domain, which is ubiquitous in bacteria and demonstrates an affinity for PG peptide stems (*47*). In *E. faecalis*, BacL1 has bacteriolytic activity when the bacteriocin, BacA, is present and contains a N-terminal domain with homology to peptidoglycan hydrolases, a NlpC/p60 domain with γ-D-

Glu-L-Lys endopeptidase activity, and three C-terminal SH3b domains (**Table 1.1**) (*48*). The PG-binding ability of the SH3b domain in BacL1 was confirmed in a binding assay whereupon *E. faecalis* PG was incubated with recombinant protein, followed by centrifugation, and SDS-PAGE analysis of the resulting pellet (bound fraction) and supernatant (unbound fraction). Spectrophotometric analysis of *E. faecalis* PG incubated with recombinant BacL1 truncation constructs of either of the two PG hydrolase domains or the entire SH3b module showed that the NlpC/p60 domain and the SH3b domains are necessary for activity, which suggests that the SH3b domains may function as a PG-targeting motif *in vivo* (Kurushima 2013). Likewise, in the cyanobacteria *Anabaena variabilis*, AvPCP contains a single N-terminal SH3b domain and a C-terminal NlpC/p60 domain with γ-D-Glu-mDAP endopeptidase activity, but only towards muropeptides with an N-terminal L-alanine (i.e. no MurNAc) (**Table 1.1**) (*33*).

NlpC/p60 proteins with peptidase activity and SH3b domains have also been identified in the human parasite *T. vaginalis* (*8*). *T. vaginalis* encodes nine NlpC/p60 proteins and two, NlpC_A1 and NlpC_A2, are highly homologous and contain a possible signal peptide for secretion, two N-terminal SH3b domains, and a C-terminal NlpC/p60 domain (*8*). *In vitro* HPLC activity assays using mixtures of *E. coli* PG revealed that NlpC_A1 and NlpC_A2 are mDAP-type D,L-endopeptidases (*8*). In particular, both enzymes exhibited specificity towards muropeptides with tetrapeptide stems as opposed to pentapeptide stems, and higher activity towards monomeric fragments (*8*). Mutation to alanine or serine of the catalytic cysteine in the NlpC/p60 domain abolished this activity (*8*).

In the Gram-positive, spore-forming pathogen *Clostridium difficile*, the PG hydrolase Acd24020 consists of a secretion signal, three N-terminal SH3b domains, and a C-terminal NlpC/p60 domain. *C. difficile* PG is characterized by direct 3-4 crosslinks between mDAP and D-Ala, and molecular docking studies with the crystal structure of the NlpC/p60 domain of Acd24020 (Acd24020-CD) suggests that it functions as a D,L-endopeptidase (*49*). Full-length Acd24020 and Acd24020-CD hydrolyzed *C. difficile* cells in a spectrophotometric lytic activity assay, suggesting that the NlpC/p60 domain alone is sufficient for lytic activity *in vitro* (*49*). Both constructs bound to *C. difficile* cells in a gel-based binding assay and showed similar lytic activity to a panel of *C. difficile* strains, but Acd24020-CD had lower lytic activity against other *Clostridium* species compared to Acd24020 (*49*). Of note, the lytic activities of Acd24020 and Acd24020-CD were diminished at high salt concentrations, but binding to *C. difficile* cells was not, indicating that the decreased activity is not related to binding to cells (*49*). Co-incubation of recombinant Acd24020 with the cysteine protease inhibitor MTSET abolished its lytic activity, thus confirming the requirement of the active site cysteine (*49*). Interestingly, the Cys-His-Ala mutant of Acd24020 showed decreased enzymatic activity, which suggests that the second histidine in the Acd24020 Cys-His-His triad may interact with mDAP in PG substrates, as identified in the docking study (*49*).

The Gram-negative bacterium *Photobacterium damselae subsp. piscicida* (*Phdp*) secretes the PG hydrolase PnpA, which contains a N-terminal C-clip followed by two SH3b domains and a C-terminal NlpC/p60/p60 domain with interesting hydrolytic activity (*50*). PnpA hydrolyzed the γ-D-Glu-mDAP bond in monomeric muropeptides with tri-, tetra-, and pentapeptide stems to produce GMDP *in vitro*. Surprisingly, PnpA did not hydrolyze intact *Phdp* PG, so Lisboa et al. hypothesized that since PnpA is secreted, then it may instead degrade PG of competing bacteria. PnpA – and not an alanine mutant at the catalytic cysteine residue – hydrolyzed the intact PG of only two of several species tested *in vitro*, *Vibrio anguillarum* and *Vibrio vulnificus* (*50*). Lisboa et al. propose this tropism is due to differences in the three-dimensional PG architecture between

the native *Phdp* and the potential competitors *V. anguillarum* and *V. vulnificus*; the PG of *V. anguillarum* and *V. vulnificus* is marked by short and highly cross-linked glycan chains that may make it more vulnerable to degradation by PnpA (*50*).

Another accessory PG-binding domain often fused to PG hydrolases is the LysM (lysine motif) domain, which demonstrates an affinity for glycan chains containing GlcNAc (*51*, *52*). LytF and CwlS are involved in cell separation in *B. subtilis* (**Table 1.1**) and have five and four N-terminal LysM domains, respectively, along with a C-terminal NlpC/p60 domain with γ-D-Glu-mDAP endopeptidase activity (*3*, *17*). The several repeats of the LysM motif likely contribute to tight binding to PG, as LytF and CwlS were shown to cleave insoluble *B. subtilis* PG extracts *in vitro* (*17*). LytF and CwlS cleave cross-linked and non-crosslinked PG fragments, as cleavage of the former yields one free amino group (α-NH2 of mDAP) that can be labeled with dinitrophenyl for spectrophotometric analysis, whereas cleavage of the latter yields two (α- and ζ-NH2 of mDAP). Wong et al. determined that the number of LysM domains linked to CwlS correlates with PG-binding and endopeptidase activity *in vitro* (*51*). A set of CwlS constructs with one to four LysM domains removed showed decreased PG-binding and endopeptidase activity *in vitro*; these analyses indicate that three LysM domains are needed for PG binding and turnover, and that the catalytic CwlS-NlpC/p60 domain alone may have low affinity for PG since it did not bind PG in a PG pulldown assay (*51*).

There are also NlpC/p60 D,L-endopeptidases with unique domain architectures that reflect their role in the survival of their endogenous bacterial species. In the human pathogen, *M. tuberculosis*, RipA and RipB have been associated with remodeling of its unique septal PG marked by the occurrence of the non-conventional N-glycolyl-muramic acid, as well as direct DAP-DAP linkages during dormancy (**Table 1.1**) (*19*, *53*, *54*). These variations are key features in the ability of *M. tuberculosis* to resist antibiotics and survive in host macrophages (*19*, *53*, *54*). Therefore, *M. tuberculosis* must encode PG hydrolases that can recognize such variation when it arises. RipA and RipB were confirmed as γ-D-Glu-mDAP endopeptidases, as they cleaved defined PG fragments terminating in DAP and yielded primary amines that were derivatized by ortho-phthalaldehyde (oPA) for spectrophotometric analysis (*54*). Interestingly, RipA exhibited higher rates of hydrolytic activity towards insoluble mDAP-type *B. subtilis* PG substrate than RipB, and it is hypothesized that this difference is due to RipB preferring substrates with the modifications found in *M. tuberculosis* PG described above (*54*). Unique features of RipA are that it contains an N-terminal module of unknown function (*7*, *21*, *54*) and that it also exhibits PG hydrolase activity towards Lys-type PG *in vitro*, as demonstrated by activity assays using fluorescein isothiocyanate (FITC)-labelled *Micrococcus lysodeikticus* cells (*7*, *21*).

While RipA has dual specificity towards PG stems, the aptly named CwlT (cell wall lytic enzyme T (Two-catalytic domains)) in *B. subtilis* exhibits dual activity; it contains a putative lipid anchor (*38*, *55*), a N-terminal muramidase domain, and a C-terminal NlpC/p60 γ-D-Glu-mDAP endopeptidase domain (**Table 1.1**) (*38*). Muramidases catalyze the cleavage of the PG glycan backbone into GlcNAc-MurNAc-peptides, and this activity was determined by incubating recombinant CwlT N-terminus with purified glycan strands of GlcNAC-MurNAc polymer from PG, followed by RP-HPLC and MS analysis (*38*). This dual activity is proposed to open the PG structure and enable conjugative transfer of transposons – like the conjugative module that encodes CwlT in several bacteria (*37*). The *cwlT* gene was found to be required for conjugation of an integrative and conjugative element (ICE) in *B. subtilis,* ICE*B*s1 (*55*). A single point mutation at the catalytic site of the muramidase domain of *cwlT* (*cwlT-E87Q*) revealed that its muramidase activity is essential for conjugative transfer of ICE*B*s1 (*55*). A single point mutation at the catalytic

site of the NlpC/p60 domain of purified full-length CwlT (CwlT-C237A) and the NlpC/p60 domain alone showed no lytic activity *in vitro* (*55*). Introduction of this mutation to the *cwlT* gene (*cwlT-C237A*) and deletion of the NlpC/p60 domain (*cwlT-Δ(207-329)*) similarly diminished the conjugation efficiency of ICE*B*s1, indicating that the peptidase activity of the NlpC/p60 domain of CwlT contributes to conjugation, but is not required (*55*).

Whereas D,L-endopeptidases cleave within the peptide stem tethered to MurNAc, D,D-endopeptidases cleave the crosslink between PG fragments, thus enabling the insertion of nascent PG polymers and the formation of new cross-links (**Fig. 1.2A**) (*9*, *45*, *56*). In *E. coli*, the redundant lipoprotein Spr and secreted YdhO specifically cleave D-Ala-mDAP cross-links, but with slight differences in the substrates they can accommodate (**Table 1.1**, **Fig. 1.2A**) (*56*). These results showed a decrease in cross-linked muropeptides, such as cross-linked disaccharide tetrapeptide (tetra-tetra), and a corresponding increase in non-crosslinked PG monomer like disaccharide-tetrapeptide (tetra). The same analysis, but with purified cross-linked and monomeric tetrapeptides as substrates, revealed that Spr also exhibits L,D-carboxypeptidase activity, as it cleaved the terminal mDAP-D-Ala peptide bond to generate disaccharide tripeptide (*56*) (**Fig. 1.2A**). While YdhO did not share this activity, it did partially cleave D-ala-mDAP cross-links in intact *E. coli* PG (*56*).

1.4.2 Amidase NlpC/p60

Bacteria also encode amidases to help tailor the cell wall during cell growth and division by cleaving either the N-acetylmuramoyl-L-Ala linkage, thus separating glycan strands from stem peptides, or the γ-D-Glu-mDAP linkage in stem peptide with a free N-terminal L-Ala. Several NlpC/p60 amidases with N-terminal SH3b domains have been biochemically and structurally characterized, such as the NlpC/p60-related AmiA from the commensal *Bacteroides uniformis* (*57*), two YkfC orthologs from the *Bacteroides thetaiotaomicron* and *Bacteroides ovatus*, and the lipoprotein DvLysin from *Desulfovibrio vulgaris Hildenborough* (**Table 1**, **Fig. 1.2A**) (*58*).

AmiA is a member of the DUF1460 protein family of unknown function, but its structure shows that it is evolutionarily related to NlpC/p60 proteins. Biochemical characterizations of AmiA demonstrate that it specifically cleaves the N-acetylmuramoyl-L-Ala bond in disaccharide containing PG fragments with either a MurNAc or 1,6-anhydro-MurNAc (anhMurNAc) at the second position tethered to di-, tri-, or tetrapeptide stems (*57*). Conversely, YkfC orthologs can only cleave stem peptides with an N-terminal L-Ala, and structural evidence shows that SH3b domain docks into the binding groove and contributes to this substrate specificity (**Table 1.1**) (*58*). GlcNAc-anhMurNAc-peptides are major cell wall turnover products, and it is suggested that since Ami and YkfC orthologs are widely distributed throughout Bacteriodes. AmiA may generate stem peptides for YfkC to further degrade (*58*). DvLysin exhibits broader substrate specificity than AmiA *in vitro* since it also cleaves stem peptides, and structural evidence suggests that this may be attributed to the extension of the substrate recognition surface of DvLysin by an N-terminal c-clip domain and the first of its two SH3b domains (**Table 1.1**) (*58*).

1.4.3 Non-catalytic NlpC/p60

The highly conserved NlpC/p60 domain in bacteria has evolved to cleave different muropeptide linkages, and there is now evidence of non-catalytic NlpC/p60s that may serve as ligand binding domains. Aside from RipA / RipB, *M. tuberculosis* also encodes RipD, which has a NlpC/p60 domain with Ala-Ser and a unique C-terminal pentapeptide repeat tail in place of the typical catalytic Cys-His residues (*39*) (**Table 1.1**). The inactivity of RipD was confirmed using

oPA labeling of PG fragments as described above for RipA and RipB. Interestingly, a PG binding assay shows that RipD exhibits binding affinity to PG (*39*). Moreover, this affinity is increased when the C-terminal tail is truncated, thus implicating it in affecting the ability of RipD to interact with PG. Although infrequent, the presence of RipD suggests that bacteria may have evolved non-catalytic NlpC/p60 domains to interact with PG and potentially stabilize it, or prevent cleavage by active PG hydrolases (*39*). Conversely, the *Mycobacterium avium* subspecies *paratuberculosis* encodes the RipD homolog, MAP_1272c, and neither full-length nor C-terminal truncation constructs bind PG *in vitro*, but structural evidence points to a prominent negatively charged groove on the surface that may be specific for binding ligands other than PG (**Table 1.1**) (*59*).

1.5 Comparative analysis of NlpC/p60 structures

Comparison of the available NlpC/p60 structures highlights their shared structural features and individual structural characteristics. To date, there are over 25 NlpC/p60 structures available in the Protein Data Bank (PDB). The first structure of an endopeptidase containing a NlpC/p60 domain, *E. coli* Spr, was reported in 2008 by using solution NMR (**Fig. 1.4A**) (*32*). Aramini et al. obtained a single-domain structure with a papain-like α+β fold of four α helices and six β-strands and identified a Cys-His-His catalytic triad in the active site of the peptidase (*32*). Since then, other NlpC/p60 structures from different bacteria have been solved and deposited to uncover substrate recognition and catalytic mechanism (**Fig. 1.4B,C**).

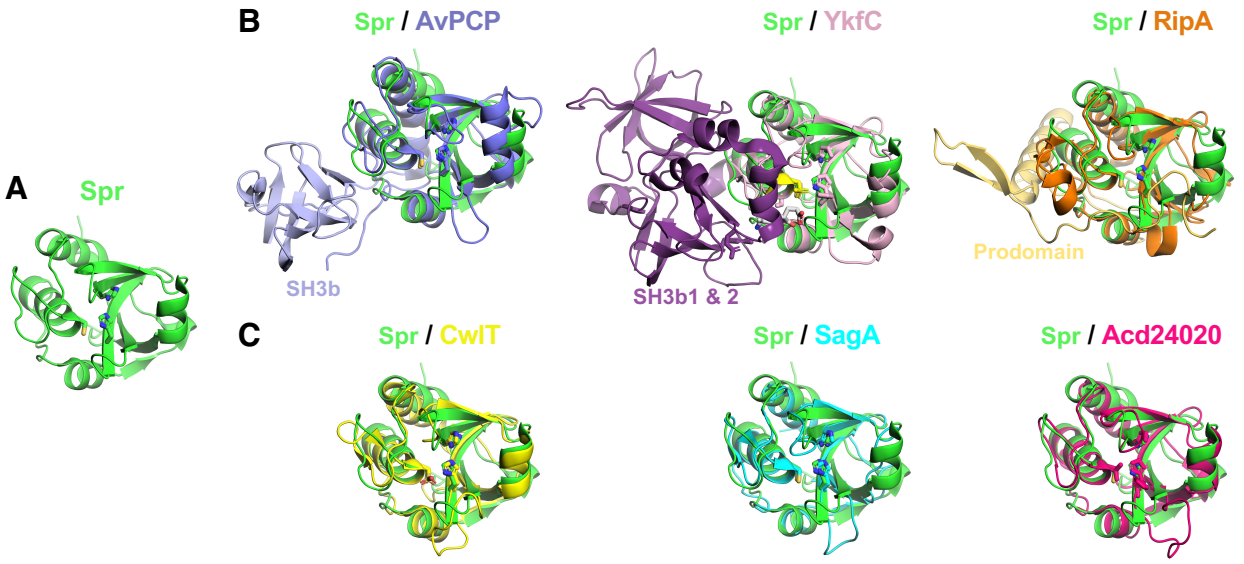


Figure 1.4. Comparative analysis of NlpC/p60 domain structures. (**A**) Conserved papain-like fold in Spr NlpC/p60 domain structure. Active site cysteine and histidine residues are represented as sticks. (**B**) Structural alignment of NlpC/p60 structures with accessory domains. (**C**) Structural alignment of NlpC/p60 domain structures used in molecular modeling studies.

1.5.1 Overall architecture of NlpC/p60 proteins

NlpC/p60 endopeptidases often consist of at least two main domains – the catalytic NlpC/p60 domain with an α/β hydrolase fold and the accessory domain(s). NlpC/p60 domains are often connected to accessory domains, such as the SH3b, LysM, and choline-binding domains.

However, several structures of NlpC/p60 domains deposited in the PDB contain only the catalytic domains. Members of these catalytic NlpC/p60 domains are approximately 150 residues and 15 kDa in size. The catalytic domain includes the highly conserved putative catalytic triad comprised of Cys, His, and His (or Asn) that is in a large cavity at the putative substrate-binding tunnel. Several structures of NlpC/p60 proteins have also been solved whose positions of the conserved Cys and His / polar residue are swapped in the primary sequence but remain in the orientation observed in other NlpC/p60 structures (*60*).

1.5.2 Substrate recognition

The recognition of the correct substrate and cleavage by NlpC/p60 catalytic activity is of great importance to the process of cell wall recycling. The NlpC/p60 substrate-binding pocket was mapped by solving the crystal structure of *Bacillus cereus* YkfC complexed with a dipeptide ligand (*61*) and by several docking studies of NlpC/p60 domains with bound muropeptides (*33*, *37*, *40*, *42*, *49*). As expected, the conserved residues include the putative catalytic triad as well as residues with key interactions with the respective ligand. The elongated groove for putative substrate-binding on NlpC/p60 surfaces has presented a conceptual hypothesis for understanding how relatively large PG substrates (e.g., cross-linked PG fragments) gain access to the active site. The observation that structures of NlpC/p60 domains have a central tunnel partially covered by hydrophilic side chains suggests that peptide stems could enter the active site through this tunnel (**Fig. 1.5A**).

A 1.79 Å atomic resolution structure with a bound dipeptide ligand has been reported for *B. cereus* YkfC (*61*). YkfC is comprised of three domains arranged in a triangle: two SH3b domains (residues 29-129, 130-207) and a C-terminal NlpC/p60 cysteine peptidase domain (residues 208-333). Interestingly, the electron density of the dipeptide L-Ala-D-Glu – which was not added during protein purification and crystallization, but may have bound during protein expression in *E. coli* – was found to be located at the interface of SH3b and NlpC/p60 domain in YkfC (*61*). This bound dipeptide is stabilized by multiple hydrogen bonds with residues, such as Glu83, Tyr118, Asp256, which are highly conserved in the YkfC subfamily of NlpC/p60 enzymes (*61*) (**Fig. 1.5B**). Ultimately, the dipeptide-bound YkfC structure facilitated molecular docking studies by serving as a template to which muropeptides and NlpC/p60 domain structures may be superimposed, as described below.

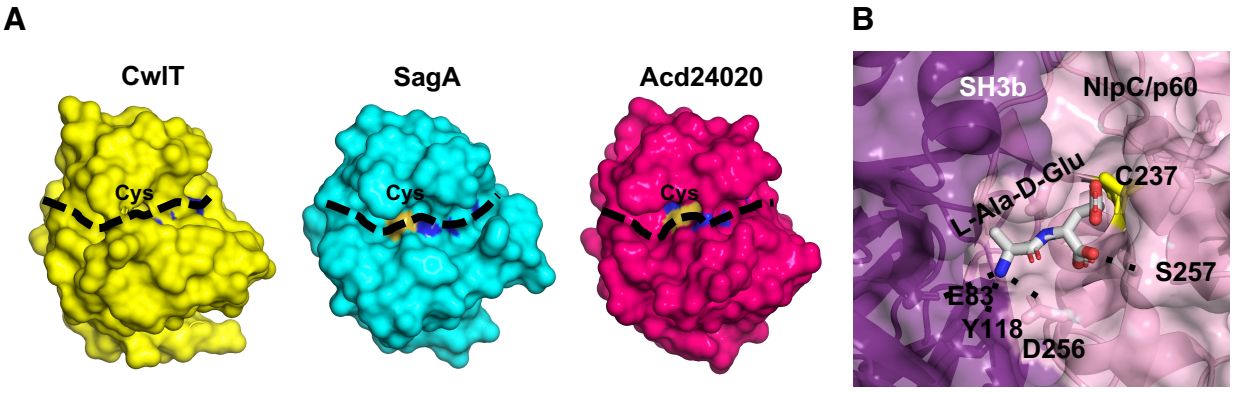


Figure 1.5. Substrate recognition in the NlpC/p60 domain. (**A**) Putative binding tunnel of the NlpC/p60 domain marked by dashed line, and active site cysteine is highlighted yellow, and conserved histidines in blue. (**B**) *Bacillus cereus* YfkC NlpC/p60 domain structure bound with L-Ala-D-Glu dipeptide. Hydrogen bonding represented as dashes.

The CwlT structures from *Staphylococcus aureus* (residue 46-340) and *Clostridium difficile* (residue 43-335) were solved to show that two domains (an N-terminal N-acetylmuramidase / bLysG domain and a C-terminal NlpC/p60 domain) are connected by a flexible linker (*37*). The overall structure and the NlpC/p60 active site of *S. aureus* CwlT is similar to the catalytic domain of *B. cereus* YkfC, which was used to guide the docking of a mDAP-type GM-tetrapeptide (*37*). Interestingly, Phe235 in *S. aureus* CwlT has the same conserved side-chain orientation as Trp228 of YkfC, which was suggested to confer substrate specificity (*37*). Both residues are located in the narrowest part of the active-site channel, indicating that conformational rearrangement of the flexible β4-β5 loop in which Phe235 resides is required for the function of the enzyme (*37*). The docked GM-tetrapeptide substrate fits the active site groove well and is stabilized by several conserved residues (Asp315, Tyr233, Thr291, and Asp315) lining the groove tunnel via hydrogen bonding and hydrophobic interactions (*37*).

Structure-function studies of the NlpC/p60 domain of *C. difficile* Acd24020 (Acd204020-CD) also highlight potential key PG substrate-protein interactions involved in *C. difficile* PG remodeling (*49*). Acd204020-CD has a unique, hourglass-shaped putative substrate-binding groove whose narrowest point is formed by the proximity of the β6-β7 loop in the top ridge of the groove to H5 in the bottom ridge (*49*). Acd204020-CD also has the flexible β4-β5 loop near the active site similar to other NlpC/p60 domain structures, which is proposed to undergo conformational change to facilitate binding of the rigid, cross-linked *C. difficile* PG (*49*). An iterative procedure consisting of superimposing and rigid-body fitting PG fragments was employed to build a model of the Acd204020-CD structure in complex with a cross-linked PG fragment. This model enabled the identification of residues that may interact with *C. difficile* PG, which is characterized by direct 3-4 cross-linking of peptide stems (mDAP1 from PG backbone 1 to D-Ala2 from PG backbone 2) (*49*). The cross-linked PG fragment resembling the letter H was docked in the putative substrate-binding groove as follows: glycan backbone 1 to the left and perpendicular to the groove with few interactions with Acd204020-CD; the attached peptide stem 1 (L-Ala1-γ-D-Glu1-mDAP1-D-Ala1) nestled within the groove spanning left to right with the most ligand-protein interactions; the mDAP1-D-Ala2 cross-link at the narrowest point of the groove with key interactions; the remainder of peptide stem 2 (mDAP2, γ-D-Glu2, and L-Ala2) to the right of this point and out into the open periphery of the groove with fewer ligand-protein interactions; and the attached glycan backbone 2 to the right and perpendicular to the groove, but away from the structure (*49*). Mutagenesis of Acd24020 and *in vitro* lytic activity assays using *C. difficile* cells were done to investigate the potential function of the ligand-interacting residues identified in the model.

In the Acd204020-CD structure, two side-chain conformations were found for the catalytic Cys299 and His354, thus supporting the hypothesis that this dyad undergoes dynamic conformational change to catalyze the cleavage of the γ-D-Glu1-mDAP1 peptide bond (*49*). Predictably, C299A, C299S, and H354A mutants of Acd24020 demonstrated no lytic activity towards *C. difficile* cells (*49*). The second histidine in the putative catalytic triad, His366, and the proximal Tyr385 formed a salt bridge and hydrogen bond with the carboxyl group of mDAP1, respectively, and several mutants (e.g., H366A, H366E, H366W, and Y385) had impaired lytic activity, implying that these residues may interact with *C. difficile* PG at this narrow point of the groove (*49*). Of the core conserved NlpC/p60 residues, Tyr287 and Asp298 interacted with L-Ala1-γ-D-Glu1, while Ser300 interacted with the carboxyl group of γ-D-Glu1 (*49*). Alanine mutants of these residues, along with neighboring residues such as the conserved PG clamp, Trp289, were inactive or showed decreased lytic activity, thus implicating them in stabilizing the

beginning of the peptide stem within the groove (*49*). Interestingly, Asp298 formed a salt bridge with Arg317 on the opposite ridge of the groove and disruption of this bond by individually switching the charge of each residue (i.e., D298R and R317D mutants) abolished lytic activity, signifying that this salt-bridge is involved in recognition of *C. difficile* PG. Furthermore, the nearby Gln309 interacted with glycan backbone 1 and was shown to be important for lytic activity (*49*).

Mutations to other structural features of the Acd24020-CD putative substrate-binding groove further illustrate how it may recognize *C. difficile* PG. A notable feature of the β6-β7 loop that contributes to the narrowest point of the Acd204020-CD groove is its prevalence of serines. Introduction of bulky residues at this narrow section (i.e., S368Y and S369Y) abrogated the lytic activity of Acd204020, whereas mutation to alanine only led to slightly diminished activity, indicating that this region of the groove needs to be accessible for proper recognition of the mDAP1-D-Ala2 cross-link (*49*). Ser371 and Ser372 are directly to the right of this point and near the remainder of peptide stem 2 (mDAP2, γ-D-Glu2, and L-Ala2), but neither mutation to alanine nor tyrosine altered the lytic activity of Acd24020, which suggests that these residues on the latter half of the β6-β7 loop are not involved in recognizing *C. difficile* PG (*49*). Further to the right of the β6-β7 loop is the putative substrate-binding groove periphery, and Thr376 and Ser378 protrude from the bottom ridge of this region. T376Y and S378Y mutants exhibited diminished lytic activity, unlike the corresponding alanine mutants (*49*). This result raises a similar hypothesis to Ser368 and Ser369 described above in that bulky residues may occlude this region of the Acd204020 groove, which should normally be accessible for effective binding of PG. Less clear is the function of the Acd204020-CD β4-β5 loop; several alanine mutants of residues in this loop showed decreased lytic activity, but, due to the inherent flexibility observed here and across other NlpC/p60 structures, the direct contribution of the β4-β5 loop to PG substrate recognition remains elusive (*49*).

Collectively, structural studies of NlpC/p60 proteins with bound PG ligands have revealed structural motifs that may be important for substrate recognition. The *B. cereus* YkfC structure set the foundation for these studies, as it provided experimental evidence of the orientation of peptide stems within the conserved central tunnel on the surface of NlpC/p60 domain structures (*61*). Subsequent docking studies with monomeric and cross-linked PG fragments have highlighted polar residues that may interact with the glycan moiety of bound muropeptides, core conserved residues for potentially stabilizing peptide stems within the groove, and residues in the groove periphery that may be involved in PG turnover (*37*, *40*, *42*, *49*). *In vitro* activity assays with *Enterococcus faecium* SagA-NlpC/p60 (*40*, *42*) and Acd24020 (*49*) have corroborated the importance of these residues for activity. NlpC/p60 structures with accessory domains have provided evidence for alternative mechanisms for determining substrate specificity, and for regulation of NlpC/p60 hydrolase activity, as described below.

1.5.3 Accessory domain(s) linked to NlpC/p60 domain structures

Several structures have been solved of NlpC/p60 hydrolases linked to SH3b accessory domains. Three crystal structures from *A. variabilis* (AvPCP; residues 14-234) and *Nostoc punctiforme* (two NpPCPs; residues 13-234) were solved by Xu *et al*, showing that the structures were almost identical to each other (*33*). Structures of AvPCP and NpPCP consist of a N-terminal SH3b domain followed by a C-terminal NlpC/p60 catalytic domain and both are highly homologous to YkfC from *B. cereus* (the rmsd is only 2.3 Å for 193 C_α atoms) (**Fig. 6**) (*61*). The SH3b domain of AvPCP also has structural similarities to the GW domains of Internalin B (*62*), and the SH3b domain of ALE-1 (*63*), both of which bind cell-wall components (*62*, *63*). In the

AvPCP structure, the single SH3b domain interacts with the NlpC/p60 domain through its distal loop, which shows a conserved domain interface. The SH3b domain of AvPCP is proposed to block part of the putative substrate-binding groove of the NlpC/p60 domain and thus contribute to its substrate specificity towards muropeptides with an N-terminal L-alanine (i.e., no MurNAc) (*33*). As mentioned previously, the crystal structure of the homologous *B. cereus* YkfC in complex with the dipeptide L-Ala-γ-D-Glu corroborated the function of the SH3b-NlpC/p60 domain interface in contributing to this substrate preference towards stem peptides (*58*). In contrast, the near identical structures of *T. vaginalis* NlpC_A1 and NlpC_A2 show an open "T" shaped putative substrate-binding groove in the NlpC/p60 domain, as the two SH3b domains are in a different arrangement than in *B. cereus* YkfC, which Pinheiro et al. note may suggest an unregulated, "toxin-like" function for NlpC_A1 and NlpC_A2 (*8*, *10*).

Structures of DvLysin (*58*) and PnpA (*50*) show an N-terminal "c-clip," two SH3b domains, and a C-terminal NlpC/p60 catalytic domain. In both structures, the c-clip wraps around the SH3b1 and NlpC/p60 domains, thus stabilizing them, while the SH3b2 domain is distant from the catalytic site (*50*, *58*). Interestingly, the c-clip-SH3b1 domain interface extends the putative substrate-binding groove of the NlpC/p60 domain to the left near the proposed glycan binding site (*50*, *58*). The NlpC/p60-SH3b1 domain interface partially covers the right side of the groove at the proposed cross-link binding site (*58*), with PnpA having an additional antiparallel β-sheet and 3_{10} helix insertion in its c-clip domain that protrudes into this region and makes it more inaccessible (*50*). As Xu et al note, the structural contributions of the DvLysin SH3b1 domain to the putative substrate-binding groove of its NlpC/p60 domain may allow for more precise PG recognition and thorough PG turnover (*58*). The conserved topology in PnpA suggests that NlpC/p60 proteins with a similar four-domain architecture may also have extended putative substrate-binding grooves (*50*).

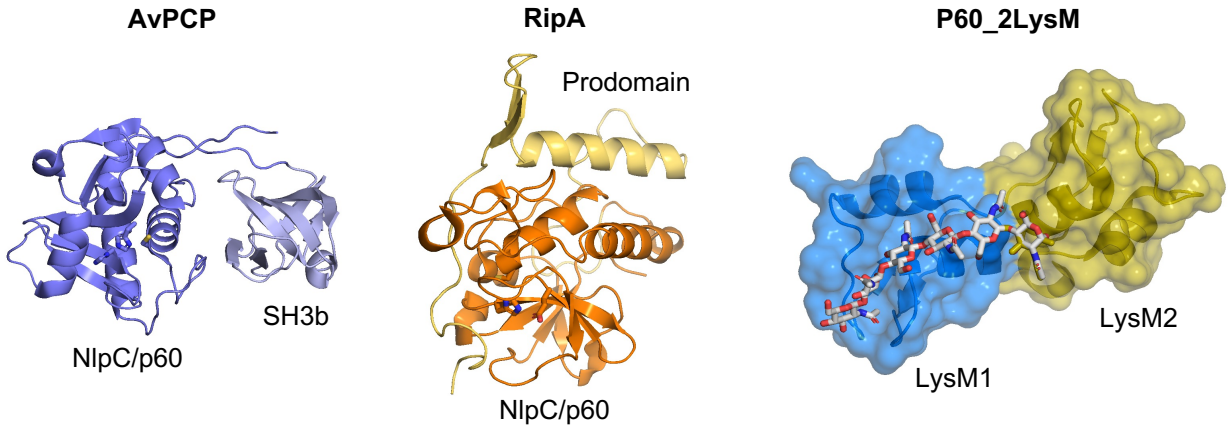


Figure 1.6. X-ray crystal structures of accessory domains of NlpC/p60 hydrolases.

Structures of mycobacterial Rip proteins have been previously reviewed in detail (*21*). An X-ray structure of truncated *M. tuberculosis* RipA (residues 263-472) showed a PB015164 domain identified in the PFAM database linked to the catalytic NlpC/p60 domain (*7*). Notably, the PB015164 domain (labeled as prodomain) blocks the catalytic cleft of the NlpC/p60 domain, suggesting a potential role as a regulator of enzyme activity (*7*) (**Fig. 1.6**). X-ray structures of full-length RipA and RipB from *M. tuberculosis* were later determined, which revealed that while the prodomains fold differently in each structure (a β-hairpin and one helix in RipA, and two helices

in RipB), they still occlude their respective putative substrate-binding grooves, further supporting the proposed regulating function of the prodomain (*54*). RipA has a unique N-terminal module and the full-length RipA structure shows that this domain forms a coiled-coil structure reminiscent of scaffolding proteins, but whose role in RipA activity is undetermined (*21*, *54*).

A structural study of P60_tth from *Thermus thermophiles* was also presented (*52*). This structure shows a D,L-endopeptidase containing an N-terminal binding domain of two LysM domains and C-terminal NlpC/p60 domain. A co-crystal structure of two LysM domains bound with chitohexaose (GlcNAc polymers) was also solved (P60_2LysM). Interestingly, chitohexaose is recognized by the two LysM domains, implying that LysM domains could act cooperatively to enhance the binding of these proteins to long carbohydrates (i.e., intermolecular dimerization) (*52*) (**Fig. 6**). This observation indicated that the two LysM domains from *T. thermophiles* LysM mediate recognition of PG for NlpC/p60 catalytic activity (*52*).

1.5.4 Structures of permuted papain-like NlpC/P60 enzymes (PPNEs)

PPNEs are members of the NlpC/p60 superfamily that have a circularly permuted active site; the relative positions of the Cys and His / polar residue of the conserved putative catalytic triad are swapped in the primary sequence (*1*, *60*). Several structures of PPNEs have been solved: *B. cereus* PPNE (BcPPNE), *E. coli* YiiX, and human PPPDE1 (*60*). In all three structures, the circularly permuted residues, along with the conserved Tyr of the NlpC/p60 core, are arranged in a conformation similar to structures of canonical papain-like NlpC/P60 enzymes (CPNEs) (*60*). Moreover, a region near the catalytic cysteine was found to be primarily hydrophobic, which differs from the corresponding hydrophilic region in CPNEs that is proposed to recognize γ-D-Glu in PG peptide stems (*60*, *61*). Interestingly, the three structures are in different bound confirmations: *B. cereus* PPNE (BcPPNE) in a more tightly closed conformation around a bound lysine compared to YiiX, which contains a bound fatty acid, and *apo*-PPPDE1 in an open confirmation (*60*). Xu et al. propose that this distinguishing hydrophobic feature in the binding site of PPNE structures, and the structural evidence of lipid-binding, may indicate that some PPNEs prefer lipid-like substrates over muropeptides (*60*).

1.6 Novel functions beyond bacterial cell division

Most of the characterized NlpC/p60 enzymes are D,L-endopeptidases. While their biochemical function seems to be conserved, the physiological roles of NlpC/p60 proteins are diverse, including involvement in cell separation, differentiation, and cell wall lysis / turnover (*64*); bacterial virulence (*65–67*); and host-pathogen interactions (*68*, *69*). Moreover, since NlpC/p60 mutant strains of pathogenic bacteria often display impaired cell division and growth phenotypes, their antibiotic sensitivity is of great interest. Several studies have shown the relationship between NlpC/P60 hydrolase activity, host interaction, and antimicrobial susceptibility, as described below.

1.6.1 Invasion and virulence, and antimicrobial susceptibility

Lenz et al. report that NlpC/p60 PG hydrolases are involved in bacterial virulence. In *L. monocytogenes*, 17 SecA2-depedent proteins were identified using mass-spectrometry, including a surface-associated N-acetylmuramidase and a secreted endopeptidase (p60) (*65*). In-frame deletions in the genes coding for p60 (Δp60) led to decreased virulence of *L. monocytogenes* (*65*). In the case of the Δp60 strain, the effect on virulence could be complemented only by expression of a full-length, catalytically active p60 (*65*). These data suggest that SecA2-dependent secretion

has evolved in part to promote the secretion of NlpC/p60 hydrolases along with other autolysins in *L. monocytogenes* virulence (*65*). Another study showed that the auto-lysin p60 of *L. monocytogenes* is involved in host-cell invasion and may serve a similar function to invasion (*70*). These findings suggest that p60 not only causes the degradation of the cell wall polymer, but also actively participates in the invasion process (*70*). In *L. monocytogenes*, p60 was shown to indirectly enhance NK cell activation in the absence of additional *L. monocytogenes* factors, contributing to bacterial pathogenicity (*71*). Purified p60 protein binds to dendritic cells (DCs) and induces IL-18 secretion indirectly, which is required for NK cell activation by p60 in co-culture (*71*). These data provide evidence that the NlpC/p60 locus is essential for invasion and virulence in bacteria.

The Rip proteins have also been shown to play an important role in the virulence of mycobacteria and their susceptibility to antibiotics, lysozyme, and detergent. In the pathogenic *Mycobacterium marinum*, the *iip* (invasion and intracellular persistence) locus consists of two genes in tandem species, *iipA* and *iipB* (*72*), which are homologs of the aforementioned *ripA* and *ripB* (Rpf interaction protein A and B) (*24*). The predicted amino acid sequences of both genes contain conserved NlpC/p60 domains at the C-terminus. The mutation of the conserved Asp / Cys / Ser / Gly (DCSG) motif to Ala in the IipA NlpC/p60 domain completely abrogated the ability to complement the *iipA*::*kan* mutant for invasion in addition to involvement in cell division, whereas expression of wild-type fully complemented the *iipA*::*kan* mutant for invasion and intracellular survival (*72*). Notably, the *iipA*::*kan* mutant had abnormal bacterial morphology and was more sensitive to antibiotics and lysozyme, suggesting that the lack of IipA renders the cell wall of *M. marinum* more permeable (*72*). Likewise, in the nonpathogenic *M. smegmatis* often used as a model to study mycobacterial infections, deletion of *ripA* (Δ*ripA*) and *ripB* (Δ*ripB*) caused increased susceptibility to antibiotics of varying sizes and to detergent compared to wild type, whereas only the Δ*ripA* strain was more susceptible to the cell wall-damaging lysozyme (*22*). Recent studies directly in the pathogenic *M. tuberculosis* revealed that inducible silencing of *ripA* sensitized *M. tuberculosis* to several cell-wall targeting drugs, such as vancomycin, and to the RNA polymerase inhibitor and frontline Tuberculosis drug rifampin (also known as rifampicin) (*23*). Moreover, this RipA depletion significantly decreased replication of *M. tuberculosis* in murine bone marrow derived macrophages *ex vivo* and led to clearance of *M. tuberculosis* from infected mice (*23*). These studies implicate RipA as a major player in the pathogenicity of mycobacteria, which is consistent with its proposed function as the main PG hydrolase required for cell division in *M. tuberculosis* (*23*).

In another study, it was shown that *M. tuberculosis* encodes Rv2190c, a "hypothetical protein" containing an NlpC/p60 domain at C-terminus (*67*) that has subsequently been renamed as RipC (*28*). Transposon insertion after the predicted NlpC/p60 domain of *Rv2190c* showed that the resulting mutant was attenuated for growth and virulence *in vivo* (*67*). The median survival for mice infected with the *Rv2190c* mutant was significantly longer than wild-type due to the defective growth of the mutant strain lacking the NlpC/p60 domain (*67*). Similar to the *ripA* mutant strains described above, the *Rv2190c* mutant strain demonstrated susceptibility to lysozyme and detergent (*67*), and a deletion mutant was hypersensitive to rifampin in a concentration-dependent manner (*28*). Collectively, these studies confirm the importance of NlpC/p60 Rip proteins for the pathogenicity of mycobacteria.

1.6.2 Functions in eukaryotes

Other well-characterized protein members of the NlpC/p60 superfamily include those belonging to the PPNE subfamily, such as the poxvirus protein G6, lecithin:retinol acyltransferase (LRAT), the nematode developmental regulator Egl-26 (egg-laying), and the tumor suppressor H-rev107 (*1*). G6 is a conserved protein with enzyme activity among poxviruses that contributes to their virulence (*66*). LRAT transfers the acyl group from the *sn-1* position of phosphatidylcholine to a retinoid acceptor, a reaction involved in the retinoid (visual) cycle. Mutations in LRAT cause visual loss and retinal degradation in animal models and human patients (*73*). Egl-26 regulates vulval morphogenesis in *Caenorhabditis elegans* and mutation of Egl-26 leads to a blocked lumen and an egg-laying defect (*74*). H-rev107 possesses phospholipase activity that inhibits tumor growth and induces apoptosis (*75*). Both Egl-26 and H-rev107 have high protein similarity to LRAT (*1*).

It has been suggested that eukaryotes acquire NlpC/p60 proteins from bacteria by lateral gene transfer. For example, a 14 kDa protein (LytFM) belonging to the NlpC/p60 superfamily has been identified in the house dust mite *Dermatophagoides pteronyssinus* (*76*, *77*). LytFM is a bacteriolytic enzyme with putative peptidase activity (*77*). Since mites do not have a cell wall, LytFM has been proposed to protect mites against pathogenic bacteria and/or utilize bacteria as a food source (*77*). The sequences flanking the *lytFM* gene contain a bacterial promoter, suggesting that the enzyme is derived from prokaryotic endosymbionts by lateral gene transfer (*76*–*79*).

Phylogenetic analyses suggested that the parasite *T. vaginalis* acquired its nine NlpC/p60 genes from bacteria by lateral gene transfer (*8*). Of the *T. vaginalis* NlpC/p60 proteins, NlpC_A1 and NlpC_A2 are active against bacterial cell wall and may help the parasite to outcompete vaginal bacteria (*8*). Expression of NlpC_A1 and NlpC_A2 was upregulated in *T. vaginalis* in the presence of bacteria in mixed culture, which corresponded with a reduction in the amount of *E. coli* in culture (*8*). Overexpression of wild type NlpC_A1 and not the inactive C179S mutant significantly increased bacterial clearance in culture (*8*). These studies further highlight an important role of NlpC/p60 proteins in host-bacterium interactions.

1.7 Introduction to the work presented in this thesis

In the more than two decades since their initial discovery, much has been learned about the evolution, genomics, and catalytic mechanism of the NlpC/p60 hydrolases. Likewise, in the over ten years since the first structure of a NlpC/p60 hydrolase was published, much has been learned about the *in vitro* function and substrate recognition of the catalytic subunit and its binding domain. However, there are still many aspects of catalysis, substrate recognition, cellular functions, and functions in modulating host immunity and health that remain to be clarified for NlpC/p60 hydrolases. In the following chapters, we describe optimized methodologies for the biochemical characterization of NlpC/p60 PG hydrolase activity, which provide a foundation for studying the roles NlpC/p60 hydrolases may play in mediating host-microbe interactions.

In Chapter 2, we develop biochemical methods to evaluate the catalytic activity of SagA, a secreted NlpC/p60 PG hydrolase from the commensal bacterium *Enterococcus faecium* Com15, in conferring protection to mice in models of enteric infection. In Chapter 3, we use fluorescence microscopy to implicate SagA in *E. faecium* PG remodeling and perform *in vitro* structure-function studies of the catalytic NlpC/p60 domain of SagA to understand how it may bind PG substrates to tailor the *E. faecium* cell wall and generate immunostimulatory cell wall fragments. In Chapter 4, we assess the utility of our optimized methods to explore how SagA PG hydrolase activity can improve responsiveness to anticancer therapy in mouse tumor models. In Chapter 5, we discuss

the acquisition of phage resistance in *E. faecium* Com12 due to mutations to *sagA* and the implications of the corresponding fitness tradeoff of enhanced antibiotic susceptibility. In Chapter 6, we apply our biochemical workflow to identify other NlpC/p60 hydrolases encoded by *E. faecium* strains as candidate targets for functional experiments in pathogenic strains.

Chapter 2. Biochemical analysis of NlpC/p60 PG hydrolase activity.

Introduction

Gut microbiota in animals consists of diverse bacterial species that promote host intestinal epithelial barrier function and serve as a source of unique anti-infective factors. Dysbiosis of these beneficial gut microbial communities is associated with inflammatory bowel diseases (IBDs) (*80*), such as Crohn's disease and colitis, and cancer (*81*, *82*). Therefore, understanding how gut microbiota and its effectors interact with host tissues and modulate health in humans is of great importance. Interestingly, cell wall fragments from commensal bacteria can activate host innate immune signaling pathways (*83–87*). The identity of these bacterial cell wall fragments, and the mechanism by which they are generated, merits further investigation due to their therapeutic potential (*87*).

PG is a vital component of the bacterial cell wall, and its dynamic remodeling is crucial for proper cell growth, division, and survival (*9*). In this regard, PG hydrolases are ubiquitous in bacteria and possess diverse activities to cleavage of internal peptide stem linkages for PG remodeling during cell growth and division (*45*, *88*). In the process of remodeling the cell wall, these PG hydrolases can also generate non-crosslinked muropeptides that are shed or recycled and available for triggering host immune responses through pattern recognition receptors such as nucleotide-binding oligomerization domain containing (NOD) 1 and 2 proteins or inflammasomes (*85–87*). Indeed, our laboratory has shown that a NlpC/p60-type PG hydrolase also termed secreted antigen A (SagA) from *Enterococcus faecium*, confers pathogen tolerance to *Caenorhabditis elegans* as well as protects mice against enteric pathogens such as *Salmonella* Typhimurium and *Clostridium difficile* (*68*, *69*). The ability of NlpC/p60 PG hydrolases to tailor the bacterial cell wall and generate distinct PG fragments make them interesting targets for studying how commensal bacteria and their effectors promote host tissue integrity and immunity (*83*, *87*).

As described above, the NlpC/p60 family is a critical PG hydrolase family involved in cell wall maintenance, survival, virulence in gram-positive bacteria, and now protection against enteric pathogens. It has been shown that NlpC/p60 proteins, among other PG hydrolases, play a modulatory role on the level of enzymatic activity and the specificity for PG substrate (*37*, *57*, *58*). Many reports have described the characterization of NlpC/p60 activity using various assays (*38*, *54*, *89*). The most commonly used procedure is the zymogram assay, which involves incubation of purified proteins with autoclaved bacterial cells as a substrate followed by renaturing SDS-PAGE (*90*). Zymography is a versatile two-stage technique involving protein separation by electrophoresis followed by detection of hydrolase activity. Protein separation by SDS-PAGE is limited by the fact that some peptidases do not renature and hence cannot be detected following treatment with SDS. Another approach is the cell wall degradation assay. It is carried out using fluorescein isothiocyanate (FITC)-labeled bacterial cells, which are incubated with purified protein, followed by measuring FITC-labeled cell wall fragments released in solution (*7*). A major shortcoming of these assays, however, is the low detection of activity of purified NlpC/p60 domains owing to a lack of suitable synthetic substrates and low activity toward protein substrates used (*39*, *54*, *91*). In this context, the characterization of activity for NlpC/p60 hydrolases, and identification of their specific substrates, is of critical importance.

Biochemical analysis methods may enable assignment of observed hydrolase activity to specific PG fragments as substrates. This analysis requires robust methods for purification of the large quantities of this protein and substrate pool. In this chapter, we describe modified and highly optimized methods for expression and purification of NlpC/p60 hydrolases that consistently yields

large quantities of pure, functional protein using an *Escherichia coli* expression system, isolation of large quantities of bacterial PG, and two enzymatic assays for the biochemical characterization of NlpC/p60 hydrolases. These methods were applied to characterize the biochemical activity of *E. faecium* SagA as a D,L-endopeptidase with a substrate preference towards pre-digested L-Lys-type PG. Moreover, molecular docking studies using the SagA-NlpC/p60 structure with a bound PG ligand revealed two Trp residues, Trp433 and Trp462, as potential PG peptide stem "clamps" in the putative substrate binding groove of SagA-NlpC/p60. Single and double alanine mutants of these residues exhibited impaired D,L-endopeptidase activity using both the in-gel and LC-MS PG profiling methods. These results were reported in this publication (*40*). The methods described here were optimized in collaboration with Dr. Byungchul Kim and published in this publication (*41*). Dr. Kim performed the PG profiling discussed below, solved the SagA-NlpC/p60 structure, executed the molecular docking studies, and performed the in-gel activity assay with intact PG and pre-digested *E. coli* PG. Dr. Yen-Chih Wang and Dr. Charles Hespen performed the *ex vivo* and *in vivo* experiments discussed below, respectively.

Results and Discussion

Bacterial expression and purification of recombinant NlpC/p60 proteins

SagA is a ~55 kDa protein comprised of a secretion-signal peptide and three domains: a coiled-coil N-terminus, serine/threonine-rich linker, and NlpC/p60 C-terminus with a conserved putative catalytic triad (Cys443, His494 and His506) (*69*). In order to determine the biochemical activity of SagA *in vitro*, we first optimized the expression and purification of full-length SagA and SagA truncation constructs. Previously, our laboratory cloned *E. faecium* Com15 *sagA* into the pET-21a(+) vector that has an IPTG-inducible T7 promoter and a C-terminal polyhistidine (6xHis) tag (*69*). This vector was transformed into BL21-CodonPlus® (DE3)-RIL competent *E. coli* (Agilent), which contains extra copies of the argU, ileY, and leuW tRNA genes that are required for heterologous expression of *sagA* in *E. coli*. IPTG-inducible expression of recombinant full-length SagA in *E. coli* led to a decrease in culture optical density (OD), suggesting cell lysis (*69*). Conversely, neither a mutant of the putative catalytic triad (SagA C443A/H494A/H506A), nor a signal sequence truncation mutant (SagA ΔSS) showed cell lysis, indicating that SagA secretion may be toxic to *E. coli*, and/or secreted SagA may hydrolyze *E. coli* PG *in trans* (*69*).

Taking these observations into account, we induced expression of C-terminally His_6-tagged SagA ΔSS and SagA for affinity purification from *E. coli* cytoplasm and culture supernatant, respectively (**Fig. 2.1**) (*41*). Protein expression was held to 2 hours to limit the expression of nontarget proteins that may carry over as contaminants throughout the purification, and to avoid spontaneous mutation in *sagA* that was observed in cells expressing and secreting SagA overnight (data not shown). Purification of SagA ΔSS required lysis of the *E. coli* cell pellet, which often involves the use of the glycosidase lysozyme – a PG hydrolase that could affect downstream activity assays if not removed during the purification of the target NlpC/p60 protein. Therefore, lysozyme was omitted from the lysis buffer and a water bath sonicator was used to lyse *E. coli* cell pellets, followed by collection of the cell-free lysate by ultracentrifugation for affinity chromatography (**Fig. 2.1B**). Sonication in a water bath consistently led to cell-free lysate with less contaminants than if using a probe sonicator (data not shown) and allowed for lysis of multiple cell pellets in parallel. Purification of SagA from *E. coli* culture supernatant proved to be more challenging to optimize. First, the decreased culture OD from cell lysis post-SagA expression and secretion (**Fig. 2.1A**) corresponded with lower protein yields and necessitated the collection of

larger volumes of culture supernatant. Second, unlike the cell-free lysate, the *E. coli* culture supernatant contains nutrients that impeded binding to the His60 Ni-superflow resin (Takara) that we found best captured His_6 tagged SagA for affinity chromatography. To address these issues, we used Jumbosep™ centrifugal devices with 30 kDa molecular weight cut-off membrane inserts (Pall Life Sciences) to both concentrate and filter the culture supernatant and allow for buffer exchange to phosphate-buffered saline (PBS). The cell-free lysate (SagA ΔSS) and processed culture supernatant (SagA) were incubated with His60 Ni-superflow resin for 1 hour and overnight, respectively, since the former had more contaminants. After elution from the resin with 300 mM imidazole, the SagA constructs were dialyzed to PBS for storage, or further purified via size exclusion chromatography (**Fig. 2.1B**) (*41*). Expression and secretion of the SagA C-terminal domain (i.e. SagA-NlpC/p60) in *E. coli* was unsuccessful, so a signal-sequence truncation construct that demonstrated robust expression was pursued for purification from *E. coli* cytoplasm, as previously described.

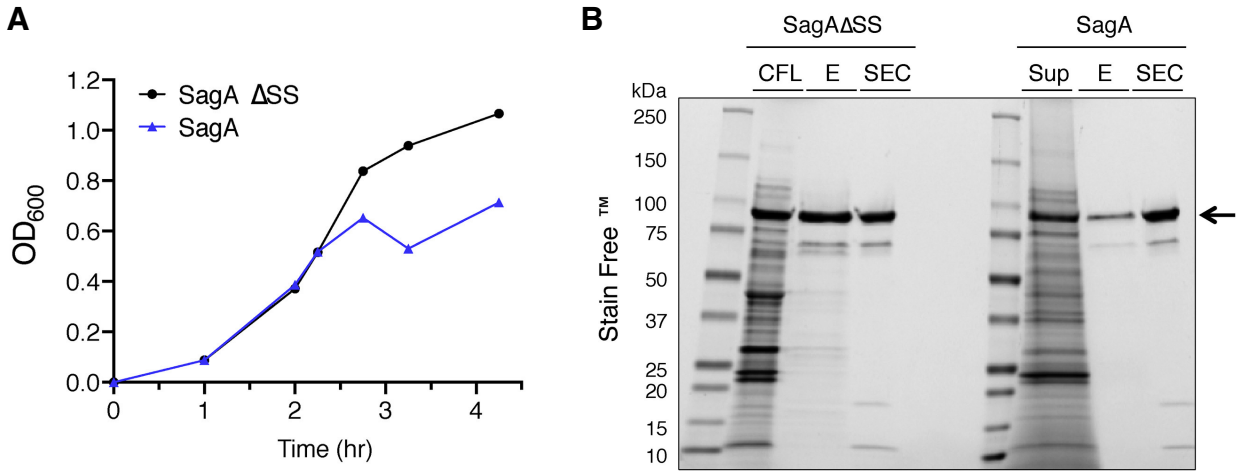


Figure 2.1. Expression (A) and purification (B) of SagA from *E. coli* cell pellet (SagA ΔSS) and supernatant (SagA). CFL, cell-free lysate; E, Ni-NTA^{2+} chromatography elution; SEC, Size exclusion chromatography. Arrow in (**B**) denotes full-length SagA.

Large scale isolation of PG

PG isolation from diverse bacterial species has been well studied throughout the last several decades (*92–96*). To study the effects of SagA expression and secretion on host strain cell wall morphology, and to deduce the biochemical activity of SagA *in vitro*, we developed a modified protocol (*96*) applicable for isolating PG from up to 3 L cultures of various Gram-positive bacteria, such as *E. faecium*, *E. faecalis*, or *Lactobacillus plantarum* (*41*). PG isolation involves harsh treatment (i.e., heat, detergent, enzymes, and acid) to separate the PG sacculus from the rest of the cell and its contents. Therefore, special care must be taken to ensure that the containers used maintain their structural integrity throughout the process, which includes many centrifugation steps that can warp plastic. Most protocols use microcentrifuge tubes, which generally withstand the PG isolation process, but limit the yield. The use of 50 or 500 mL conical bottom polypropylene centrifuge tubes (Corning) fulfilled these requirements for the large scale isolation of PG from multiple bacteria in parallel. In addition, a key step in PG isolation is the removal of the detergent used to lyse cells because it can inhibit the subsequent enzyme treatment needed to degrade

released nucleic acids and proteins (*96*). Repeated washes are therefore necessary and entail resuspending the insoluble cell wall preparation with water, centrifugation, and then removal of supernatant. However, adding too much water often turns the cell wall preparation flocculent, which makes it harder for it to pellet after each centrifugation step and leads to sample loss. Therefore, we determined that washes at a ratio of 120 mL water to every 3 L bacterial cell pellet was sufficient to both remove detergent and allow the insoluble cell wall preparation to pellet after centrifugation (*41*). This method consistently produced near-gram quantities of insoluble, purified PG that facilitated profiling of the cell wall composition of individual strains, and served as substrate for enzymatic assays, as described below.

Insoluble, purified PG isolated from bacteria were treated with muramidases, such as mutanolysin, which catalyze the cleavage of the β-1,4-linkages between MurNAc and GlcNAc residues in PG, thus yielding soluble non-crosslinked and crosslinked PG fragments. These PG fragments could then be resolved by LC-MS for identification and purification (**Fig. 2.2**, top), or used as potential substrates for *in vitro* activity assays. Using these techniques, we determined that *E. faecium* PG contains smaller non-crosslinked muropeptides compared to the non-protective *E. faecalis* (*40*, *68*, *69*). Moreover, LC-MS PG profiling of the protective *E. faecalis* strain engineered to express and secrete SagA revealed that this strain had less cross-linked PG than the wild type (*40*).

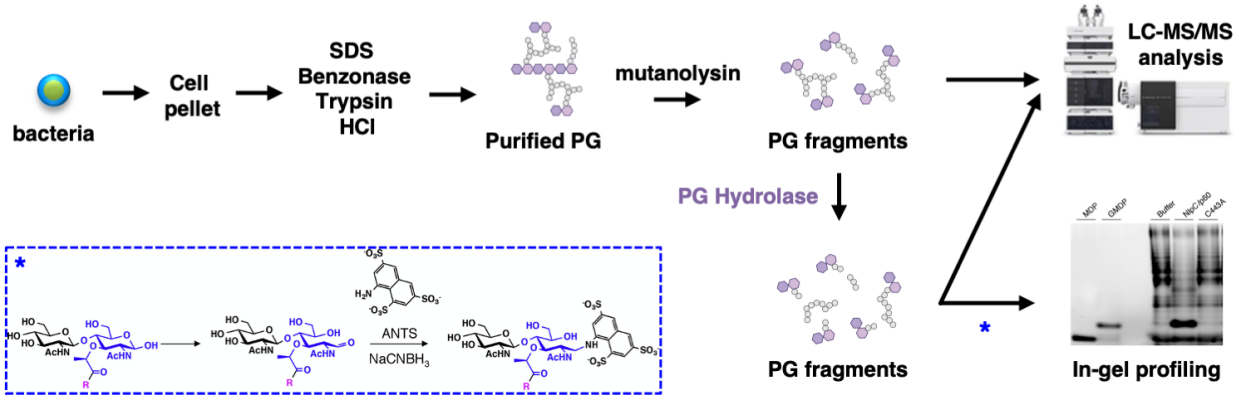


Figure 2.2. Schematics of in-gel and LC-MS profiling assays of NlpC/p60 hydrolase activity. Blue inset represents the ANTS-labeling of reducing end of MurNAc via reductive amination with sodium cyanoborohydride ($NaCNBH_3$). ANTS, 8-aminonaphthalene-1,3,6-trisulfonic acid, disodium salt (Invitrogen).

Biochemical characterization of NlpC/p60 PG hydrolase activity

PG profiling of engineered SagA+ bacteria highlighted the potential PG remodeling ability of SagA in its host strain, but elucidation of its precise enzymatic function was necessary to understand the biochemical mechanism(s) that govern SagA-mediated protection. To that end, low and high-resolution methods were optimized to evaluate the biochemical activity of NlpC/p60 PG hydrolases *in vitro* (*40*, *41*). These methods entail the incubation of purified protein with known soluble PG fragments followed by in-gel profiling or LC-MS analysis (**Fig. 2.2**). For the low-resolution in-gel profiling assay using the ANTS fluorophore to covalently label soluble PG fragments via reductive amination (*97*), muropeptide controls were custom synthesized or obtained commercially. For the high-resolution assay, we used an Agilent 1200 LC with a C18

column connected in-line with a Thermo LTQ-Orbitrap mass spectrometer equipped with an electrospray ionization source (*40*, *41*).

To characterize the PG hydrolase activity of SagA, recombinant full-length SagA and SagA-NlpC/p60 constructs were purified from *E. coli* supernatant and cytoplasm, respectively, as detailed previously (**Fig. 2.3A,B**) (*40*). The in-gel profiling assay showed that SagA-NlpC/p60 could not cleave *E. faecium* Com15 PG unless pre-digested and solubilized by mutanolysin *in vitro* (**Fig. 2.3C**) (*40*). Furthermore, full-length SagA and SagA-NlpC/p60 only hydrolyzed pre-digested *E. faecium* Com15 and *E. faecalis* OG1RF PG (L-Lys-type), but not pre-digested *E. coli* PG (mDAP-type) *in vitro* (**Fig. 2.3D**) (*40*). The product band exclusive to the SagA-treated samples co-migrated with synthesized GlcNAc-MurNAc-L-Ala-D-isoGln (GlcNAc-MDP) (*40*). These results suggest that SagA is a L-Lys-type D,L-endopeptidase that can accommodate different PG cross-bridges (Asp/Asn *Efm* Com15 and Ala-Ala in *Efs* OG1RF), and that the SagA-NlpC/p60 domain alone is sufficient for enzymatic activity (*40*). Next, we pursued LC-MS profiling of pre-digested *E. faecalis* OG1RF PG after incubation with SagA constructs to identify the precise substrates and products of SagA PG hydrolase activity *in vitro*, which would in turn uncover the cleavage bond specificity of SagA. *E. faecalis* OG1RF PG was used as a heterogenous substrate mixture since *E. faecalis* OG1RF does not express nor secrete a SagA ortholog and we assumed it would not contribute to high background of SagA enzymatic products compared to the native *E. faecium* Com15 PG. LC-MS profiling of the resulting SagA-generated muropeptides revealed SagA as a D,L-endopeptidase, as it preferentially cleaved cross-linked L-Lys PG fragments with penta- and tetrapeptide stems *in vitro* (**Fig. 2.3E**, **Table 2.1 and 2.2**) (*40*). The identity of the expected, common D,L-endopeptidase product fragment under these conditions, GlcNAc-MDP, was confirmed by MS/MS analysis (*40*). In addition, MS/MS analysis of muropeptide products after incubation of a purified, crosslinked PG fragment from *E. faecium* Com15 with SagA-NlpC/p60 showed that SagA-NlpC/p60 cleaves within the tetrapeptide stem as opposed to tripeptide stem to generate GMDP (*40*). Alanine mutants of the conserved cysteine (C443A) of full-length SagA and SagA-NlpC/p60 were inactive under these conditions and confirmed that the observed D,L-endopeptidase activity is dependent on the catalytic cysteine (**Fig. 2.3E**, **Table 2.1 and 2.2**).

Figure 2.3. Full-length SagA and SagA-NlpC/p60 exhibit D,L-endopeptidase activity towards cross-linked, Lys-type PG with tetra- and pentapeptide stems. (**A**) Schematic representation of SagA and SagA-NlpC/p60 domain. (**B**) Stain-free SDS-PAGE of full-length SagA and NlpC/p60 construct. (**C**) ANTS visualization of isolated peptidoglycan from *E. faecium* treated with buffer or purified SagA-NlpC/p60 (intact *E. faecium*) and peptidoglycan treated with mutanolysin followed by buffer or purified SagA-NlpC/p60 (predigested *E. faecium*). Muropeptides were isolated as previously described (*98*), dried, labeled with ANTS, separated by native PAGE and then visualized by UV. ANTS-labeled synthetic fragments MDP (MurNAc-L- Ala-D-iGln), GlcNAc-MDP (GlcNAc-MurNAc-L-Ala-D-iGln) were analyzed in parallel for comparison. Bold asterisks indicate product formation from overnight incubation of peptidoglycan digested with purified SagA construct. (**D**) ANTS visualization of isolated peptidoglycan from *E. faecium* (*Efm*), *E. faecalis* (*Efs*) and *E. coli* (*Ec*) treated with mutanolysin followed by each purified SagA construct (SagA-FL: full-length SagA), NlpC/p60: truncated NlpC/p60 domain) grown from E. coli. (**E**) LC-MS analysis of peptidoglycan isolated from E. faecalis digested by mutanolysin followed by buffer, purified full-length SagA, full-length SagA-C443A, NlpC/p60 domain, and NlpC/p60-C443A. Enzymatic products (a,b,c,d,e) were generated only after incubation with purified full-length SagA and purified NlpC/p60 domain as shown in Table 2.2. The in-gel profiling analysis in panels (C) and (D) were performed by Dr. Byungchul Kim.

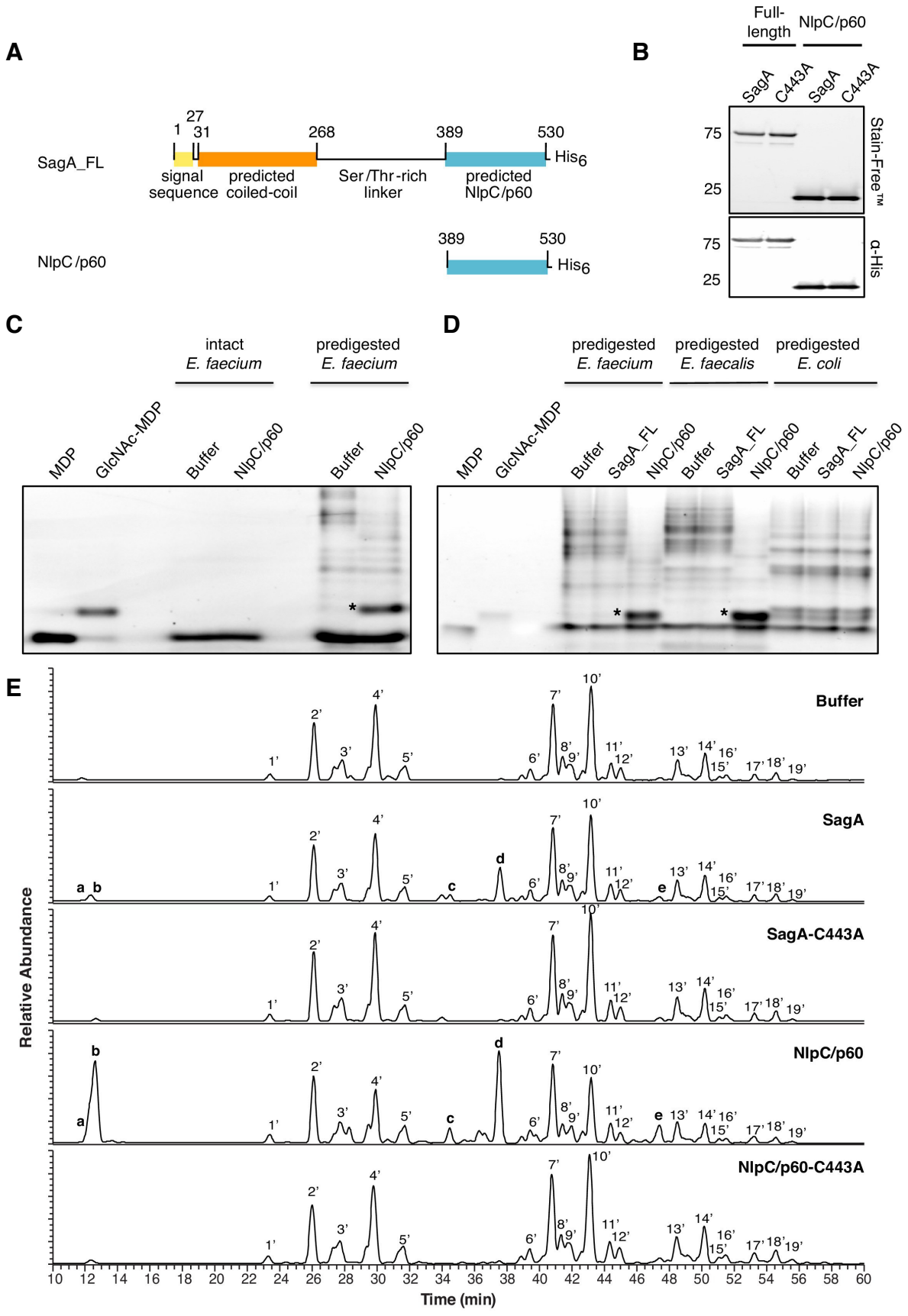

A
SagA_FL
1
27
31
268
389
530
His6
signal sequence
predicted coiled-coil
Ser/Thr-rich linker
predicted NlpC/p60
NlpC/p60
389
530
His6
B
Full-length
NlpC/p60
SagA
C443A
SagA
C443A
75
25
Stain-Free™
75
25
α-His
C
intact E. faecium
predigested E. faecium
MDP
GlcNAc-MDP
Buffer
NlpC/p60
Buffer
NlpC/p60
D
predigested E. faecium
predigested E. faecalis
predigested E. coli
MDP
GlcNAc-MDP
Buffer
SagA_FL
NlpC/p60
Buffer
SagA_FL
NlpC/p60
Buffer
SagA_FL
NlpC/p60
E
Buffer
SagA
SagA-C443A
NlpC/p60
NlpC/p60-C443A
Relative Abundance
Time (min)

Table 2.1. Molecular mass and composition of muropeptides from *E. faecalis* OG1RF V1.

Peak[a]	RT (min)	calculated $[M+H]^+$	observed $[M+H]^+$	Proposed structure[b]
1	25.3	897.44	897.44	GM-tetra
2	27.9	968.48	968.48	GM-tri (AA)
3	29.8	969.46	969.46	GM-tri (AA), Gln/Glu[d]
4	31.9	1110.55	1110.55	GM-penta (AA)
5	33.8	1111.53	1111.53	GM-penta (AA), Gln/Glu[d]
6	41.5		2070.95	ND[c]
7	43.1	1989.98	1989.98	2GM-tri (AA) - tetra (AA)
8	43.8		2190.12	ND[c]
9	44.3	1991.95	1991.97	2GM-tri (AA) - tetra (AA), Gln/Glu x2[d]
10	45.6	2132.06	2133.06	2GM-penta (AA) - tetra (AA)
11	47.0	2134.03	2134.05	2GM-penta (AA) - tetra (AA), Gln/Glu x2[d]
12	47.6	2134.03	2134.05	2GM-penta (AA) - tetra (AA), Gln/Glu x2[d]
13	51.0	3012.47	3012.49	3GM-tri (AA) - tetra (AA) - tetra (AA), Gln/Glu
14	52.8	3154.55	3154.57	3GM-penta (AA) - tetra (AA) - tetra (AA), Gln/Glu[d]
15	53.8	3155.53	3155.55	3GM-penta (AA) - tetra (AA) - tetra (AA), Gln/Glu x2[d]
16	54.2	3155.53	3155.55	3GM-penta (AA) - tetra (AA) - tetra (AA), Gln/Glu x2[d]
17	55.9	4033.98	4033.99	4GM-tri (AA) - tetra (AA) - tetra (AA) - tetra (AA), Gln/Glu[d]
18	57.3	4176.05	4176.07	4GM-penta (AA) - tetra (AA) - tetra (AA) - tetra (AA), Gln/Glu[d]
19	58.4	4177.03	4177.05	4GM-penta (AA) - tetra (AA) - tetra (AA) - tetra (AA), Gln/Glu x2[d]
20		4178.02	4178.04	4GM-penta (AA) - tetra (AA) - tetra (AA) - tetra (AA), Gln/Glu x3[d]

a. Peak numbers refer to Figures 2.3 and 2.4.

b. GM, disaccharide (GlcNAc-MurNAc); 2GM, disaccharide-disaccharide (GlcNAc-MurNAc-GlcNAc-MurNAc); 3GM, disaccharide-disaccharide-disaccharide (GlcNAc-MurNAc-GlcNAc-MurNAc- GlcNAc-MurNAc); GM-Tri, disaccharide tripeptide (L-Ala-D-iGln-L-Lys); GM-Tetra, disaccharide tetrapeptide (L-Ala-D-iGln-L-Lys-D-Ala); GM-Penta, disaccharide pentapeptide (L-Ala-D-iGln-L-Lys-D-Ala -D-Ala).

c. ND: Precise structure unknown.

d. The assignment of the amide and the hydroxyl functions to either peptide stem is arbitrary.

Table 2.2. Molecular mass and composition of enzymatic products from incubation of *E. faecalis* OG1RF PG with purified SagA constructs V1.

Peak[a]	RT (min)	calculated $[M+H]^+$	observed $[M+H]^+$	Proposed structure[b]
a	12.0	431.26	431.26	Tri (AA)
b	12.1	698.31	698.31	GM-di (AA)
c	34.4	1310.69	1310.69	GM-tri-di (AA x2)
d	37.4	1452.77	1452.77	GM-tetra-tri (AA x2)
e	47.3	2474.27	2474.27	GM-tetra-GM-tetra-tri (AA x3)

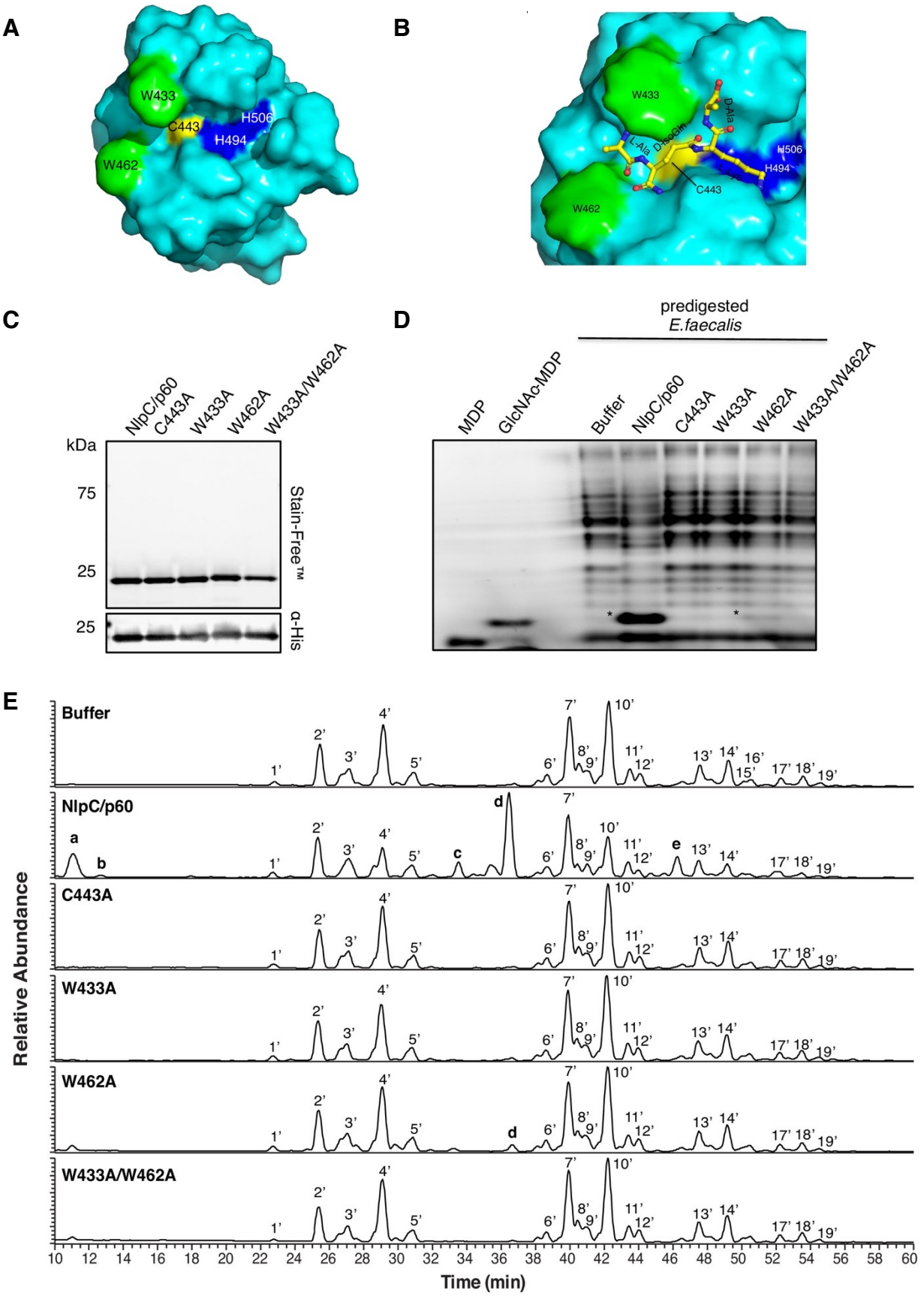
a. Peak numbers refer to Figures 2.3 and 2.4.
b. GM, disaccharide (GlcNAc-MurNAc); GM-di, disaccharide dipeptide (L-Ala-D-iGln); GM-Tri, disaccharide tripeptide (L-Ala-D-iGln-L-Lys); GM-Tetra, disaccharide tetrapeptide (L-Ala-D-iGln-L-Lys-D-Ala).

To characterize the key amino acid residues necessary for SagA PG hydrolytic activity, our laboratory embarked on structural and biochemical studies of recombinant SagA-NlpC/p60 (*40*). The structure of the SagA-NlpC/p60 domain (PDB 6B8C) was solved to 2.4 Å resolution by X-ray crystallography and displayed a conserved arrangement of the putative catalytic triad similar to known NlpC/p60 domain structures (**Fig. 2.4A**) (*40*). Molecular modeling studies using Schödinger produced a model of SagA-NlpC/p60 with a bound tetrapeptide ligand in the surface-exposed, putative substrate-binding groove (**Fig. 2.4B**) (*40*). We observed that Trp433 and Trp462 protrude prominently from the SagA-NlpC/p60 structure and may serve as clamps for binding potential peptidoglycan substrates (*40*). Alanine mutants of Cys443, Trp433, and Trp462 abolished the D,L-endopeptidase activity of SagA-NlpC/p60, thus supporting our hypothesis that the Trp residues may play a crucial role in SagA-NlpC/p60 PG hydrolase activity (**Fig. 2.4C-E**, **Table 2.1 and 2.2**) (*40*). Trp433 is conserved in other NlpC/p60 proteins and suggests that its requirement for PG hydrolase activity may be conserved as well (*40*).

Our optimized biochemical methods enabled the discovery of SagA as a D,L-endopeptidase and we next sought to uncover the link between SagA PG hydrolase activity and SagA-mediated protection. We determined that SagA produces GlcNAc-MDP *in vitro*, which is similar to other small muropeptides that are known to trigger host immune responses through the NOD 1 and 2 pattern recognition receptors (*85–87*). Interestingly, GlcNAc-MDP and a related cross-linked muropeptide product isolated from enzymatic reactions with SagA specifically activated NOD2 in mammalian cells in a dose-dependent manner, as did synthetic GlcNAc-MDP and commercial MDP (*40*). The finding that SagA-generated muropeptides activate NOD2 *ex vivo* informed our proposed model of SagA protective activity: SagA D,L-endopeptidase activity may generate small muropeptides containing MDP that may in turn bind pattern recognition receptors such as NOD2 for downstream activation of host immune pathways *in vivo* (*40*). To determine if SagA PG hydrolase activity is required for protection *in vivo*, strains of the probiotic bacterium *L. plantarum* were engineered to heterologously express different SagA constructs and evaluated in a mouse model of *C. difficile* infection (*40*). Remarkably, only the *L. plantarum* strain engineered to express and secrete wild type SagA (*Lpl*-SagA) was protective, but not strains expressing secretion deficient (*Lpl*-ΔSS) or catalytically inactive (*Lpl*-C443A) SagA. Although heterologous expression of SagA in these *L. plantarum* strains did not alter their cell wall composition, these

results affirm that secretion of catalytically active SagA is sufficient to confer protection against enteric pathogens *in vivo*, and further highlight the therapeutic potential of SagA (*40*).

Figure 2.4. Trp433 and Trp462 may serve as clamps for binding potential PG substrates and are crucial for SagA-NlpC/p60 PG hydrolase activity. (A) Surface representation of the SagA-NlpC/p60 structure (PDB entry: 6B8C). Catalytic triad of Cys443 (yellow), His494 (blue), His506 (blue), Trp433 (green), and Trp462 (green) are highlighted. (**B**) Model of tetrapeptide peptidoglycan fragment (L-Ala-D-isoGln-L-Lys-D-Ala) bound to the SagA-NlpC/p60 structure. (**C**) Stain-Free (Bio-Rad) SDS-PAGE of SagA-NlpC/p60 domain, NlpC/p60-C443A, NlpC/p60-W433A, NlpC/p60-W462A, and NlpC/p60-W433A/W462A constructs. (**D**) ANTS visualization of peptidoglycan treated with mutanolysin followed by buffer or purified SagA-NlpC/p60 constructs (predigested *E. faecalis*). Bold asterisks indicate product formation from overnight incubation of peptidoglycan digested with purified SagA constructs. (**E**) LC-MS analysis of isolated peptidoglycan from *E. faecalis* were digested with mutanolysin, followed by incubating with buffer, NlpC/p60, C443A, W433A, W462A, and W433A/W462A mutants at 37 °C for 16 hr. Enzymatic products (a, b, c, d, and e) from incubation of *E. faecalis* peptidoglycan by purified NlpC/p60 and NlpC/p60-W462A mutant were observed as shown in Table 2.2. Dr. Byungchul Kim solved the SagA-NlpC/p60 structure in panel (A) and performed the molecular docking study in panel (B).


A
W433
C443
H506
H494
W462
B
W433
D-Ala
L-Ala
D-Glu
Dap
H506
H494
C443
W462
C
NlpC/p60
C443A
W433A
W462A
W433A/W462A
kDa
75
25
25
Stain-Free™
α-His
D
predigested
E.faecalis
MDP
GlcNAc-MDP
Buffer
NlpC/p60
C443A
W433A
W462A
W433A/W462A
E
Buffer
NlpC/p60
C443A
W433A
W462A
W433A/W462A
Relative Abundance
Time (min)

The methodologies that we have described in this chapter enabled the preparation of SagA and the biochemical characterization of its enzymatic activity. We optimized protocols to prepare SagA constructs from *E. coli* culture supernatant or cytoplasm, and to isolate PG from Gram-positive bacteria on a large scale. Low and high resolution assays were then employed to study the cell wall composition of SagA+ bacteria and to identify SagA as a D,L-endopeptidase with a substrate preference towards L-Lys type PG. Molecular docking studies highlighted structural features in the catalytic SagA-NlpC/p60 domain that were determined to be important for its PG hydrolase activity using our enzymatic assays. Ultimately, these biochemical studies greatly improved our understanding of the role SagA PG hydrolase activity may play in the observed protective ability of SagA *in vivo*. Moreover, these methods should be generally applicable for the preparation of NlpC/p60 proteins from other bacterial species and the biochemical characterization of their enzymatic activities. The purified, recombinant proteins that can now be obtained for many isoforms and constructs can be used for many other *in vitro* applications. *In vitro* assays offer powerful tools to study PG hydrolases. They can be used to identify and analyze substrates and determine their specific activities. The present chapter can serve as a reference for any such studies employing recombinant NlpC/p60 proteins for studying function and regulation of this unique family in PG remodeling.

Acknowledgements

We thank Dr. Kavita Rangan for her foundational studies with SagA. We thank Dr. Joseph P Fernandez, Dr. Brian D Dill and Dr. Henrik Molina for assistance with peptidoglycan analysis (The Rockefeller University Proteomics Resource Center). Dr. Byungchul Kim also thanks Uhn-Soo Cho (University of Michigan, Ann Arbor), Yehuda Goldgur (MSKCC) for assistance in SagA-NlpC/p60 structural determination.

Chapter 3. *Enterococcus* NlpC/p60 PG hydrolase SagA localizes to sites of cell division and only requires catalytic dyad for protease activity

Introduction

The findings regarding SagA PG hydrolase activity detailed in the previous chapter have improved our understanding of the mechanisms that govern SagA-mediated protection, yet the localization of SagA in *E. faecium* (*Efm*) and its ligand-protein interactions remain elusive. A major constraint toward elucidating the precise function of SagA is its essential role in *Efm* division and growth, as complementation of *sagA* was required for disruption of the chromosomal gene (*99*). Due to the essentiality of *sagA*, we cannot ascertain if impaired cell division phenotypes arise in *Efm* upon disruption of expression of functional SagA similar to NlpC/p60-mutant strains of other bacteria, as discussed in Chapter 1. Furthermore, evidence exists for auto-regulation in NlpC/p60-type (*27*) and related (*89*) PG hydrolases. The constraints regarding the genomic dissection of *sagA* described here hampers our ability to verify whether the unique domain architecture of SagA begets auto-inhibition of SagA PG hydrolase activity. Also, the structural basis of the direct contacts between PG substrates and SagA is key to understanding how the D,L-endopeptidase activity of *Efm* SagA functions *in vivo*, possibly on other bacteria or dietary PG fragments. Obtaining a muropeptide-bound structure of SagA to reveal these significant contacts, however, has proven challenging; the preferred muropeptide substrates of SagA are large and cross-linked, and purification of these fragments for crystallization trials is limiting.

To address these outstanding questions, we embarked on imaging studies of enterococci and SagA as well as additional biochemical and structure activity studies. Using a fluorescent analog of D-alanine to probe for actively growing cells as well as transmission electron microscopy (TEM), we showed that SagA contributes to a tri-septal structure in dividing SagA+ enterococci cells. Fluorescence microscopy of tetracysteine (Cys4)-tagged SagA and its truncation constructs overexpressed in *Efm* Com15 revealed that SagA is localized at the cell division site via its coiled-coil N-terminus, which suggests that SagA is involved in PG remodeling during cell division. Moreover, comparison of the biochemical activity of recombinant full-length SagA and the SagA-NlpC/p60 showed that the N-terminus does not appear to have an auto-inhibitory function. Molecular docking studies of the SagA-NlpC/p60 structure with a representative monomeric pentapeptide PG fragment revealed key features of the putative SagA-NlpC/p60 substrate-binding groove. Subsequent alanine-scanning mutagenesis of potential key residues of SagA-NlpC/p60 *in vitro* afforded additional insight to SagA D,L-endopeptidase activity in enterococci during cell division. In addition, we discovered that full-length SagA orthologs from several other *Enterococcus* species exhibit D,L-endopeptidase activity *in vitro* via a Cys-His catalytic dyad. Our investigation of SagA localization and biochemical activity highlights how SagA is involved in remodeling the cell wall in *Enterococcus*, which is important for understanding its roles in bacterial cell division and modulation of host immune responses. These results were reported in this publication (*42*). Dr. Ti-Yu Lin performed all imaging studies. Yadyvic Estrella aided with in-gel profiling of PG hydrolase activity of SagA-NlpC/p60 mutants. Dr. Byungchul Kim aided in designing mutants of SagA-NlpC/p60, and Dr. Henrik Molina with sample submission for LC-MS and analysis.

Results and Discussion

According to the template model of PG assembly in enterococci, actively dividing cells generate nascent PG chains at the septum with non-crosslinked pentapeptide stems, which are

trimmed in mature PG by an L,D-carboxypeptidase (*100–102*). To determine if SagA plays a role in PG remodeling, we stained *Efm* cells with HADA (*103*), a D-alanine analog containing a fluorescent coumarin moiety that has been used to probe newly synthesized PG in many bacterial species (**Fig. 3.1A**). We found that dividing *Efm* cells show a tri-septal HADA staining pattern labeling the septa and equators (**Fig. 3.1B-D**). In contrast, *Enterococcus faecalis* (*Efs*), which does not have a *sagA* gene, only exhibited HADA staining at the division septa. However, an *Efs* OG1RF strain expressing SagA (*Efs-sagA* (*69*)) revealed HADA staining at both septa and equators, suggesting a role of SagA in PG remodeling during cell division. The tri-septal structure was also confirmed by transmission electron microscopy (TEM) (**Fig. 3.1E)**. These indentations at the equators indicate that daughter cells of SagA+ enterococci commence cell division before separation from one another. Interestingly, HADA staining (**Fig. 3.2A**) and TEM (**Fig. 3.2B**) of various enterococci species encoding SagA orthologs revealed that SagA may broadly contribute to this tri-septal structure. Other *Efs* strains (775 and V583) do not encode SagA-like NlpC/p60 orthologs, but also display faint and pronounced HADA staining at the equators and division septa, respectively (**Fig. 3.2A**), suggesting the focused HADA staining may be specific to *Efs* OG1RF. Whether the NlpC/p60 domain of SagA plays a structural or catalytic role in this tri-septal phenotype is unclear, as overexpression of the catalytically inactive SagA C443A mutant appeared to be toxic to *Efm* Com15 cells and yielded no colonies, and chromosomal insertion of SagA C443A in *Efs* OG1RF yielded low or undetectable levels of protein expression (*40*). Nevertheless, due to the faster growth rate and abundance of non-crosslinked PG in *Efm* Com15 compared to *Efs* OG1RF (*40*), it is possible that SagA simultaneously contributes to cell division and separation.

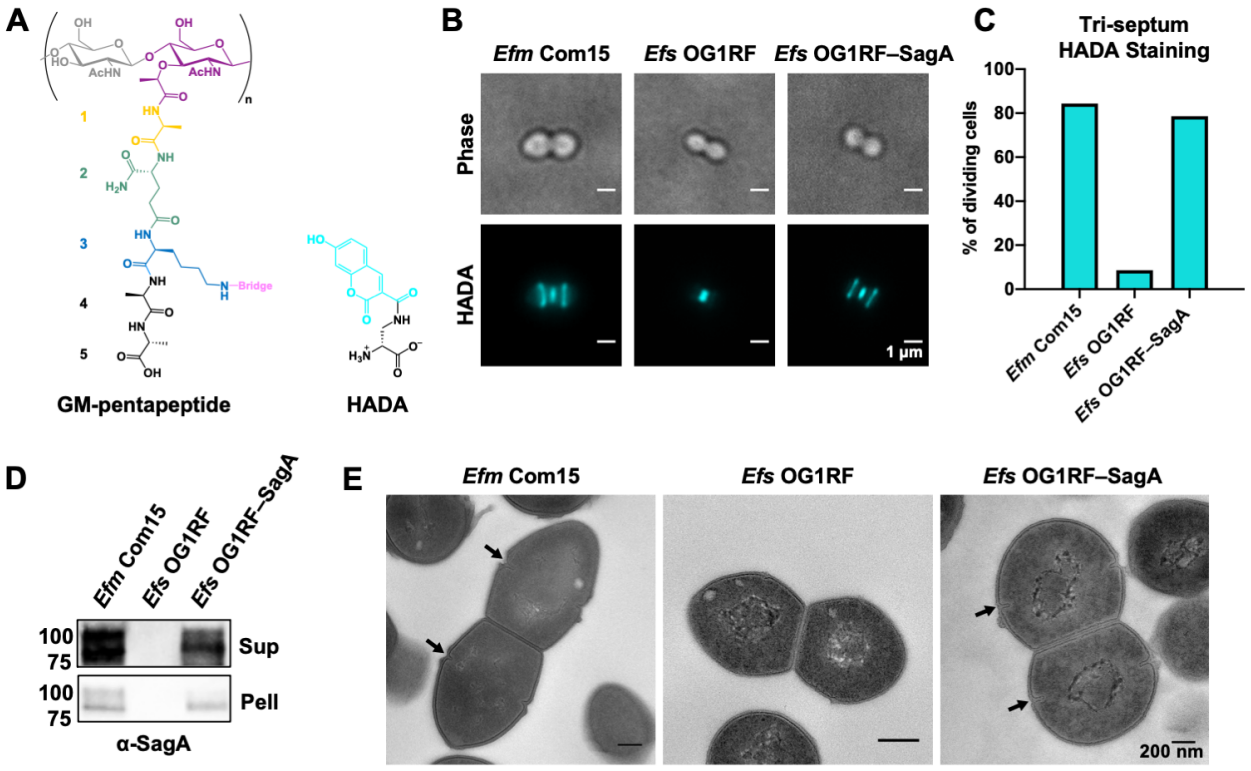


Figure 3.1. SagA contributes to tri-septum HADA staining in dividing cells. (**A**) Chemical structures of GM-pentapeptide (GlcNAc-MurNAc-L-Ala-D-isoGln-L-Lys-D-Ala-D-Ala, left) and HADA (right). For GM-pentapeptide, the position of each amino acid in the peptide stem is numbered on the left. The covalent bridge on the L-Lys is most often D-Asn/Asp in SagA+ enterococci, or L-Ala-L-Ala in *Enterococcus faecalis* strains. (**B**) HADA staining of enterococci. Scale bar = 1 µm. (**C**) Percentage of dividing cells with tri-septum HADA staining. 300 cells counted for each strain. (**D**) Expression of SagA confirmed by α-SagA immunoblotting of culture media supernatant (Sup) and cell pellet lysate (Pell). (**E**) TEM imaging reveals obvious tri-septum indentations (arrows) in SagA+ enterococci (*Efm* Com15 and *Efs* OG1RF–SagA), but not in SagA- enterococci (*Efs* OG1RF). Scale bar = 200 nm. *Efm* Com15, *Enterococcus faecium* Com15; *Efs* OG1RF, *Enterococcus faecalis* OG1RF; *Efs* OG1RF–SagA, *Enterococcus faecalis* OG1RF–SagA chromosomal insertion. The imaging and protein expression analysis shown here were the work of Dr. Ti-Yu Lin.

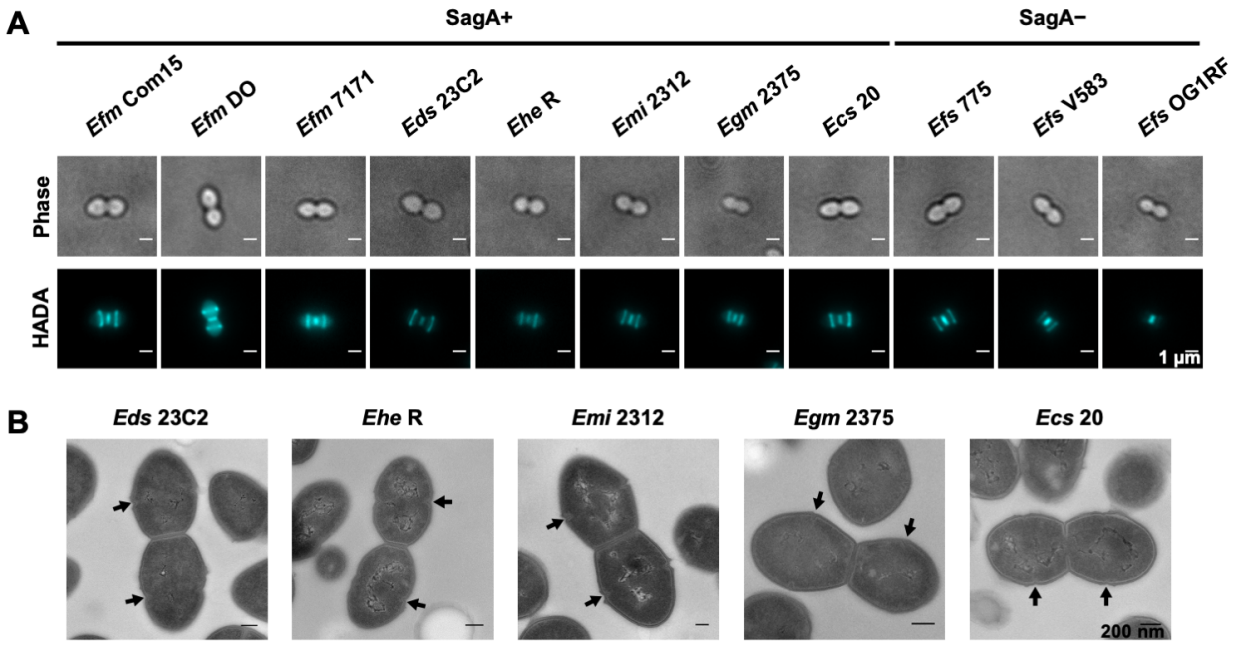


Figure 3.2. Tri-septum phenotype in dividing cells of SagA+ enterococci species. (**A**) HADA staining of SagA+ enterococci and SagA- enterococci. Scale bar = 1 µm. (**B**) TEM imaging reveals obvious tri-septum indentations (arrows) in SagA+ enterococci. Scale bar = 200 nm. *Efm* Com15, *Enterococcus faecium* Com15; *Efm* DO, *Enterococcus faecium* DO; *Efm* 7171, *Enterococcus faecium* 7171; *Eds* 23C2, *Enterococcus durans* 23C2; *Ehe* R, *Enterococcus hirae* R; *Emi* 2312, *Enterococcus mundtii* NCDO 2312; *Egm* 2375*, Enterococcus gallinarum* NCDO 2375; *Ecs* 20, *Enterococcus casseliflavus* 20; *Efs* 775, *Enterococcus faecalis* 775; *Efs* V583, *Enterococcus faecalis* V583; *Efs* OG1RF, *Enterococcus faecalis* OG1RF. Dr. Ti-Yu Lin performed the fluorescence and transmission-electron microscopy shown here.

To study the subcellular localization of SagA, we inserted a tetracysteine (Cys4) tag at its C-terminus and ectopically expressed it in *Efm* cells (**Fig. 3.3A**). We found no significant growth defects in cells overexpressing SagA compared to control cells containing the empty vector (**Fig. 3.3B**). We then visualized the protein fusion using the biarsenical dye ReAsH, which covalently binds the Cys4 amino acid motif (*104*, *105*). We found that SagA localizes at the septa and equators in dividing *Efm* cells (**Fig. 3.3C**), further suggesting a regulatory role of SagA in bacterial cell division. We also found SagA C-terminal truncation ($SagA^{\Delta C}$) and SagA N-terminus ($SagA^{Nter}$) localize at the septa and equators using the Cys4-tagging and ReAsH labeling (**Fig. 3.4**), suggesting that the coiled-coil N-terminal domain is sufficient for the subcellular localization of SagA. The differential colocalization of these two constructs compared to wild-type may be attributed to lower expression in *Efm* (**Fig. 3.4D**) and may also indicate that correct folding of the full-length protein is important for targeting septal PG.

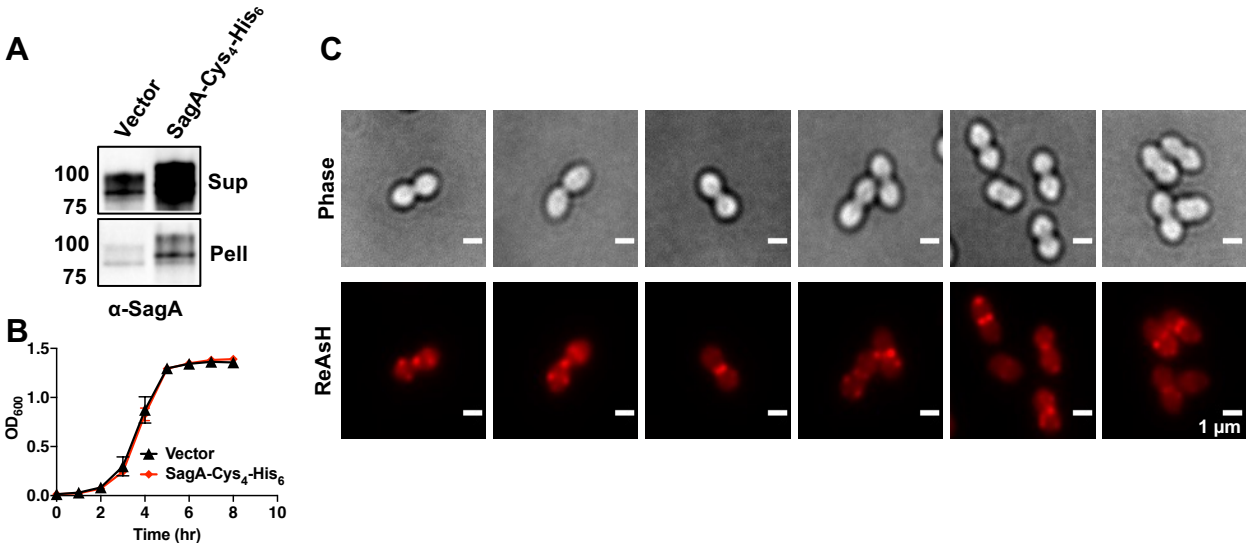


Figure 3.3. SagA is localized at the cell division site of *E. faecium* Com15. (**A**) Expression of SagA confirmed by α-SagA immunoblotting of culture media supernatant (Sup) and cell pellet lysate (Pell). (**B**) Growth curve of *E. faecium* Com15 strains (n = 6). (**C**) Fluorescence microscopy of *E. faecium* Com15 overexpressing SagA-Cys_4-His_6. Scale bar = 1 μm. Vector, *E. faecium* Com15 transformed with empty pAM401 overexpression vector; SagA-Cys_4-His_6, *E. faecium* Com15 transformed with pAM401:SagA-Cys_4-His_6 overexpression vector. The protein expression, cell growth, and imaging analysis shown here were performed by Dr. Ti-Yu Lin.

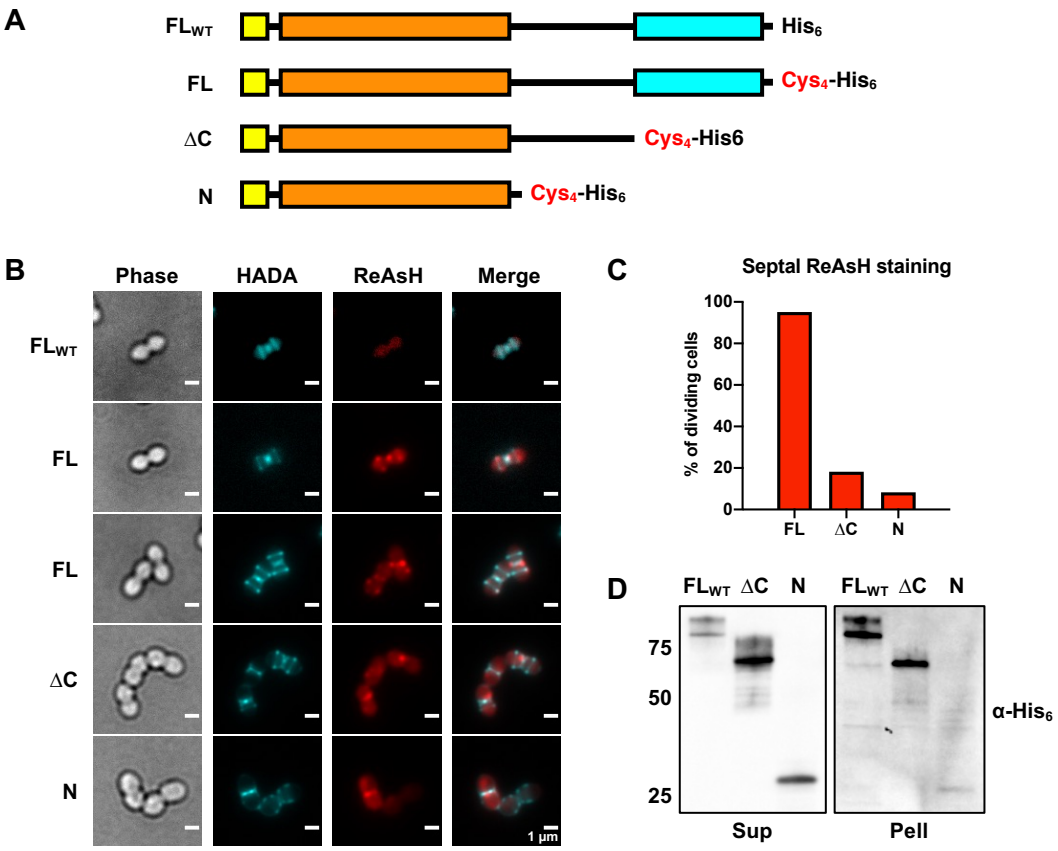


Figure 3.4. Localization of SagA truncation constructs at the cell division site of *E. faecium* Com15. (**A**) Domain architectures of SagA constructs cloned into the pAM401 overexpression vector. Signal sequence (yellow), coiled-coil (orange), NlpC/p60 (cyan) domains are highlighted. (**B**) Fluorescence microscopy of *E. faecium* Com15 over-expressing SagA truncation constructs. Scale bar = 1 μm. (**C**) Percentage of dividing cells with septal ReAsH staining. 300 cells counted for each strain. (**D**) Expression of SagA constructs confirmed by α-His_6 immunoblotting of culture media supernatant (Sup) and cell pellet lysate (Pell). The imaging and protein expression analysis shown here were the work of Dr. Ti-Yu Lin.

Protein fold analysis using Phyre2 (*106*) shows that the SagA N-terminus bears 99% secondary structure homology to the inactive, auto-inhibitory module (*89*) of the essential PG hydrolase PcsB (*107*) from *Streptococcus pneumoniae* (*89*). Structural studies suggest PcsB adopts an interlocking, autoinhibitory dimeric structure where the N-terminal coiled-coil domain of one monomer interacts with the catalytic C-terminus of the other in an inactive conformation (*89*). This homology between the N-termini of SagA and PcsB suggests that the SagA N-terminus may have a regulatory function aside from contributing to the subcellular localization of SagA. However, intact protein analysis by MALDI-TOF mass spectrometry revealed that full-length SagA exists in a monomeric state (**Appendix 3.1**). Additionally, the unusual linker length of SagA (120 amino acids) raises the possibility for the N and C-termini to interact. We were unable to ectopically express an N-terminal truncation of SagA ($SagA^{\Delta Nter}$) in *Efm* Com15 for imaging studies. Therefore, to discern the role of the SagA N-terminus in SagA D,L-endopeptidase activity, we compared the *in vitro* enzymatic activity of recombinant full-length SagA and SagA-NlpC/p60

(**Fig. 3.5A**) on soluble muramidase-digested *Efs* PG. Full-length SagA and SagA-NlpC/p60 demonstrated comparable D,L-endopeptidase activity towards model PG fragments similar to nascent and template PG in *Efm* (**Fig. 3.5B-D and 3.6, Table 3.1 and 3.2**) (*100–102*). Full-length SagA D,L-endopeptidase activity more readily produced smaller cross-linked fragments ($Product_e$ in **Fig. 3.5D** and $Product_f$ in **Fig. 3.6C**), but there was no significant difference in overall GlcNAc-MurNAc-dipeptide production compared to the SagA-NlpC/p60 domain alone (**Fig. 3.6E**), suggesting that the N-terminus does not play an autoinhibitory role, but perhaps a scaffolding and targeting function in enterococci. Our results also suggest additional enzymes such as muramidases or other glycosidases in *Efm* are needed to generate smaller PG fragments that can then be cleaved by SagA.

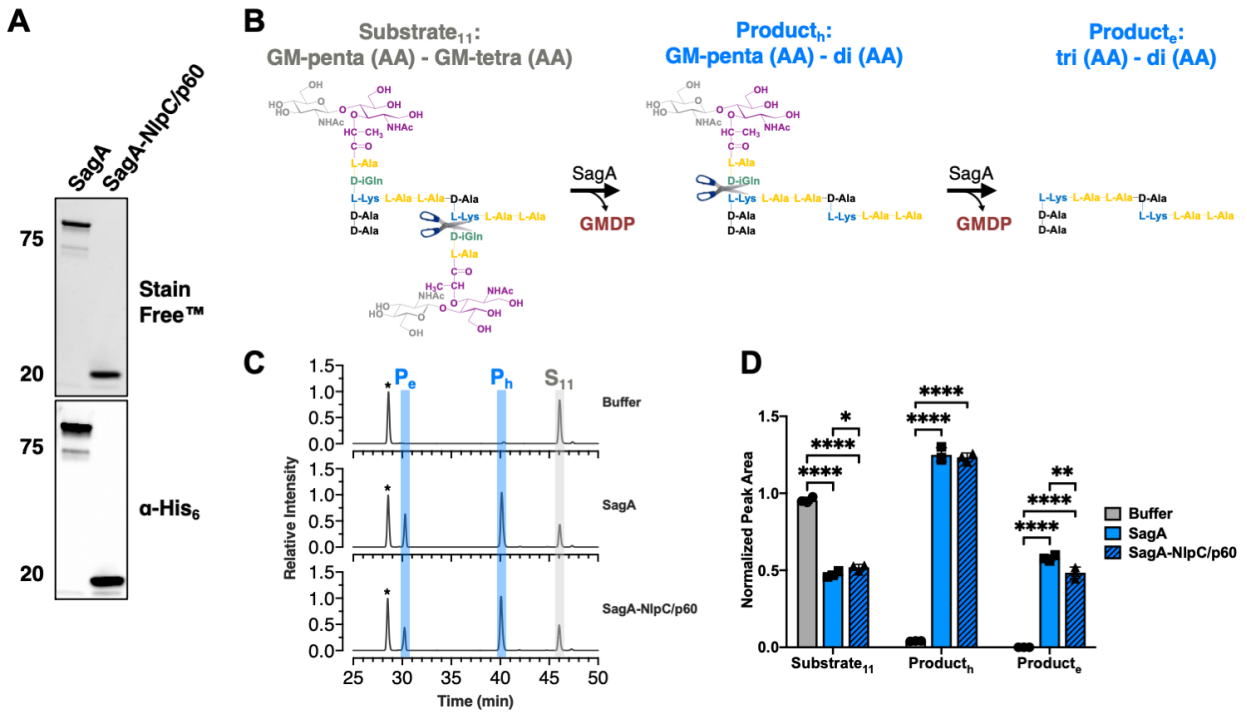


Figure 3.5. Comparison of *in vitro* activity of SagA truncation constructs. (**A**) Stain-Free (Bio-Rad) SDS-PAGE and α-His_6 immunoblotting of SagA constructs purified from the cell-free lysate of BL21(DE3)-RIL *Escherichia coli*. (**B**) Schematic of *in vitro* SagA D,L-endopeptidase activity. $Substrate_{11}$ represents the most abundant of the soluble, cross-linked muropeptides released following mutanolysin digestion of intact *E. faecalis* OG1RF PG. Iterative cleavage of $Substrate_{11}$ by SagA generates GlcNAc-MurNAc-dipeptide (GMDP) and two fragments, $Product_h$ and $Product_e$. $Substrate_{11}$ and $Product_h$ are illustrated with reduced anomeric carbon of MurNAc. Numbers and letters in subscript denote the LC-MS peak labels described in Table 3.1 and 3.2, respectively. (**C**) LC-MS analysis of soluble mutanolysin-digested *E. faecalis* OG1RF PG incubated with buffer, SagA, or SagA-NlpC/p60 at 37˚C for 16 hr. Data are shown as extracted ion chromatograms. The intensity of extracted ion abundance was normalized to the unchanging GM-tripeptide (AA) peak marked with an asterisk. Colored columns indicate PG substrate and products as described in panel (B). See Figure 3.6 for the corresponding total ion chromatograms. (**D**) Quantification of substrate depletion and product formation by SagA truncation constructs *in vitro* (n = 3). For each condition, peak areas were recorded from extracted ion chromatograms of substrate and products and normalized to that of the unchanging GM-tri-(AA). Data were analyzed with one-way ANOVA and Tukey's multiple comparisons post-hoc test. Centerline, average. Error bar, SD ****$P < 0.0001$, ***$P < 0.0002$, **$P < 0.0021$. Comparisons with no asterisk had $P > 0.05$ and were not considered significant.

Figure 3.6. Further comparison of *in vitro* activity of SagA truncation constructs. (**A**) LC-MS analysis of soluble mutanolysin-digested *E. faecalis* OG1RF PG incubated with buffer, SagA, or SagA-NlpC/p60 at 37˚C for 16 hr. Data are shown as total ion chromatograms corresponding to Figure 3.5C. The peak labels and identities are indicated in Table 3.1 and 3.2. The intensity of total ion abundance was normalized to the unchanging GM-tripeptide (AA) peak (2). Peak labels for confirmed SagA D,L-endopeptidase products are in bold with colors representing the different enzymatic reactions illustrated in panels (B-E). Peak labels for fragments unique to SagA-treated samples, but of unknown structure, are in bold italics. (**B-E**) Schematics (left) and quantifications (right) of *in vitro* SagA D,L-endopeptidase activity (n = 3). The substrates represent soluble *E. faecalis* OG1RF PG fragments released following mutanolysin digestion of intact *E. faecalis* OG1RF PG. Cleavage of each substrate by SagA generates GlcNAc-MurNAc-dipeptide (GMDP, panel E) and a product fragment containing any remaining disaccharides, peptide stems, and bridges (panels B-D). All fragments illustrated with reduced anomeric carbon of MurNAc. For data shown in panel (A), extracted ion chromatograms of substrates and products were obtained, and peak areas were recorded and normalized to that of the unchanging GM-tri-(AA). Data were analyzed with one-way ANOVA and Tukey's multiple comparisons post-hoc test. Centerline, average. Error bar, SD ****$P < 0.0001$, ***$P < 0.0002$, **$P < 0.0021$. Comparisons with no asterisk had $P > 0.05$ and were not considered significant.

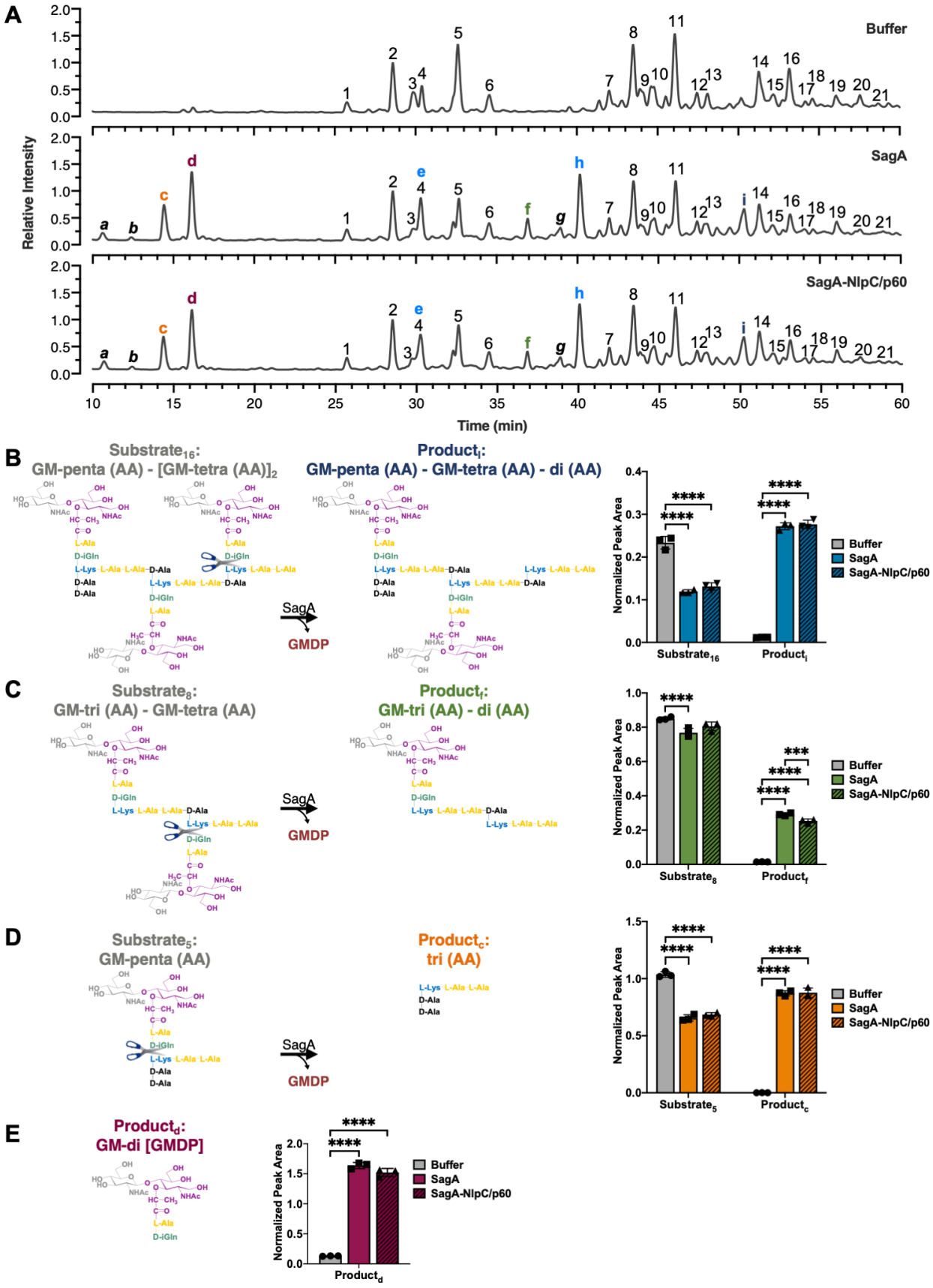

A
Buffer
SagA
SagA-NlpC/p60
Relative Intensity
Time (min)
B
Substrate16:
GM-penta (AA) - [GM-tetra (AA)]2
Producti:
GM-penta (AA) - GM-tetra (AA) - di (AA)
SagA
GMDP
Normalized Peak Area
Substrate16
Producti
Buffer
SagA
SagA-NlpC/p60
C
Substrate8:
GM-tri (AA) - GM-tetra (AA)
Productf:
GM-tri (AA) - di (AA)
SagA
GMDP
Substrate8
Productf
D
Substrate5:
GM-penta (AA)
Productc:
tri (AA)
SagA
GMDP
Substrate5
Productc
E
Productd:
GM-di [GMDP]
Productd

Table 3.1. Molecular mass and composition of soluble muropeptides from *E. faecalis* OG1RF V2.

Peak[a]	RT (min)	calculated $[M+H]^+$	observed $[M+H]^+$	Proposed structure[b]
1	25.7	897.4411	897.4412	GM-tetra
2	28.6	968.4782	968.4775	GM-tri (AA)
3	29.9		1167.6100	ND[c]
4	30.4	969.4623	969.4619	GM-tri (AA), Glu
5	32.6	1110.5525	1110.5517	GM-penta (AA)
6	34.6	1111.5365	1111.5364	GM-penta (AA), Glu
7	42.0		2069.9494	ND[c]
8	43.5	1989.9830	1989.9828	GM-tri (AA) - GM-tetra (AA)
9	43.9		2190.1174	ND[c]
10	44.6	1990.9671	1990.9672	GM-tri (AA) - GM-tetra (AA), Gln/Glu[d]
11	46.1	2132.0573	2132.0566	GM-penta (AA) - GM-tetra (AA)
12	47.4	2133.0413	2133.0422	GM-penta (AA) - GM-tetra (AA), Gln/Glu[d]
13	48.1	2133.0413	2133.0410	GM-penta (AA) - GM-tetra (AA), Gln/Glu[d]
14	51.3	3011.4878	3011.4885	GM-tri (AA) - $[GM\text{-}tetra\ (AA)]_2$
15	52.2	3012.4719	3012.4737	GM-tri (AA) - $[GM\text{-}tetra\ (AA)]_2$, Gln/Glu[d]
16	53.2	3153.5621	3153.5625	GM-penta (AA) - $[GM\text{-}tetra\ (AA)]_2$
17	54.1	3154.5461	3154.5492	GM-penta (AA) - $[GM\text{-}tetra\ (AA)]_2$, Gln/Glu[d]
18	54.6	3155.5301	3155.5420	GM-penta (AA) - $[GM\text{-}tetra\ (AA)]_2$, Gln/Glu x2[d]
19	56.1	4032.9926	4032.9852	GM-tri (AA) - $[GM\text{-}tetra\ (AA)]_3$
20	57.5	4175.0669	4175.0598	GM-penta (AA) - $[GM\text{-}tetra\ (AA)]_3$
21	58.4	4176.0509	4176.0468	GM-penta (AA) - $[GM\text{-}tetra\ (AA)]_3$, Gln/Glu[d]

a. Peak labels refer to Figure 3.6 and Appendix 3.2.
b. GM, disaccharide (GlcNAc-MurNAc); GM-tri, disaccharide tripeptide (L-Ala-D-isoGln-L-Lys); GM-tetra, disaccharide tetrapeptide (L-Ala-D-isoGln-L-Lys-D-Ala); GM-penta, disaccharide pentapeptide (L-Ala-D-isoGln-L-Lys-D-Ala-D-Ala); AA, bridge (L-Ala-L-Ala).
c. ND: Precise structure unknown.
d. The assignment of the amide and the hydroxyl functions to either peptide stem is arbitrary.

Table 3.2. Molecular mass and composition of enzymatic products from incubation of soluble *E. faecalis* OG1RF PG with purified SagA constructs V2.

Peak[a]	RT (min)	calculated $[M+H]^+$	observed $[M+H]^+$	Proposed structure[b]
a	10.6		778.2752	ND[c]
b	12.4		495.2295	ND[c]
c	14.4	431.2613	431.2610	tri (AA)
d	16.1	698.3091	698.3086	GM-di [GMDP]
e	30.3	772.4676	772.4673	tri (AA) - di (AA)
f	36.9	1310.6919	1310.6922	GM-tri (AA) - di (AA)
g	38.9		1509.8236	ND[c]
h	40.1	1452.7661	1452.7658	GM-penta (AA) - di (AA)
i	50.3	2473.2636	2473.2620	GM-penta (AA) - GM-tetra (AA) - di (AA)

a. Peak labels and formatting refer to Figure 3.6 and Appendix 3.2.
b. GM, disaccharide (GlcNAc-MurNAc); GM-di, disaccharide dipeptide (L-Ala-D-isoGln); GM-tri, disaccharide tripeptide (L-Ala-D-isoGln-L-Lys); GM-tetra, disaccharide tetrapeptide (L-Ala-D-isoGln-L-Lys-D-Ala); GM-penta, disaccharide pentapeptide (L-Ala-D-isoGln-L-Lys-D-Ala-D-Ala); AA, bridge (L-Ala-L-Ala).
c. ND: Precise structure unknown

To further understand the mechanisms involved in SagA activity, we performed alanine scanning mutagenesis of SagA-NlpC/p60 using our optimized expression and purification protocol from *E. coli* (*41*). Our previous studies show that W433 and W462 are required for SagA-NlpC/p60 activity and may contribute to substrate binding based on the SagA-NlpC/p60 structure and modeling studies (*40*), but other amino acid residues involved in activity and potential substrate binding have not been evaluated. We therefore performed docking studies of the monomeric pentapeptide PG fragment, GlcNAc-MurNAc-L-Ala-D-isoGln-L-Lys-D-Ala-D-Ala (GlcNAc-M5P) with the SagA-NlpC/p60 x-ray structure using the program Glide (Schrödinger) (*108*, *109*). GlcNAc-M5P is a common PG fragment in *Efm* Com15 and the SagA- *Efs* OG1RF with the only difference being the bridge link at L-Lys: D-Asn/D-Asp in the former (*102*) and L-Ala-L-Ala in the latter (*110*). Docking of larger cross-linked PG fragments and products was unsuccessful. The docking of GlcNAc-M5P to SagA-NlpC/p60 was constrained so that the L-Ala-γ-D-isoGln dipeptide in the pentapeptide stem was positioned in proximity to the catalytic Cys443 residue like the L-Ala-γ-D-isoGlu dipeptide in the *Bacillus cereus* YkfC NlpC/p60 structure (*61*). The ligand-docked model revealed potential key residues for D,L-endopeptidase activity inside and surrounding the putative substrate binding groove (**Fig. 3.7A and 3.8A**). Subsequent alanine scanning of these residues using an in-gel profiling assay to observe the production of GlcNAc-MurNAc-dipeptide (*40*) demonstrated the impact of these mutations on SagA-NlpC/p60 D,L-endopeptidase activity *in vitro* (**Fig. 3.7B,C and 3.8E**).

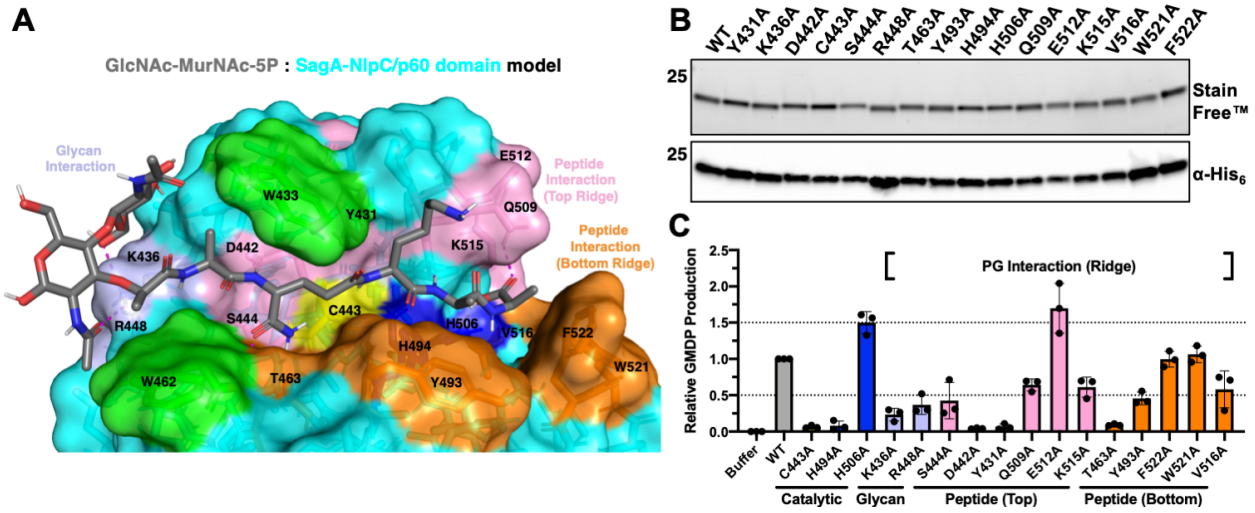
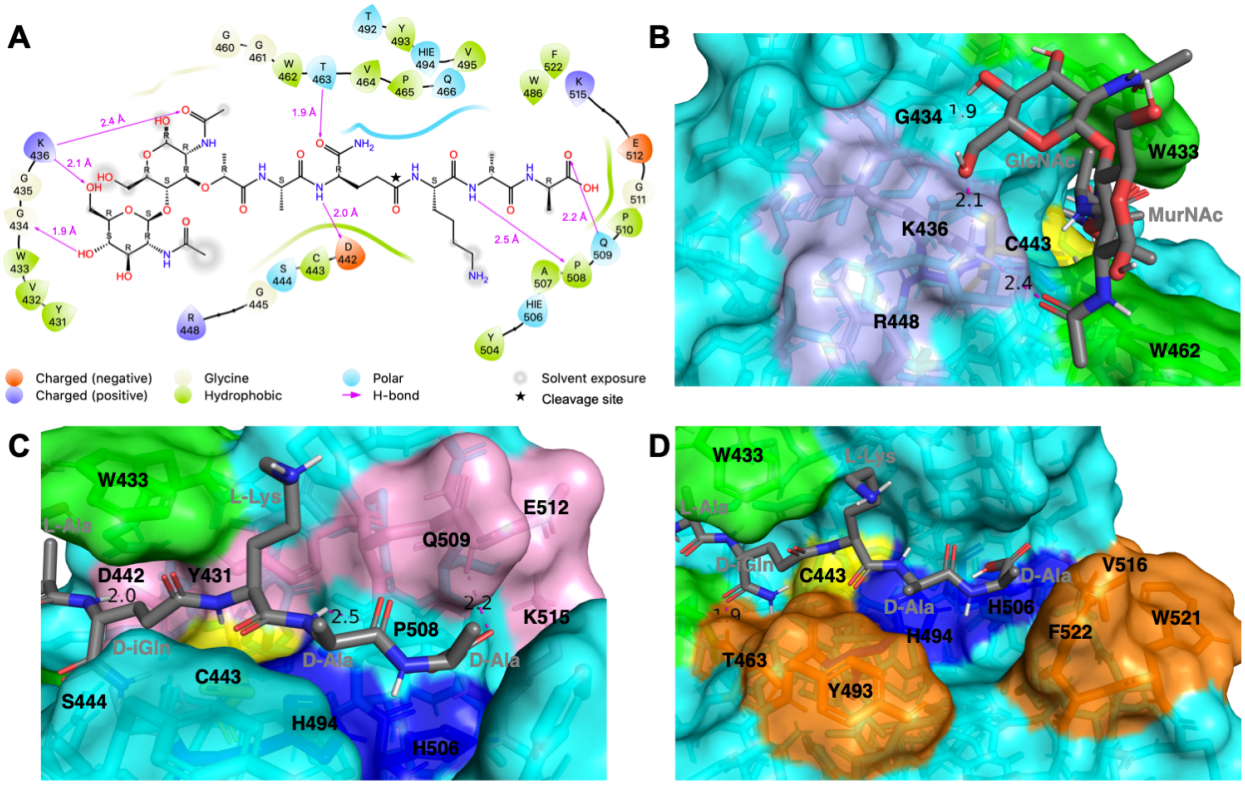


Figure 3.7. Alanine scanning of the putative SagA-NlpC/p60 substrate binding groove reveals key residues for endopeptidase activity *in vitro*. (**A**) Model of GlcNAc-MurNAc-pentapeptide (5P = L-Ala-D-isoGln-L-Lys-D-Ala-D-Ala) docked into the SagA-NlpC/p60 structure with optimized hydrogen bonding network at pH 7.0 using Glide (Schrödinger, LLC, New York, NY). Residues in the putative SagA-NlpC/p60 binding groove with potential interactions with PG are highlighted: positively charged residues at the entrance of the groove (lavender, glycan interaction); mostly polar and charged residues along the top ridge of the binding groove (pink, peptide interaction); mostly aromatic residues along the bottom ridge of the binding groove (orange, peptide interaction). Putative catalytic triad of Cys443 (yellow), His494 (blue), His506 (blue), and the required Trp433 (green), and Trp462 (green) are highlighted. Hydrogen bonding marked by dashed magenta lines. (**B**) Stain-Free (Bio-Rad) SDS-PAGE and α-His_6 immunoblotting of SagA-NlpC/p60 alanine mutants purified from the cell-free lysate of BL21(DE3)-RIL *Escherichia coli*. (**C**) Quantification of relative GlcNAc-MurNAc-dipeptide (GMDP) production by SagA-NlpC/p60 alanine mutants. Mutanolysin-digested *E. faecalis* OG1RF PG was incubated with buffer or SagA-NlpC/p60 alanine mutants at 37˚C for 16 hr, followed by incubation of the reaction mixture with ANTS fluorophore at 37˚C for 16 hr and in-gel profiling of the PG fragments. Intensities of the GMDP product bands were recorded and normalized to the WT band intensity (n = 3). Columns are highlighted according to docked model in panel (A). Top horizontal line, ≥ 50% WT activity. Bottom horizontal line, ≤ 50% WT activity. Column centerline, average. Error bar, SD.

Figure 3.8. Zoom-in views of GlcNAc-MurNAc-5P: SagA-NlpC/p60 domain model (A-D) and (E) in-gel profiling of PG hydrolase activity of SagA-NlpC/p60 mutants. (**A**) 2D ligand interaction diagram of the GlcNAc-MurNAc-5P: SagA-NlpC/p60 model. GlcNAc-MurNAc-5P (GlcNAc-M5P) ligand is displayed as a 2D structure. Residues within 6 Å of the ligand are represented as teardrop shapes labeled with their 1 letter residue code and residue number. The teardrop shapes are colored according to the properties of the residues, and the points on the shapes are oriented in the direction of the sidechains. Black lines represent chains of connected residues with residues at distances beyond 3 Å of ligand represented as black dots. Hydrogen bonds between the residues and the ligand are drawn as purple arrows going from donor to acceptor with the distances labeled above. Residues present in the diagram without hydrogen bond interactions with the ligand are involved in nonspecific hydrophobic interactions with the ligand. The binding pocket is indicated by a line drawn around the ligand that is colored by the color (i.e. property) of the nearest residue. Solvent exposure is indicated on the ligand atoms by a gray circle, and by the break in the line drawn around the binding pocket. D,L-endopeptidase cleavage site marked by black star. G434 and P508 omitted from the alanine screen in favor of the nearby polar residues protruding from the putative SagA-NlpC/p60 binding groove. (**B-D**) Zoomed in view of potential interactions between putative SagA-NlpC/p60 binding groove and docked GlcNAc-M5P ligand. Hydrogen bonds shown as dashed magenta lines and distances (Å) labeled in black. Potential interactions between SagA-NlpC/p60 and the GlcNAc-MurNAc moiety of docked GlcNAc-M5P ligand in panel (B). Potential interactions between residues in the top (C) and bottom (D) ridges of the putative SagA-NlpC/p60 substrate binding groove with the peptide stem of docked GlcNAc-M5P ligand. (**E**) In-gel profiling of PG fragments generated by SagA-NlpC/p60 mutants *in vitro* corresponding to Figure 3.7C.


A
B
C
D
E
Charged (negative)
Charged (positive)
Glycine
Hydrophobic
Polar
H-bond
Solvent exposure
Cleavage site
G434
W433
K436
C443
R448
W462
GlcNAc
MurNAc
L-Lys
L-Ala
D-iGln
D-Ala
D442
Y431
E512
Q509
K515
P508
S444
H494
H506
T463
Y493
F522
W521
V516
PG Interaction (Ridge)
Cat. Triad
Glycan
Peptide (Top)
Peptide (Bottom)
Buffer
WT
C443A
H494A
H506A
R448A
K436A
S444A
D442A
Y431A
Q509A
E512A
K515A
T463A
Y493A
F522A
W521A
V516A
Native
GMDP

We then evaluated the activity of individual alanine mutants of the residues in the putative catalytic triad of SagA-NlpC/p60, which showed that the C443A and H494A mutants were inactive. Surprisingly, not only did the H506A mutant retain its activity but it was also slightly more active than wild-type SagA-NlpC/p60 (**Fig. 3.7C**). The disaccharide moiety of the GlcNAc-M5P docked to a region of the substrate binding groove that contained electropositive residues including Lys436 and Arg448. In the docked structure, these residues formed hydrogen bonds and polar interactions with the carbohydrates, respectively, suggesting a stabilizing role (**Fig. 3.8B**). Accordingly, mutation of either of these residues to Ala diminished SagA hydrolytic activity. We next examined the residues that lined the solvent exposed binding cleft of the pentapeptide stem. With the docked model oriented as shown in **Fig. 3.8C,D**, we observed two prominent ridges of the binding cleft. The first half of the top ridge consists of Ser444, Asp442 and Tyr431, which are highly conserved in NlpC/p60 hydrolases and clustered around the catalytic Cys443 in the SagA-NlpC/p60 structure (**Fig. 3.8C**) (*1*, *33*). The S444A mutant showed diminished activity, whereas the D442A and Y431A mutants were inactive (**Fig. 3.7C**), likely due to the disruption of the interaction of each residue with γ-D-isoGln shown in the docked model (**Fig. 3.8C**). While Ser444 did not directly bind with the PG stem, Asp442 hydrogen bonded with the amide NH group of γ-D-isoGln (*36*, *38*, *61*). Moreover, Tyr431 is oriented towards the carbonyl of the γ-D-isoGln-L-Lys amide bond, supporting its proposed function in creating an "oxyanion hole" to stabilize the tetrahedral intermediate formed during catalysis (*32*, *33*, *37*, *58*). The second half of the top ridge consists of Gln509 and Glu512 protruding over the groove, and Lys515 at the seam of the groove (**Fig. 3.8C**). The observed lower activity of the Q509A mutant may be due to the lack of the hydrogen bond shown between Gln509 and the carboxyl carbonyl of the terminal D-Ala which may help anchor peptide stems in the groove. Surprisingly, the E512A mutant demonstrated enhanced D,L-endopeptidase activity whereas the activity of the K515A mutant was on par with that of the Q509A mutant (**Fig. 3.7C**). This improved activity for the E512A mutant was interesting, as no direct bond was observed between Glu512 and the terminal D-Ala-D-Ala in the docked model. Likewise, Lys515 did not directly interact with the bound ligand, but it does contribute to an exposed pocket at the seam of the groove, which may be important for proper binding of PG substrates and may explain the weaker activity of the K515A mutant.

In the bottom ridge (**Fig. 3.8D**), Thr463 is the only polar residue near Cys443, and the abrogated activity of the T463A mutant (**Fig. 3.7C**) may be due to the disruption of the hydrogen bond formed between the backbone carbonyl of Thr463 and the amino group of γ-D-isoGln. The first aromatic residue in the bottom ridge, Tyr493, is positioned near the branching point of L-Lys and D-Ala-D-Ala in the middle of the groove and the Y493A mutant showed lower activity. *In silico* alanine scanning of Tyr493 using Schrödinger generated the SagA-NlpC/p60 Y493A mutant structure (**Fig. 3.9B**). The binding groove is more solvent-exposed (**Fig. 3.9A,B**), which suggests that this aromatic residue may hold the terminal amino acids within the groove via van der Waals interactions. Mutants of residues towards the end of the bottom ridge, F522A and W521A, had similar activity to wild-type, suggesting that these residues are not crucial for activity even though they are positioned near the terminal D-Ala (**Fig. 3.8D**). In contrast, the V516A mutant exhibited weaker activity, perhaps because it contributes to the exposed pocket at the seam of the groove like Lys515.

To explain the enhanced activity observed with the H506A and E512A mutants, further *in silico* alanine scanning was performed with particular attention paid to the periphery of the SagA-NlpC/p60. Interestingly, His506 juts into this pocket and Glu512 is directly above it. *In silico* alanine scanning of these residues generated SagA-NlpC/p60 structures with alterations of the

pocket: the H506A mutant structure contributed a pronounced cavity to the pocket, thus deepening it (**Fig. 3.9C**), and the E512A mutant structure flattened the space above the pocket (**Fig. 3.9D)**. Therefore, it is conceivable that these structural alterations allow for better binding and hydrolysis of PG as observed in our *in vitro* activity assays (**Fig. 3.7C**). Collectively, these results indicate that PG substrates are oriented towards the top ridge of the putative SagA-NlpC/p60 substrate binding groove for proper D,L-endopeptidase activity and production of GlcNAc-MurNAc-dipeptide.

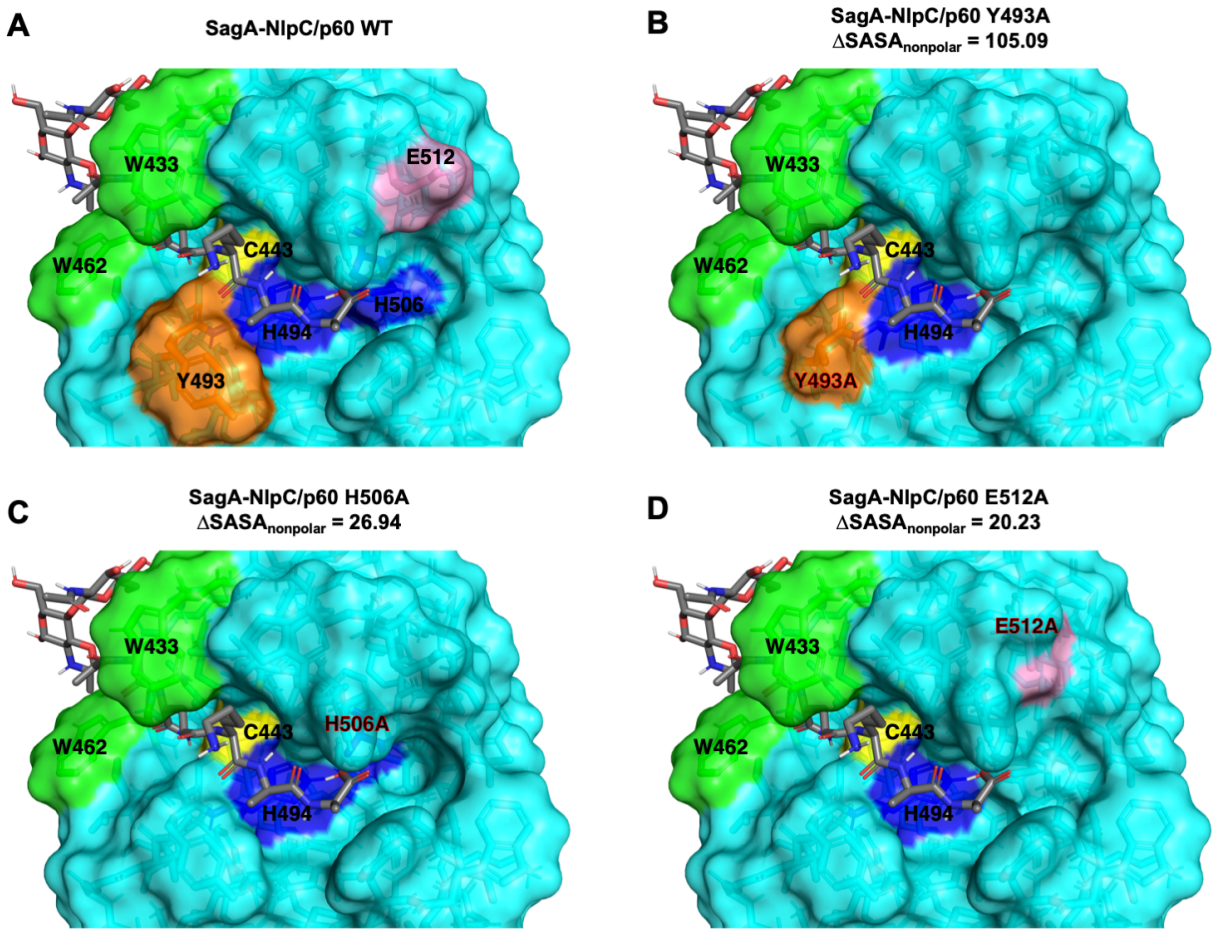


Figure 3.9. *In silico* alanine scanning of the SagA-NlpC/p60 structure using Glide generates mutant structures with altered substrate binding groove peripheries. (**A**) Wild-type SagA-NlpC/p60 structure. (**B-D**) Mutant SagA-NlpC/p60 structures generated using the residue scanning panel in Glide. SagA-NlpC/p60 Y493A (B) H506A (C) and E512A (D) mutant structures shown. The change in surface area of nonpolar atoms due to the mutation ($\Delta SASA_{nonpolar}$) included for each mutant structure. GlcNAc-M5P ligand from docking study with wild-type SagA-NlpC/p60 structure superimposed to mutant structures for reference.

Genomic analysis and primary sequence alignment revealed that the Cys-His-His triad is conserved across multiple SagA orthologs from other enterococci (**Fig. 3.10A**). To address if SagA-like proteins in general used catalytic dyads or if our finding was unique to *Efm* Com15 SagA, we prepared recombinant Cys-His-Ala mutants of full-length SagA orthologs from *Efm* Com15, *Eds* 2RC2, and *Emi* 2312 (**Fig. 3.10B**) and evaluated their biochemical activity *in vitro*. Compared to the buffer control and the inactive *Efm* Com15 SagA C443A mutant, all Cys-His-Ala mutants demonstrated D,L-endopeptidase activity (**Fig. 3.10C, Appendix 3.2, Table 3.1 and 3.2**), suggesting that the conserved second histidine is not required for *in vitro* activity. Inspection of the docked SagA-NlpC/p60 structure did not reveal any obvious residues that may pair with Cys443-His494 to comprise a catalytic triad, thus indicating that SagA-NlpC/p60 may function via a catalytic Cys-His dyad. Although most structurally similar NlpC/p60 hydrolases contain residues for a Cys-His-His catalytic triad in a similar orientation to SagA-NlpC/p60 (*40*), the functional role of these residues has not been completely established. Therefore, these results suggest that the putative catalytic Cys-His-His triad should be reevaluated for other NlpC/p60 hydrolases. A notable exception is RipA-NlpC/p60, which has a required catalytic Cys-His-Glu triad (*36*). The glutamic acid residue is proposed to form a salt bridge with a neighboring arginine residue to lock in the orientation of the catalytic cysteine, which Squeglia et al note an equivalent histidine residue is unable to form (*36*). The lack of this interaction in NlpC/p60 domains with a conserved Cys-His-His triad, coupled with the limited flexibility of the catalytic site in these domains (*36*), raises the possibility that the second histidine may not have a role in catalysis.

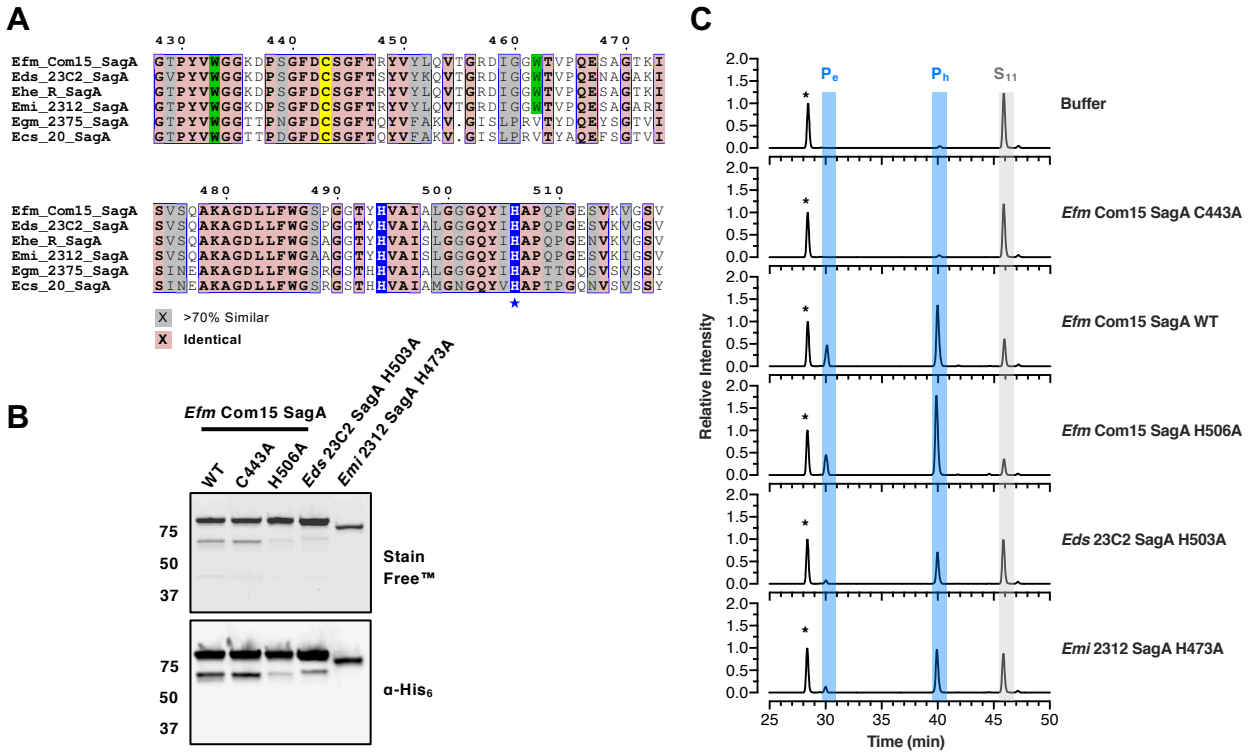


Figure 3.10. Conserved His506 not required for in vitro activity of SagA orthologs from other enterococci. (**A**) Multiple sequence alignment of *Efm* Com15 SagA-NlpC/p60 with the NlpC/p60 domains of SagA orthologs. SagA ortholog sequences were located downstream of *mreD* locus using BioCyc. NlpC/p60 domains were determined using InterPro. Multiple sequence alignment was performed using ClustalOmega and analyzed with ESPript. Conserved H506 marked with blue star. (**B**) Stain-Free (Bio-Rad) SDS-PAGE and α-His_6 immunoblotting of SagA orthologs purified from the cell-free lysate of BL21(DE3)-RIL *Escherichia coli*. (**C**) LC-MS analysis of soluble mutanolysin-digested *E. faecalis* OG1RF PG incubated with buffer or full-length SagA orthologs at 37˚C for 16 hr. Data are shown as extracted ion chromatograms with colored columns representing $Substrate_{11}$, $Product_h$, and $Product_e$, as in Figure 3.5B,C. See Appendix 3.2 for the corresponding total ion chromatograms. *Efm* Com15, *E. faecium* Com15; *Eds* 23C2, *E. durans* 23C2; *Ehe* R, *E. hirae* R; *Emi* 2312, *E. mundtii* NCDO 2312; *Egm* 2375, *E. gallinarum* NCDO 2375; *Ecs* 20, *E. casseliflavus* 20.

To better understand structural features that govern NlpC/p60 hydrolase function, we compared the topology of the putative SagA-NlpC/p60 substrate binding groove with that of other biochemically and structurally characterized NlpC/p60-type hydrolases. These representative NlpC/p60 hydrolases have varied domain architectures (**Fig. 3.11A**) and their NlpCp60 domains display a broad range of enzymatic and hydrolytic specificities *in vitro*. Sequence alignment of the NlpCp60 domains show conservation of the SagA-NlpC/p60 core residues (i.e. Y431, D442, C443, S444, H494 and H506), but there is poor similarity of several of the residues that we have demonstrated to modulate the *in vitro* endopeptidase activity of SagA-NlpC/p60 (**Fig. 3.11B**). However, structural alignment of SagA-NlpC/p60 with these NlpC/p60 domains highlights key differences in their putative binding grooves that may dictate their respective functions (**Fig. 3.11C and 3.12**). For instance, unlike SagA-NlpC/p60, the RipA NlpC/p60 domain has a prominent loop

occluding its entire binding groove (**Fig. 3.12A**). This loop is a putative inhibitory prodomain that is proposed to be proteolytically cleaved *in vivo* for activation of RipA at the *M. tuberculosis* septum (*7*, *21*, *22*, *26*, *54*). The *Staphylococcus aureus* CwlT NlpC /p60 domain, on the other hand, shows a loop protruding from its bottom ridge partially blocking the groove near the site of Tyr493 in SagA-NlpC/p60, but it also has a flatter and more open groove periphery (**Fig. 3.12B**) (*37*). The *E. coli* Spr NlpC/p60 domain contains a similar loop, but it actually points into the groove, as do several bulky residues at the periphery (**Fig. 3.12C**) (*32*). The binding groove of the ligand bound NlpC/p60 domain of *Bacillus cereus* YkfC is slightly more open, except for two features at the beginning and end of the groove (**Fig. 3.12D**) (*61*). First, the SH3b1 domain of YkfC blocks the entrance of its NlpC/p60 binding groove, which is proposed to contribute to the substrate specificity of YkfC towards stem peptides with an N-terminal L-Ala (i.e. no disaccharide moiety) (*58*). Second, Arg309 and Tyr321 of YkfC are positioned near Glu512 and Phe522 in the SagA-NlpC/p60 structure and they occlude the pocket in the YkfC binding groove periphery. Interestingly, the NlpC/p60 domain whose binding groove topology is most similar to SagA-NlpC/p60 is of *M. tuberculosis* RipD (**Fig. 3.12E**), which is a non-catalytic NlpC/p60 domain that retains its ability to bind PG *in vitro* (*39*). Overall, the diversity in the binding groove topologies of these NlpC/p60 hydrolases illustrates the challenge of elucidating the exact structural motifs and substrate-protein interactions that govern their enzymatic and hydrolytic specificities. Nevertheless, our structure-function studies of SagA-NlpC/p60 have identified residues and hotspots in the binding groove that may be involved in NlpC/p60 hydrolase activity, but may not have been identified via sequence homology alone, such as the site of the Tyr493 lid and the less studied binding groove periphery.

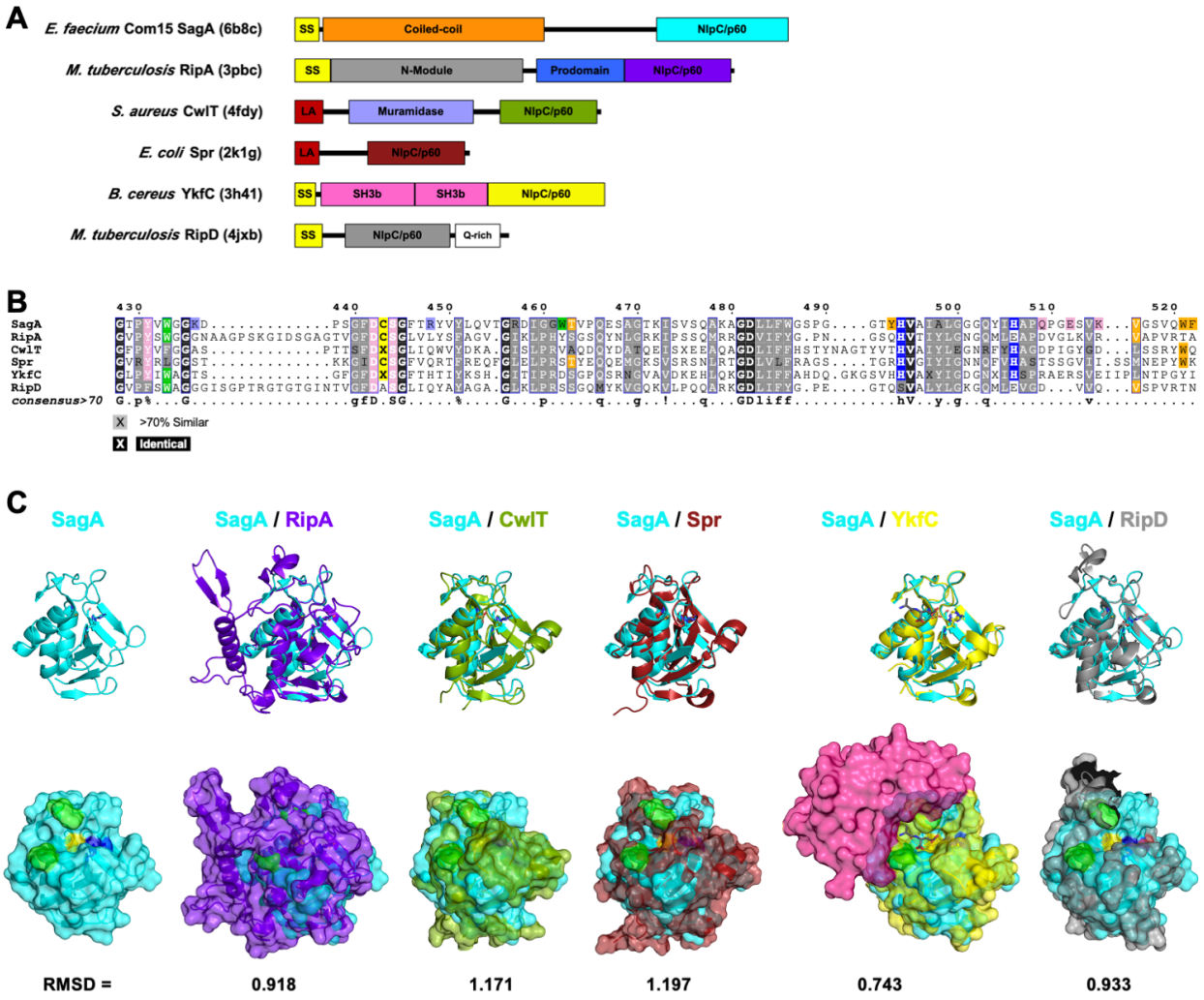


Figure 3.11. NlpC/p60 orthologs exhibit diverse substrate binding groove topologies. (**A**) Domain architectures of biochemically and structurally characterized NlpC/p60 orthologs. (**B**) Multiple sequence alignment of SagA-NlpC/p60 with the NlpC/p60 domains of characterized NlpC/p60 orthologs. NlpC/p60 domains were determined using InterPro. Multiple sequence alignment was performed using ClustalOmega and analyzed with ESPript. Color scheme follows that of Figure 3.7, with grey denoting residues with >70% consensus, and black denoting identical residues. The catalytic cysteine in the NlpC/p60 domain of the *S. aureus* CwlT and *B. cereus* YkfC structures is oxidized and marked with an X in the sequence alignment. (**C**) Structural alignment of SagA-NlpC/p60 with the NlpC/p60 domains of characterized NlpC/p60 orthologs. Ribbon (top) and surface (bottom) representations shown. RMSD of the aligned atoms to SagA-NlpC/p60 displayed below the structures. The *in vitro* biochemical activity and hydrolytic specificity of each NlpC/p60 ortholog is as follows: *Enterococcus faecium* Com15 SagA (6b8c), D,L-endopeptidase activity, L-Lys-type PG; *Mycobacterium tuberculosis* RipA (3pbc), D,L-endopeptidase activity, mDAP and L-Lys-type PG; *Staphylococcus aureus* CwlT (4fdy), D,L-endopeptidase activity, mDAP-type PG; *Escherichia coli* Spr (2k1g), D,D-endopeptidase and slight L,D-carboxypeptidase activity, mDAP-type PG; *Bacillus cereus* YkfC (3h41), amidase activity, L-Ala peptides; *Mycobacterium tuberculosis* RipD (4jxb), non-catalytic.

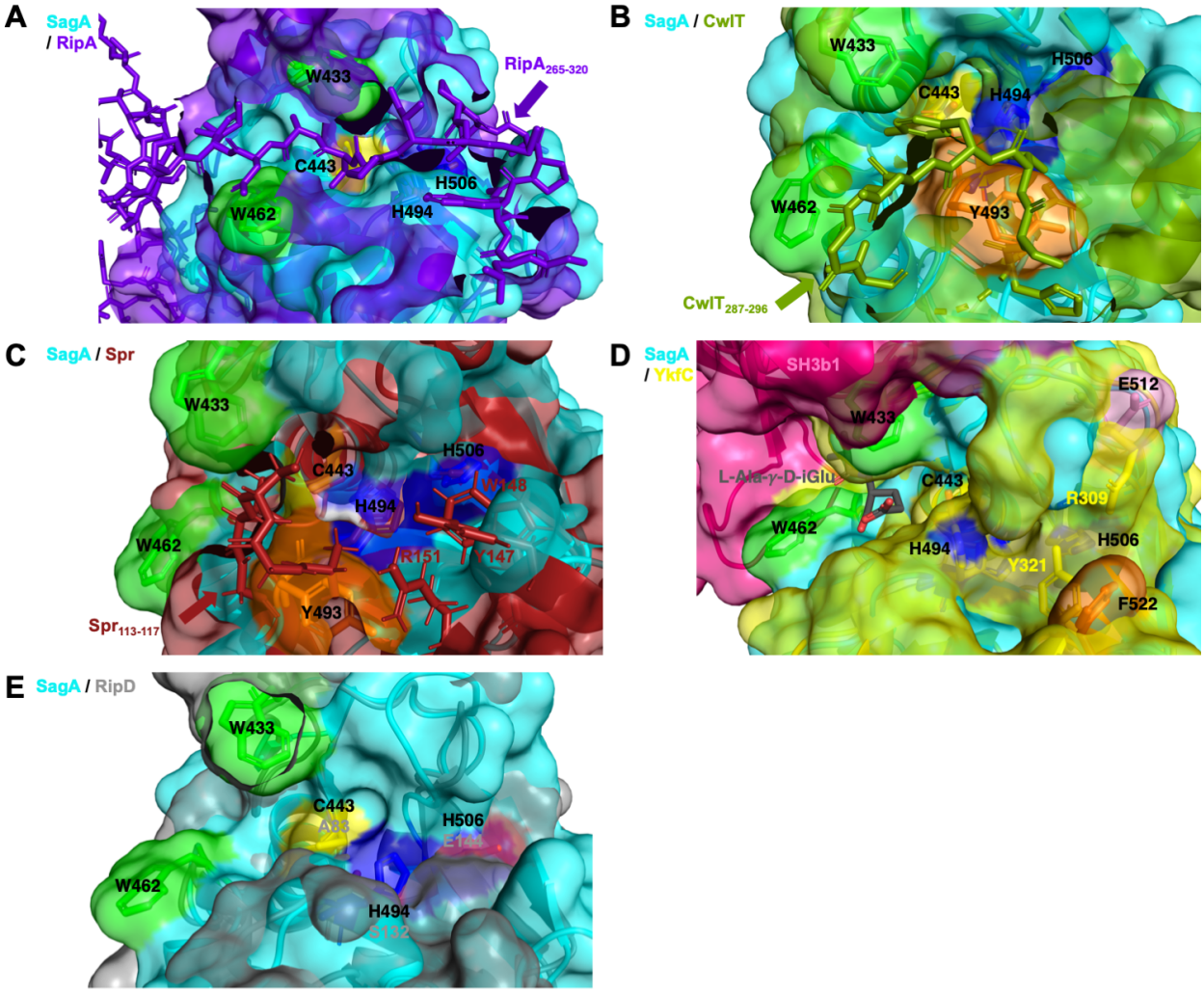


Figure 3.12. Zoomed in view of structural alignment of SagA-NlpC/p60 with the NlpC/p60 domains of characterized NlpC/p60 orthologs. For all structural alignments, SagA-NlpC/p60 colored as in Figure 3.7 with residues labeled in black. Each NlpC/p60 ortholog colored as in Figure 3.11C. (**A**) RipA-NlpC/p60$_{265\text{-}320}$ loop shown as purple sticks and marked by arrows. (**B**) CwlT-NlpC/p60$_{287\text{-}296}$ shown as dark green sticks and marked by arrow. (**C**) Spr-NlpC/p60 residues shown as burgundy sticks with burgundy labels. Spr$_{113\text{-}117}$ loop marked by burgundy arrow. (**D**) YkfC-SH3b1 domain shown in dark pink and bound L-Ala-γ-D-iGlu in dark grey. YkfC-NlpC/p60 residues shown as yellow sticks with yellow labels. (**E**) RipD residues labeled in grey.

In summary, our data indicate that SagA is colocalized with zones of active PG remodeling in enterococci, and that the NlpC/p60 hydrolase activity of SagA may be governed by potential key interactions between septal PG and the putative SagA-NlpC/p60 substrate binding groove. We previously characterized the hydrolytic specificity of SagA as a D,L-endopeptidase with a preference towards cross-linked L-Lys-type PG with penta- and tetrapeptide stems,(*40*) which comprises nascent and template PG chains during PG biosynthesis in *Efm* (*100–102*). Given that *sagA* is essential for *Efm* growth (*99*), our discovery of the NlpC/p60 hydrolase activity of SagA (*40*) and the subcellular localization of SagA at the septa of dividing enterococci cells implicate

SagA in PG turnover during cell growth and division. To identify potential key contacts between PG substrates and SagA, we docked a representative GlcNAc-MurNAc-pentapeptide into the putative SagA-NlpC/p60 binding groove and analyzed the potential substrate interacting residues of SagA-NlpC/p60 via alanine scanning *in vitro* and *in silico*. Our structural and biochemical studies suggest that PG substrates are oriented towards the top ridge of the putative SagA-NlpC/p60 substrate binding groove, and that substrate interactions at the periphery of the groove may be important for PG turnover. Moreover, our biochemical analysis suggests that SagA may function via a catalytic Cys-His dyad, which was consistent in full-length SagA orthologs from other enterococci species. While structural alignment of SagA-NlpC/p60 with other structurally and biochemically characterized NlpC/p60 orthologs exhibits the diversity in their putative substrate binding groove topologies, our studies suggest that substrate interactions at the center and periphery of the grooves may govern PG recognition, binding and hydrolysis. Ultimately, this work highlights how the activity of SagA may be involved in the viability of *Efm* and provides a greater understanding of how SagA generates the defined muropeptides that we have shown to have therapeutic potential (*40*, *68*, *69*).

Acknowledgements

We thank Paul Dominic B. Olinares for intact protein analysis by Matrix-assisted Laser Desorption Time-of-Flight Mass Spectrometry (Brian T. Chait laboratory, The Rockefeller University). We also thank Deena Oren for assistance with *in silico* residue scanning of SagA-NlpC/p60 (The Rockefeller University Structural Biology Resource Center).

Chapter 4. *Enterococcus* peptidoglycan remodeling promotes immune checkpoint inhibitor therapy.

Introduction

In studying specific host-microbe interactions, our laboratory has shown how commensal *Efm* Com15 and other SagA+ bacteria trigger host immune pathways to limit enteric infection (*40*, *68*, *69*). Pathogenic *Enterococcus* species exist, such as antibiotic-resistant *Efm* and *Efs* (*111*, *112*), but commensal enterococci in particular have been associated with several functions in host immunity in addition to what we have reported (*40*, *68*, *69*). For instance, commensal enterococci strains are currently applied as probiotics in animals and humans to treat and prevent gastrointestinal diseases (*113*). Furthermore, commensal enterococci have been found to regulate autoimmunity (*114*) and graft-versus-host disease in mice and humans (*115*). It is apparent that among the gut microbiota, commensal *Enterococcus* species may be implicated in modulating human health, and new roles for these bacteria have been discovered in the context of cancer and the efficacy of antitumor therapeutics.

To elaborate, the CTLA-4 and PD-1/PD-L pathways serve as regulatory checkpoints for the immune system and can limit antitumor T cell response (*116*). In cancer immunotherapy using checkpoint blockade, antibodies such as anti-CTLA-4 and anti-PD-1 or anti-PD-L1 block these respective pathways, which allows for continued activation and proliferation of antitumor T cells (*116*). Recent studies suggest that the human commensal microbiota may improve anticancer treatment in patients, as microbiota from responders to anti-PD-1 immunotherapy enhanced tumor clearance in mouse models (*117*, *118*). Notably, different commensal *Enterococcus* species were recovered in anti-PD-1 immunotherapy-responsive cancer patients (*117*, *118*), but their specific activity and mechanism in anticancer immunotherapy had not been described.

In this chapter, we describe our investigation into how specific *Enterococcus* species and strains influence the efficacy of checkpoint blockade agents. To evaluate how enterococci may elicit this improved antitumor activity, we used a mouse tumor model and discovered that specific enterococci – including several strains of *Efm* but not *Efs* – enhanced the antitumor response of anti–PD-L1. To ascribe the distinguishing feature(s) of these enterococci that endow them with this antitumor activity, we assessed the immunostimulatory *Efm* phenotypes described in this thesis, namely SagA expression and secretion, *Enterococcus* cell wall composition, and SagA PG hydrolase activity. We confirmed that the protective enterococci express and abundantly secrete SagA orthologs with highly conserved NlpC/p60 domains, and that their cell wall is composed of more non-crosslinked PG than the non-protective *Efs*. Using our established *in vitro* activity assays for NlpC/p60 PG hydrolase activity, we demonstrated that the SagA orthologs encoded by the protective enterococci are D,L-endopeptidases with similar activity to *Efm* Com15 SagA. The previously uncharacterized *Efs* OG1RF SalA included in this analysis, on the other hand, was revealed as a D,L-endopeptidase with different cleavage bond specificity. Heterologous expression of SagA in the probiotic bacterium *L. lactis* ultimately confirmed that secretion of active, functional SagA is sufficient to promote immune checkpoint inhibitor therapy. Together, these studies provide direct evidence of how specific bacteria play a role in mediating host immunity in the context of cancer and serve as further evidence of the therapeutic potential of SagA PG hydrolase activity. This work was uploaded to bioRxiv (*43*). Dr. Matthew Griffin first identified and confirmed the expression of the SagA orthologs in the enterococci and executed all *in vivo* experiments.

Results and Discussion

To evaluate how specific *Enterococcus* species and strains may improve checkpoint inhibitor blockade in cancer immunotherapy, we used a murine melanoma model and orally administered bacteria. Briefly, specific pathogen-free (SPF) C57BL/6 mice from The Jackson Laboratory were pretreated for two weeks with a broad-spectrum antibiotic cocktail to clear the resident gut microbiota. Afterwards, mice were given water supplemented with the bacteria of interest. Supplemented mice were then subcutaneously implanted with B16-F10 melanoma cells and treated with anti–PD-L1 (**Fig. 4.1A**). We found that supplementation of *Efm* Com15 alone did not decrease tumor growth, whereas treatment of anti-PD-L1 only had modest activity *in vivo* (**Fig. 4.1B**). However, *Efm* Com15 supplementation and anti-PD-L1 treatment at both low and high doses synergistically suppressed tumor growth *in vivo* (**Fig. 4.1B**). Therefore, we proceeded to test whether individual *Enterococcus* species and strains would maintain this synergistic phenotype at the low anti-PD-L1 dose. First, we evaluated human-isolated, type, and multi-drug-resistant strains of *Efm* and *Efs* given that they are the most prevalent in the human microbiota, and that we wanted to determine whether this antitumor effect was unique to *Efm* Com15. Only mice supplemented with the *Efm* strains had decreased tumor volumes (**Fig. 4.1C**). Quantification of bacteria in the fecal samples harvested from animals showed that these results were not due to differential colonization in the gut between the *Efm* and *Efs* strains (**Fig. 4.1D**). As stated previously, multiple commensal *Enterococcus* species including *Efm* were recovered in anti-PD-1 immunotherapy-responsive cancer patients, such as *Eds*, *Ehe*, *Emi* and *Egm* (*117*). Using the same murine melanoma model described here, we found that only certain species from this group shared the antitumor activity observed with *Efm* (**Fig. 4.1E**), and that differences in bacterial load again did not correlate this activity (**Fig. 4.1F**) (*43*).

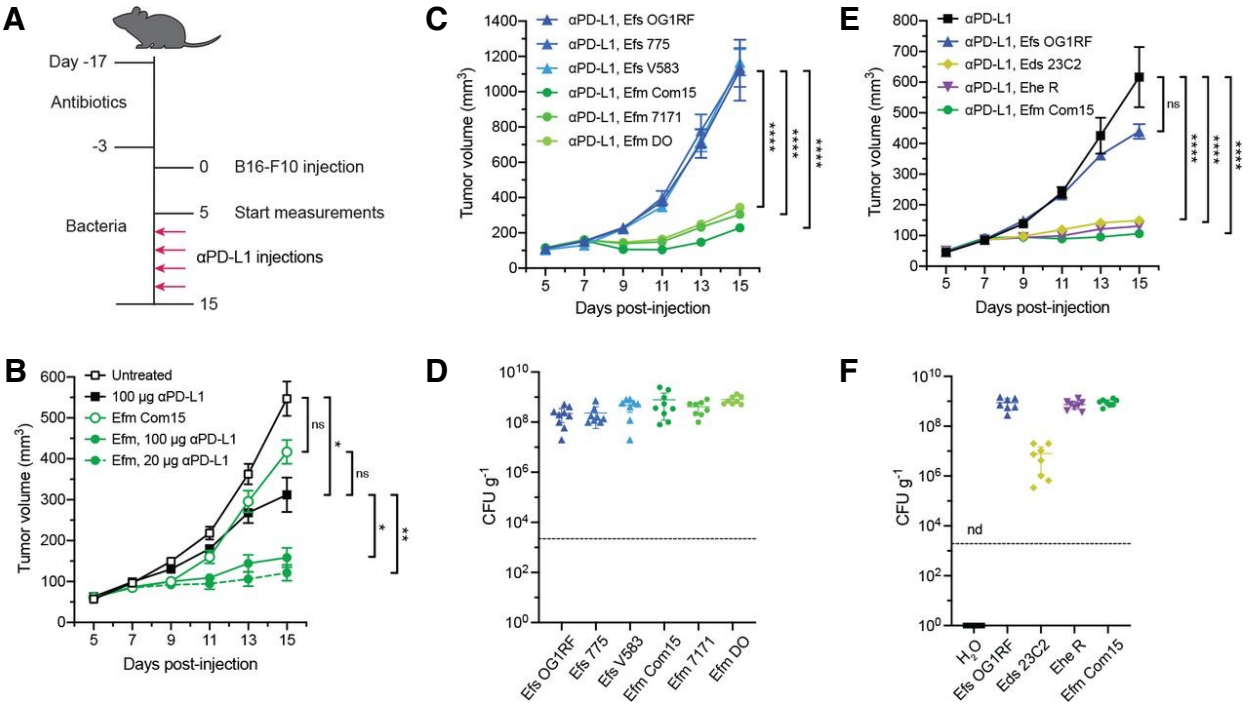
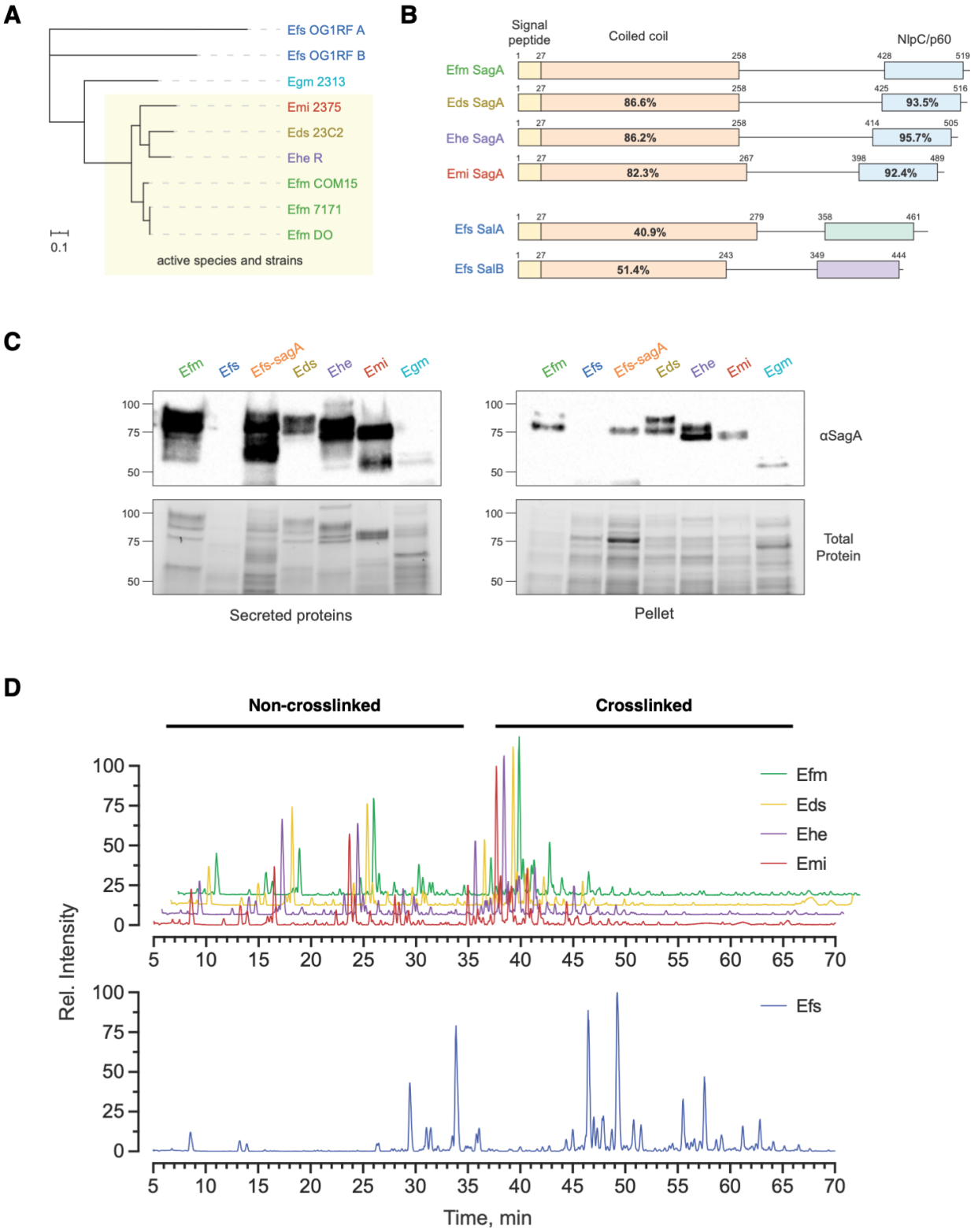


Figure 4.1. Specific enterococci improve anti–PD-L1 efficacy in B16-F10 melanoma model. (**A**) Schematic of tumor growth model in SPF mice with antibiotic treatment and oral enterococci supplementation. Days are indexed based on the day of tumor injection. Mice were provided antibiotic-containing water *ab libitum* for two weeks followed by water supplemented with the indicated enterococci for the remainder of the experiment. Animals were then subcutaneously implanted with B16-F10 melanoma cells, and tumor volume measurements started when tumors reached ~50-100 cm^3 (day 5). Mice were treated with anti–PD-L1 by intraperitoneal injection every other day starting two days after the start of measurement. For all data except for (B), 20 μg anti–PD-L1 was used for each injection. (**B**) B16-F10 tumor growth in antibiotic-treated animals that were supplemented with or without *E. faecium* (*Efm*) Com15 and treated with or without anti–PD-L1 starting on day 7 at doses indicated in the legend. n = 7-8 mice per group. (**C**) B16-F10 tumor growth in antibiotic-treated mice that were supplemented with the indicated *E. faecalis* (*Efs*) and *Efm* strains and treated with anti–PD-L1 starting on day 7. n = 7-8 mice per group. (**D**) Colony forming unit (CFU) analysis of *Efs* and *Efm* strains in fecal samples harvested from animals as treated in (C). (**E**) B16-F10 tumor growth in antibiotic-treated mice that were supplemented with the indicated enterococci and treated with anti–PD-L1 starting in day 7. n = 8-9 mice per group. (**F**) CFU analysis of enterococci in fecal samples harvested from animals as treated in (E). nd = not detected. For (B), (C), and (E), data represent as mean ± s.e.m. and were analyzed by mixed-effects model with Tukey's multiple comparisons post-hoc test. *$P < 0.05$, **$P < 0.01$, ****$P < 0.0001$, ns = not significant. For (D) and (F), each symbol represents one mouse. Dotted lines indicate the limit of detection (2,000 CFU g^{-1}). Data represent means ± 95% confidence interval. The data shown here are the work of Dr. Matthew Griffin.

Previous chapters in this thesis discuss the immunomodulatory effects of *Efm* SagA secretion and PG hydrolase activity, which led us to hypothesize that the conserved antitumor activity in the active enterococci correlated with expression of functional SagA. Genomic analysis revealed that the active enterococci encoded SagA orthologs downstream of the *mreD* locus (**Fig. 4.2A**), which we first noted in Chapter 3 upon the observation of the triseptal phenotype in SagA+ enterococci. Multiple sequence alignment of these SagA orthologs showed a conserved domain architecture and over 90% sequence identity in the catalytic NlpC/p60 domain (**Fig. 4.2B**). Conversely, the non-protective *Egm* and *Efs* encode distinct proteins. The closest related SagA-like protein in *Egm* showed only 67% sequence identity in the predicted catalytic C-terminal domain (**Fig. 4.3A**). Likewise, no direct SagA orthologs were found in *Efs*, as we reported previously (*40*, *69*), and the two most similar conserved proteins, *sagA*-like proteins A and B (SalA and SalB) (*119*), possessed distinct C-terminal domains (**Fig. 4.2B**). Using polyclonal SagA serum raised against purified, recombinant *Efm* Com15 SagA (*40*), we confirmed that only the active enterococci express SagA, as evidenced by the abundant signal of secreted SagA in culture media and of cytoplasmic SagA in bacterial cell pellets (**Fig. 4.2C**). Accordingly, comparative analysis of the cell wall composition of the active enterococci displayed a conserved phenotype; the active, SagA+ enterococci had more non-crosslinked PG fragments compared to the non-active *Efs* (**Fig. 4.2D**) or *Egm* (**Fig. 4.3**). Our proposed model of SagA-mediated protection against enteric pathogens posits that *Efm* Com15 has more small, non-crosslinked muropeptides in its cell wall that may stimulate the host immune system (*40*), which these findings regarding the SagA+ enterococci further reinforce.

Figure 4.2. Immunotherapy-active enterococci possess highly conserved orthologs to *Efm* SagA and more non-crosslinked PG fragments than *Efs*. (**A**) Unrooted phylogenetic clustering of putative SagA ortholog protein sequences identified by global PG peptidase analysis of enterococci species and strains along with the closest entries from *Egm* and *Efs* based on IQ-Tree analysis. Active strains are indicated by the yellow box. Scale bar represents sequence distance. (**B**) Comparison of primary sequence homology and domain architecture of SagA orthologs and two conserved *sagA*-like proteins from *Efs*. Numbers above each bar denote amino acid residue boundaries of the indicated domains, and percentages represent sequence identity compared to *Efm* SagA. (**C**) Western blot detection of SagA orthologs in secreted protein and cell pellet fractions harvested from overnight cultures of the indicated enterococci using antiserum raised against *Efm* Com15 SagA. Bottom panels show total protein loading. Numbers indicate estimated molecular weight (kDa). (**D**) LC-MS ion chromatograms of mutanolysin-treated PG harvested from *Enterococcus* species. Data are shown as relative intensity of base peak ion abundance. Dr. Matthew Griffin performed the genomic analysis of SagA orthologs in enterococci and the SagA expression and secretion profiling.


A
Efs OG1RF A
Efs OG1RF B
Egm 2313
Emi 2375
Eds 23C2
Ehe R
Efm COM15
Efm 7171
Efm DO
0.1
active species and strains
B
Signal peptide
Coiled coil
NlpC/p60
Efm SagA
Eds SagA
86.6%
93.5%
Ehe SagA
86.2%
95.7%
Emi SagA
82.3%
92.4%
Efs SalA
40.9%
Efs SalB
51.4%
C
Efm
Efs
Efs-sagA
Eds
Ehe
Emi
Egm
αSagA
Total Protein
Secreted proteins
Pellet
D
Non-crosslinked
Crosslinked
Efm
Eds
Ehe
Emi
Efs
Rel. Intensity
Time, min

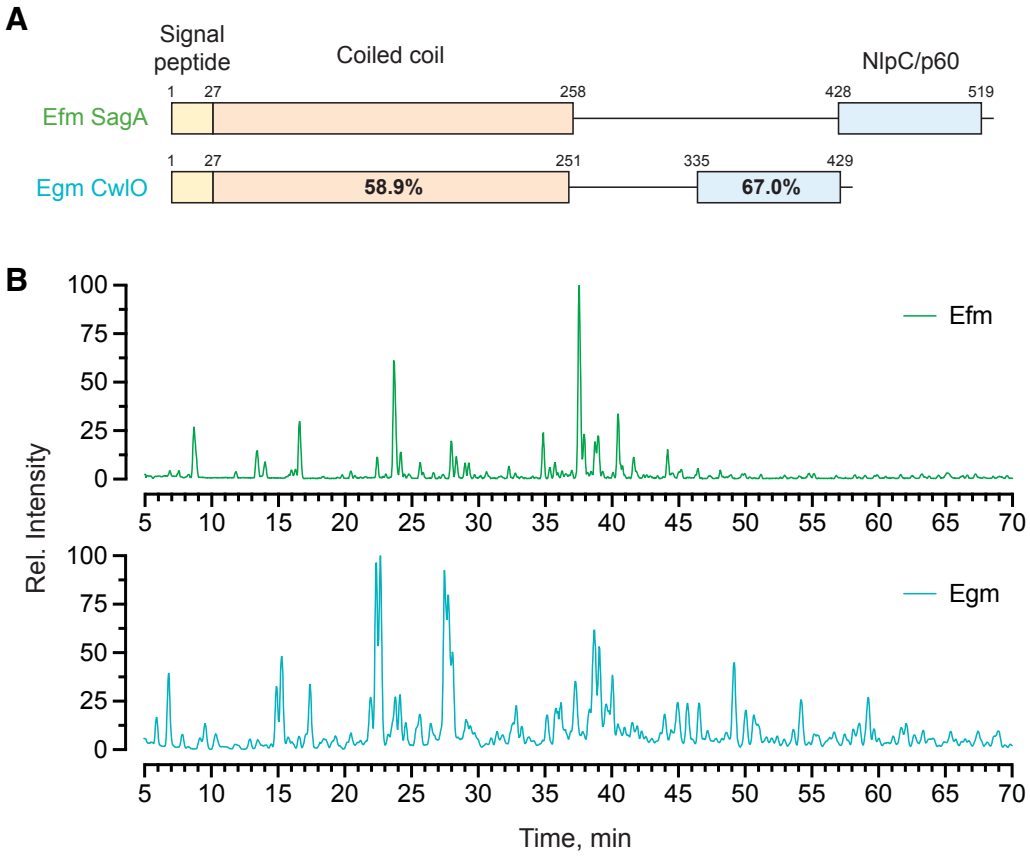


Figure 4.3. Closest NlpC/p60 homolog and PG profile are distinct in *E. gallinarum*. (**A**) Comparison of primary sequence homology and domain architecture of *E. faecium* Com15 SagA and the closest ortholog from *Egm* 2375 similar to Fig. 4.2B. Numbers above each bar denote amino acid residue boundaries of the indicated domains, and percentages represent sequence identity. (**B**) LC-MS ion chromatograms of mutanolysin-treated PG harvested from *E. faecium* Com15 (*Efm*) and *E. gallinarum* 2375 (*Egm*) similar to Fig. 4.2D. Data are shown as relative intensity of base peak ion abundance.

Next, we investigated whether the active enterococci encode catalytically active SagA orthologs. Full-length SagA orthologs or related proteins from the respective *Enterococcus* species were expressed and purified from *E. coli* (**Fig 4.4A**) and incubated with mutanolysin-digested *Efs* OG1RF PG as a heterogenous substrate mixture. The resulting soluble muropeptide products were analyzed by LC-MS, as detailed in previous chapters. In Chapter 3, we showed that Cys-His-Ala mutants of *Eds* and *Emi* SagA displayed PG hydrolase activity (*42*). Predictably, compared to wild type *Efm* SagA as the positive control, wild type *Eds*, *Ehe*, and *Emi* SagA shared similar D,L-endopeptidase activity towards predigested L-Lys type PG *in vitro* (**Fig 4.4B,C and 4.5**). Expression of the *Egm* ortholog in *E. coli* was unsuccessful. *Efs* SalB did not show PG hydrolase activity under our conditions, but SalA was confirmed as a D,L-endopeptidase with different cleavage bond specificity than the SagA orthologs. Whereas the SagA orthologs cleaved the γ-D-Glu-L-Lys bond within peptide stems (**Fig. 4.4C**), SalA hydrolyzed the L-Ala-D-Ala crosslink between muropeptide stems, thus generating monomeric PG fragments (**Fig 4.6**). We speculate that the more crosslinked cell wall composition of *Efs* compared to the SagA+ enterococci may be attributed to SalA PG hydrolase activity, as the resulting monomeric muropeptide products we identified can in theory be recycled and cross-linked again. These results show that the enterococci with enhanced antitumor activity share similar cell wall composition and encode catalytically active SagA orthologs involved in PG remodeling.

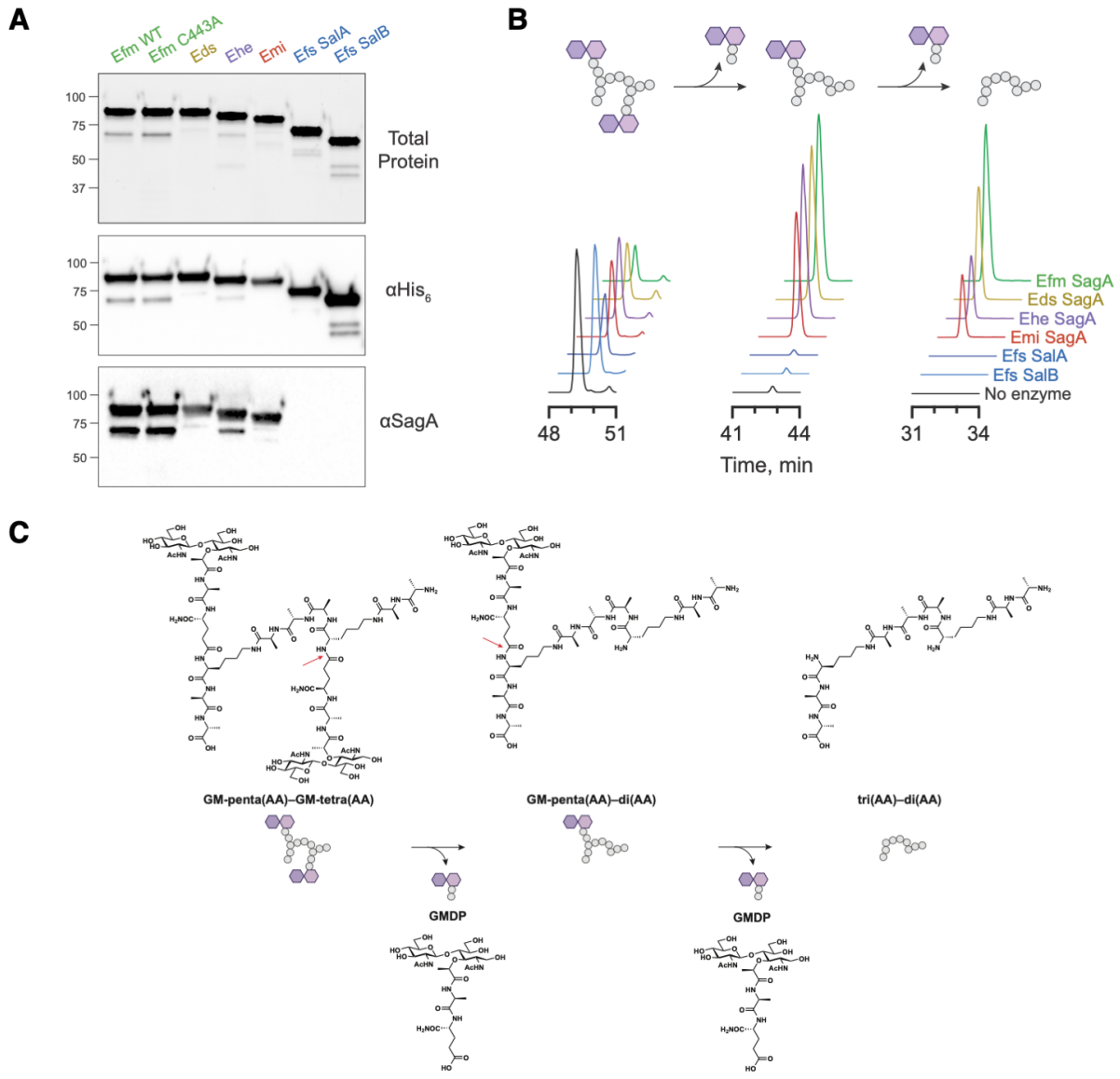


Figure 4.4. SagA orthologs exhibit similar D,L-endopeptidase activity *in vitro*. (**A**) Total protein and Western blotting of full-length SagA orthologs after expression and purification from *E. coli*. The two SagA-like proteins from *Efs* OG1RF, SalA and SalB, do not show reactivity using antiserum raised against *Efm* Com15 SagA. (**B**) *In vitro* activity of purified SagA orthologs. Data are shown as extracted LC-MS ion chromatograms of a crosslinked PG substrate and two iterative hydrolysis products after incubating a mixture of PG fragments with purified SagA orthologs from the indicated species for 16 hours at 37 °C. Peak heights are shown as relative intensity of ion abundance, and all chromatograms are shown on the same scale. (**C**) Chemical structures of PG substrate and products of SagA activity. Chemical structures of the model crosslinked PG substrate (GM-penta(AA)–GM-tetra(AA)) and the two products (GM-penta(AA)–di(AA); tri(AA)–di(AA)) produced by SagA and its orthologs through the iterative hydrolysis of D-isoglutamine–L-lysine amide bonds indicated by the red arrows. Each hydrolysis step releases one equivalent of GlcNAc-muramyl dipeptide (GMDP).

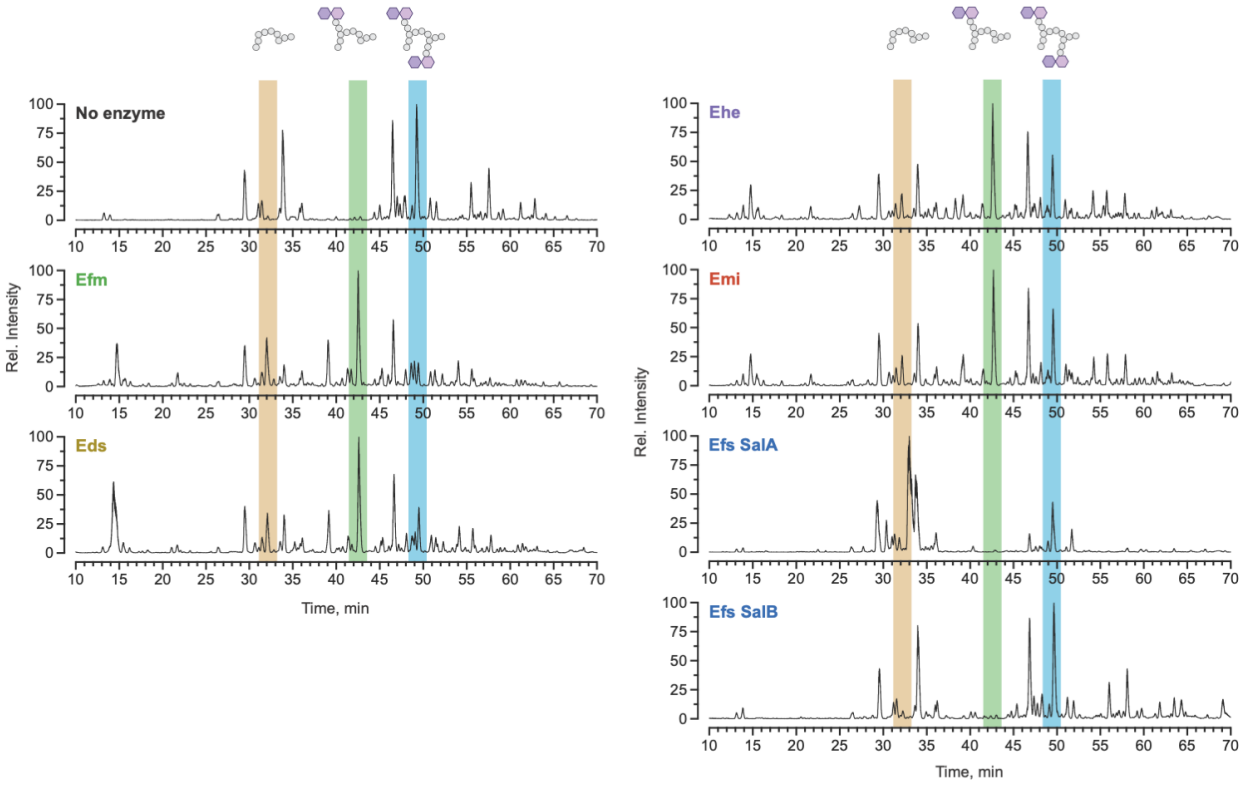


Figure 4.5. Biochemical characterization of SagA orthologs. *In vitro* activity of purified SagA orthologs corresponding to Fig. 4.4. Full LC-MS ion chromatograms of a mixture of PG fragments incubated with purified SagA orthologs from the indicated species for 16 hours at 37 °C. Data are shown as relative intensity of base peak ion abundance. The blue column indicates a crosslinked substrate for SagA (GM-penta(AA)–GM-tetra(AA)) that is cleaved iteratively into two products indicated by the green (GM-penta(AA)–di(AA)) and orange (tri(AA)–di(AA)) columns, corresponding to the loss of one or two GlcNAc-muramyl dipeptide (GMDP) molecules, respectively. See Fig. 4.4C for the corresponding chemical structures.

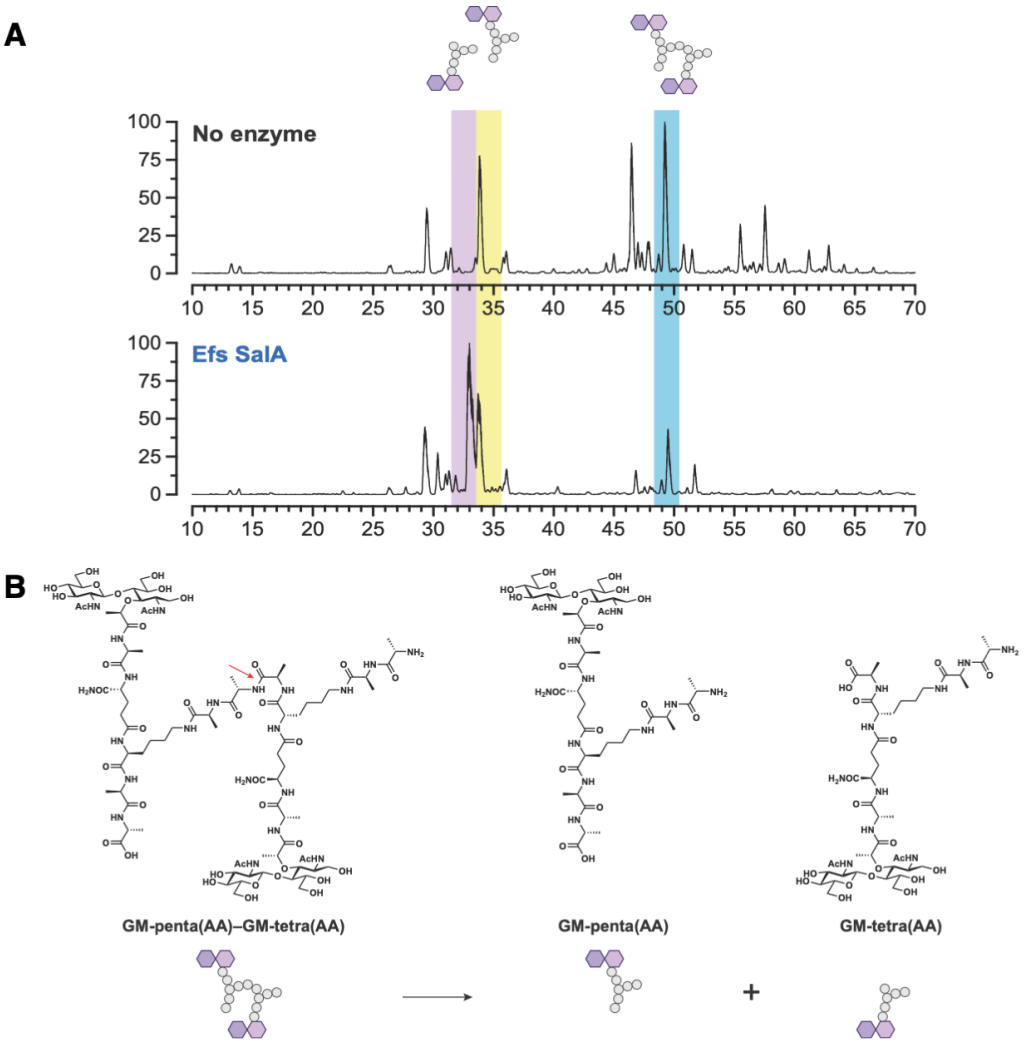


Figure 4.6. *Efs* OG1RF SalA exhibits distinct D,L-endopeptidase activity to generate monomeric PG fragments *in vitro*. (**A**) *In vitro* activity of purified *Efs* OG1RF SalA corresponding to Fig. 4.4B and 4.5. Full LC-MS ion chromatograms of a mixture of PG fragments incubated with purified *Efs* SalA for 16 hours at 37 °C. Data are shown as relative intensity of base peak ion abundance. The blue column indicates a SalA substrate that is cleaved at the cross-bridge into two products indicated by the yellow and purple columns. (**B**) Chemical structures of the model crosslinked PG substrate (GM-penta(AA)–GM-tetra(AA), blue column in LC-MS chromatograms) and the two products (GM-penta(AA), yellow column; GM-tetra(AA), purple column) produced by *Efs* SalA.

The engineered *Efs–sagA* strain expresses and secretes SagA (**Fig 4.2C**) and has less cross-linked PG fragments than parental SagA– *Efs* (*40*), so we used this strain in mouse models of enteric infection to show that SagA was sufficient to confer protection (*40*, *68*, *69*). To determine whether SagA is also sufficient to enhance the efficacy of checkpoint blockade therapy, we compared *Efs–sagA* with non-active *Efs* and active *Efm* in the mouse tumor models described here. Mice supplemented with *Efs–sagA* and not *Efs* had decreased tumor volumes upon anti-PD-L1 treatment similar to mice supplemented with *Efm*, with no differences observed in bacterial load between the animals (*43*). Similar results were found when mice with intact, complex microbiota were used as opposed to antibiotic-pretreated mice (*43*). Although less enterococci were recovered from fecal microbiota samples than before – suggesting that the active enterococci did not need to dominate in the gut for improved antitumor activity – 16S rRNA sequencing showed that the colonization of the enterococci did not significantly alter the composition of the native microbiota (*43*). The versatility of SagA+ bacteria in improving response to checkpoint blockade antibodies was then examined by injecting mice with different cancer cell types and antibodies. Remarkably, mice with subcutaneous tumors consisting of MCA205 fibrosarcoma or MC38 colorectal carcinoma cells had a significant decrease in tumor growth upon injection of anti–PD-1 or anti–CTLA-4, respectively, and supplementation with *Efs–sagA* or *Efm* (*43*). These studies demonstrate that SagA is sufficient for the antitumor activity phenotype in the various mouse tumor models, but the precise mechanism(s) by which SagA modulates that improved response to checkpoint blockade therapy was unclear.

Work from our laboratory showed that the pattern recognition receptor NOD2 was required for the protective activity of *Efm* and SagA in mice (*68*), and we observed that SagA-generated muropeptides could activate NOD2 in mammalian cells *ex vivo* (*40*). Therefore, to interrogate whether SagA PG remodeling was involved in the improved antitumor activity of checkpoint blockade agents, we focused on identifying SagA-mediated activation of NOD2 in the B16-F10 melanoma mouse model. As expected, *Nod2* expression was required for *Efs–sagA* to improve responsiveness to anti–PD-L1 (*43*). However, whether SagA PG hydrolase activity was directly contributing to this phenotype was unknown.

We noted in Chapter 3 that overexpression of the catalytically inactive SagA C443A mutant in *Efm* was unsuccessful, as was the chromosomal insertion of SagA C443A in *Efs* (*40*), which would have enabled us to link SagA PG hydrolase activity to the protective phenotypes observed *in vivo*. The engineered *L. plantarum* (*Lpl*) strains that contain plasmids for overexpression of *sagA* have been instrumental in our previous *in vivo* studies (*40*, *68*, *69*), but chromosomal integration of *sagA* is necessary for the development of stable SagA+ bacteria as therapeutics. Moreover, therapeutic application of *Enterococcus* is limited by concerns regarding the acquisition of antibiotic resistance in *Efm* (*120–122*). To circumvent these issues, and to broaden the applicability of SagA+ bacteria, we collaborated with Rise Therapeutics to engineer *L. lactis* (*Lls*) strains with chromosomal *sagA* insertions for evaluation in the B16-F10 melanoma mouse model. Lactic acid bacteria (LAB) such as *Lls* are commonly used as oral probiotics to deliver therapeutic molecules, such as proteins and enzymes (*123–125*). The engineered *Lls* strains were designed to express secreted, catalytically active (*Lls–sagA WT*) and inactive (*Lls–sagA C443A*) SagA, as well as a secretion-deficient construct (*Lls–sagA* Δ*SS*). These strains abundantly expressed SagA with the WT and C443A constructs showing robust secretion and the ΔSS construct releasing some SagA in the secreted fraction possibly due to cell lysis (**Fig. 4.7A**). Similar to the engineered *Lpl* (*40*), the PG profiles for the engineered *Lls* were the same, suggesting that heterologous SagA expression does not alter the *Lls* cell wall composition in culture (**Fig.**

4.7B). Likewise, expression and secretion of functional SagA was required for the antitumor phenotype; mice supplemented with *Lls–sagA WT* – and not parental *Lls* or *Lls–sagA C443A* – exhibited significantly decreased tumor growth upon anti–PD-L1 treatment similar to *Efm* (**Fig. 4.7C**). Interestingly, supplementation with *Lls–sagA* Δ*SS* led to a significant suppression in tumor growth in mice as well, but not to the extent of *Lls–sagA WT* or *Efm* (**Fig. 4.7C**). The presence of SagA in the culture media of *Lls–sagA* Δ*SS* (**Fig. 4.7B**) suggests that cell lysis and release of SagA may also occur in the gut to lead to this improved anti–PD-L1 response (*43*). These results validate that changes to the native cell wall composition upon heterologous SagA expression are not required for therapeutic effect of engineered SagA+ bacteria. Moreover, SagA PG hydrolase activity may directly contribute to enhanced efficacy of cancer immunotherapy by checkpoint blockade, although whether SagA acts on PG from the host strain, dead bacteria, and or diet from remains to be determined. To further implicate SagA PG hydrolase activity in this improved antitumor response, NOD2-active and inactive isomers of MDP, which is reminiscent of a major SagA D,L-endopeptidase product, GMDP (*40*, *42*), were co-injected with anti–PD-L1, and only the NOD2-active MDP-L,D isomer decreased tumor growth in mice (*43*). Therefore, our results collectively suggest that SagA-generated muropeptides may activate NOD2 to stimulate the host immune system and lead to the synergistic effect in suppressing tumor growth during cancer immunotherapy (*43*).

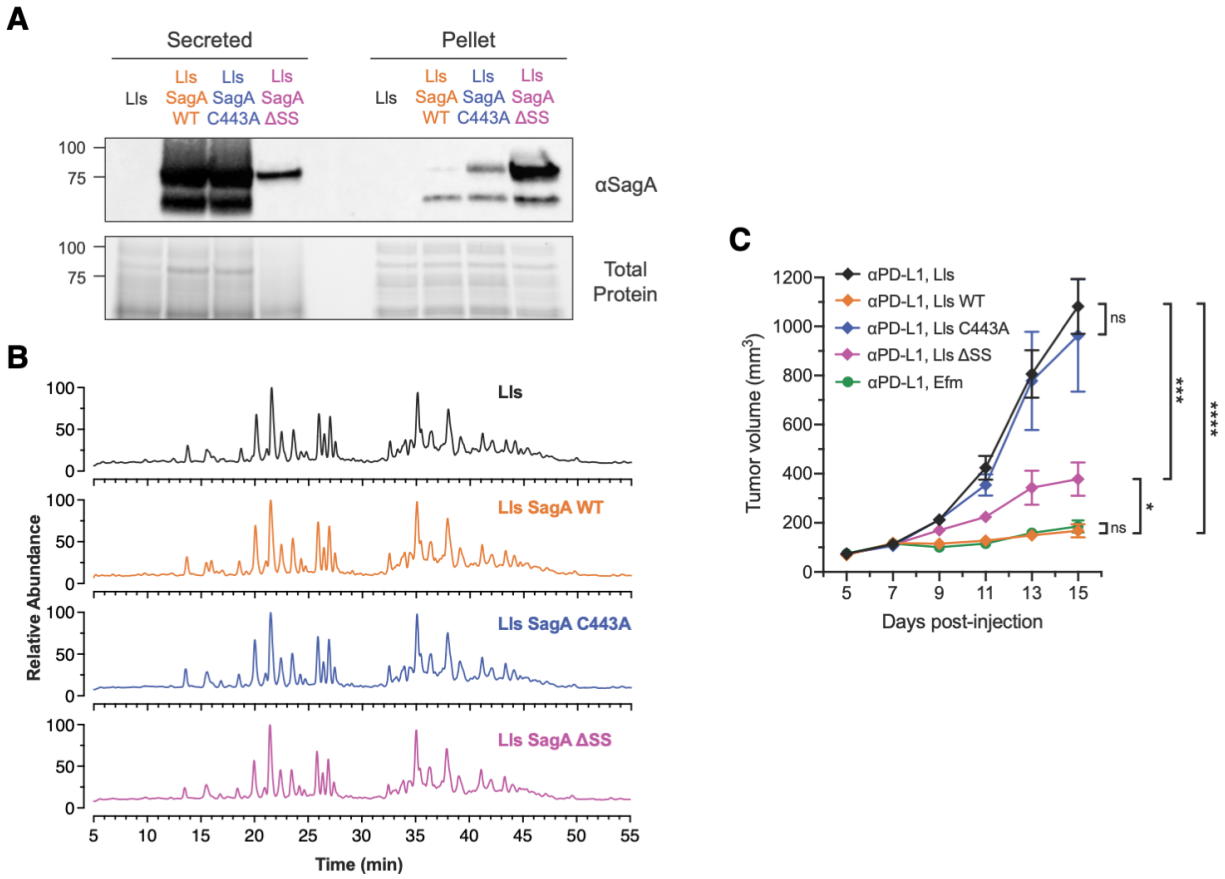


Figure 4.7. *L. lactis-SagA* abundantly secretes SagA without altering the native cell wall profile in culture and enhances immunotherapy activity *in vivo*. (**A**) Western blot detection of ectopically expressed Efm Com15 SagA in secreted protein and cell pellet fractions harvested from overnight cultures of the indicated engineered *Lactococcus lactis* (*Lls*) strains using antiserum raised against *Efm* Com15 SagA. Bottom panels show total protein loading. Numbers indicate estimated molecular weight (kDa). WT = wild-type, ΔSS = signal sequence deletion. (**B**) LC-MS ion chromatograms of mutanolysin-treated PG harvested from *Lls* strains shown in (A). (**C**) B16-F10 tumor growth in antibiotic-treated mice that were supplemented with the indicated Lls strains and treated with anti–PD-L1 starting on day 7. *n* = 9-11 mice per group.

In conclusion, our exploration of the role *Enterococcus* species may play in cancer immunotherapy revealed that *Enterococcus* and SagA can enhance the efficacy of checkpoint blockade agents. Using a mouse tumor model, we demonstrated that SagA+ enterococci in particular have a synergistic effect with checkpoint blockade antibodies such as anti–PD-L1. These enterococci abundantly secrete SagA orthologs with highly conserved catalytic NlpC/p60 domains. Moreover, the cell wall of the active, SagA+ enterococci is composed of more non-crosslinked muropeptides compared to non-active *Efs*, which we previously found can have an immunostimulatory effect *ex vivo* (*40*). Using our established protocols for characterizing NlpC/p60 hydrolase activity, we confirmed that the full-length, wild type SagA orthologs exhibit D,L-endopeptidase activity towards predigested L-Lys-type PG *in vitro*. This analysis also showed *Efs* SalA can cleave at the cross-link site of multimeric PG fragments to generate monomeric muropeptides, which we propose may be cross-linked again and reincorporated into the *Efs* cell wall, thus distinguishing its cell wall composition from the active SagA+ enterococci. The use of an engineered *Efs–sagA* strain expressing SagA ultimately confirmed that SagA is sufficient to improve responsiveness to multiple checkpoint blockade antibodies in tumors of various cancer cell types, and that the antitumor phenotype is dependent on the pattern recognition receptor NOD2. Engineered *Lls* strains with chromosomal *sagA* insertions allowed us to determine that secretion of catalytically active SagA is required for improved antitumor activity, which indicated that SagA-generated muropeptides may be responsible for stimulating the host immune system. Accordingly, co-injection of anti–PD-L1with the NOD2-active MDP-L,D that is similar to SagA-generated GMDP suppressed tumor growth in mice, further linking SagA PG hydrolase activity with improved checkpoint blockade response. Collectively, these studies reinforce the therapeutic potential of *Enterococcus* and SagA and provide a framework for studying the impact of secreted NlpC/p60 hydrolases from commensal bacteria in other disease models.

Acknowledgements

We thank Dr. Caitlin Steckler and Dr. Henrik Molina of the Rockefeller Proteomics Resource Center for assistance with PG analysis.

Chapter 5. Investigation of SagA-NlpC/p60 mutants from phage resistant *E. faecium* Com12 with enhanced antibiotic susceptibility.

Introduction

The genus *Enterococcus* is highly adaptable and prevalent in nature, as it is found in many animals and throughout the environment (*111*, *122*). In this thesis, we have discussed how commensal *Enterococcus* species have been developed as probiotics for animals and humans and have demonstrated beneficial effects in mouse models of enteric infection (*40*, *68*, *69*) and cancer (*43*). Of these commensal enterococci, *Efm* and *Efs* are most common in the human gut microbiota. Unfortunately, pathogenic *Enterococcus* species can acquire resistance to drugs in health care settings (*120*), and pathogenic *Efm* and *Efs* strains in particular are increasingly present in this context (*122*, *126*). Pathogenic enterococci have developed mechanisms to evade susceptibility to multiple antibiotics, and the speed with which they adapt to new therapies has accelerated, limiting clinicians in the options they have to treat patients (*122*). The antibiotic vancomycin, or even combination antibiotic therapy, have been used as last resorts to treat severe bacterial infections due to dysbiosis of the gut microbiota and the ensuing adverse health effects (*127*, *128*). Alarmingly, vancomycin-resistant *Enterococcus* (VRE) strains of *Efm* and *Efs* are becoming more widespread and pose a major health risk especially towards immunocompromised individuals, such as cancer patients (*129*) and residents of nursing facilities (*126*). VRE can also serve as a reservoir for the acquisition of vancomycin resistance in other pathogenic bacteria (*126*), further compounding their pathogenicity. Therefore, there is a pressing need for alternative therapies to treat VRE and multi-drug resistant (MDR) enterococci.

Bacteriophage-based therapy has recently emerged as a promising option to treat MDR bacterial infection. Bacteriophage (phage) with lytic activity against MDR bacteria have been identified (*130–133*), and even used in combination with low doses of antibiotics in efforts to prevent the acquisition of antibiotic resistance in the targeted pathogens (*134*). What remains elusive, however, is how exactly phage therapy functions in this context, which is imperative to know so that phage therapy may be applied towards combating VRE infection while ensuring that resistance does not develop. In this chapter, we describe our collaboration with the laboratory of Dr. Breck Duerkop at the University of Colorado School of Medicine in which we investigated the mechanisms underlying the acquisition of resistance to phage therapy in *Efm* in order to establish a foundation for the development of targeted phage therapy against VRE. We found that mutations to the conserved NlpC/p60 domain of *sagA* in *Efm* Com12 afforded the bacteria resistance to a particular phage. These phage resistant strains abundantly secreted SagA and shared PG profiles similar to wild type *Efm* Com12 in culture. Remarkably, *sagA* mutation and phage resistance arose with a fitness tradeoff by sensitizing the strains to antibiotics. Most important, phage therapy synergized with both cell wall and membrane-acting antibiotics and validated SagA as a target for potentially treating VRE. This work was published (*135*), and follow up biochemical studies have revealed that the SagA mutants are inactive *in vitro*, which raises the possibility that *Efm* with impaired SagA may be viable, but more vulnerable to currently available antibiotics.

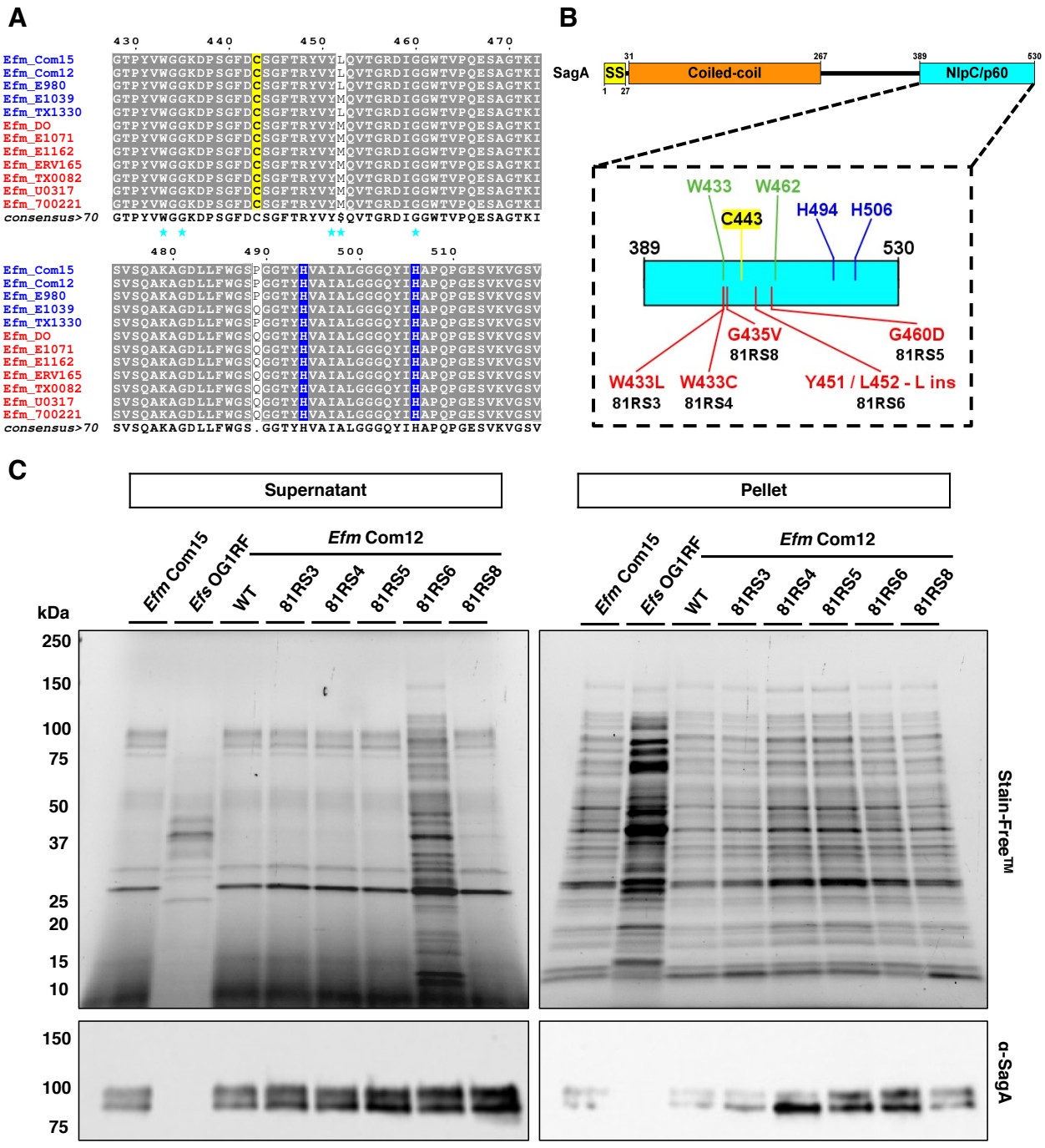
Results and Discussion

To identify phage with lytic activity towards *Efm*, the Duerkop laboratory isolated three phages from raw sewage (9181, 9183, and 9184) (*135*) using plaque assays with commensal, clade B strains of *Efm* (Com12 and 1,141,733) (*136*), which have been shown to act as antibiotic-resistance reservoirs for pathogenic, hospital-associated clade A *Efm* strains (*137*). Of note, phage

9181 showed a narrow host range, as it specifically infected and killed *Efm* Com12 and *Efm* Com15 in a phage susceptibility assay using a panel of laboratory and MDR clinical isolates of *Efm* (*135*). To identify genes involved in the phage 9181-enterococcus interaction, the Duerkop laboratory generated spontaneous phage 9181 resistant *Efm* Com12 strains, which involved using phage 9181-embedded Todd Hewitt Broth (THB) agar to recover stable, resistant isolates (*135*). Six phage 9181 resistant *Efm* Com12 strains were isolated (*Efm* Com12 81R3-R8), four of which had mutations localized to the NlpC/p60 domain of *Efm* Com12 (*135*).

Multiple sequence analysis of *sagA* across commensal and MDR *Efm* shows that the catalytic NlpC/p60 domain is virtually identical (**Fig 5.1A**) and implies that the findings with regard to the *sagA* mutations in the phage 9181 resistant *Efm* Com12 strains may be applied to target other *Efm* strains. Specifically, *Efm* Com12 81R3 and *Efm* Com12 81R4 contain mutations at the conserved Trp433 (W433L and W433C, respectively) that we previously determined may serve as a clamp for binding PG peptide stems and is vital for SagA-NlpC/p60 PG hydrolase activity *in vitro* (**Fig 5.1B**) (*40*). Intriguingly, *Efm* Com12 81R8 has a mutation (G435V) adjacent to the required Trp433, while the *Efm* Com12 81R5 mutation (G460D) is proximal to the second required PG peptide clamp, Trp462 (**Fig 5.1B**) (*40*). *Efm* Com12 81R5, on the other hand, has a codon duplication where a Leu is embedded between the uncharacterized Tyr451 and Leu452 (**Fig 5.1B**). *In silico* analysis using the SagA-NlpC/p60 structure showed that these mutations may not have a negative impact on SagA structure and function, so the individual strains may still be able to express SagA (*135*). We subsequently confirmed that the phage 9181 resistant *Efm* Com12 strains abundantly secrete SagA (supernatant) and retain intracellular SagA (pellet) similar to wild type (**Fig 5.1C**) (*135*). Since the *sagA* mutations are localized at or near residues deemed important for SagA PG hydrolase activity, we hypothesized that the phage 9181 resistant *Efm* Com12 strains expressed catalytically inactive or impaired SagA, which would somehow manifest as phage 9181 resistance. Complementation of these strains with the pAM401 plasmid the Hang laboratory designed for overexpression of wild type SagA under its native promoter (p*sagA*) recovered phage susceptibility in the *Efm* Com12 strains, suggesting that catalytic SagA may be dispensable for *Efm* viability, but key for susceptibility to phage 9181 (*135*).

Figure 5.1. Phage resistant *Efm* Com12 strains encode *sagA* mutants and abundantly express SagA. (**A**) The SagA-NlpC/p60 domain is highly conserved in commensal and multi-drug resistant strains of *Efm*. SagA ortholog sequences were located downstream of *mreD* locus in commensal and multi-drug resistant strains of *Efm* using BioCyc. NlpC/p60 domains were determined using InterPro. Multiple sequence alignment was performed using ClustalOmega and analyzed with ESPript. Conserved residues relevant to SagA mutations in phage 9181 resistant *Efm* Com12 strains marked with cyan star. (**B**) Domain architecture schematic of *Efm* Com15 SagA with inset representing conserved SagA-NlpC/p60 domain. Biochemically characterized residues labeled on the top side of the SagA-NlpC/p60 domain as follows: PG peptide stem clamps (W433 and 4462) in green, catalytic cysteine (C443A) in yellow, and histidines (H494 and H506) in blue. SagA mutations labeled in red below the SagA-NlpC/p60 domain with the corresponding phage 9181 resistant *Efm* Com12 strain labeled in black. (**C**) Expression of SagA in secreted protein (supernatant) and cell pellet (pellet) fractions harvested from midlog phase (OD_{600} 0.6) cultures of the indicated *Enterococcus* species and strains. Top panel (Stain-Free^{TM}) represents total protein loading. Bottom panel (α-SagA) represents Western blot detection of SagA using antiserum raised against recombinant *Efm* Com15 SagA. Numbers indicate estimated molecular weight (kDa).


A
B
C
SagA
SS
Coiled-coil
NlpC/p60
W433
W462
C443
H494
H506
389
530
G435V
81RS8
G460D
81RS5
W433L
81RS3
W433C
81RS4
Y451 / L452 - L ins
81RS6
Supernatant
Pellet
Efm Com15
Efs OG1RF
Efm Com12
WT
81RS3
81RS4
81RS5
81RS6
81RS8
kDa
250
150
100
75
50
37
25
20
15
10
Stain-Free™
α-SagA

To study if *sagA* mutation and phage 9181 resistance coincide with a fitness tradeoff, the Duerkop laboratory evaluated whether the phage 9181 resistant mutants were more susceptible to antibiotics compared to parental *Efm* Com12. The phage 9181 resistant *Efm* Com12 harboring *sagA* mutations displayed enhanced susceptibility to the cell wall-targeting β-lactam antibiotics ampicillin and ceftriaxone, and to the membrane-targeting lipopeptide drug daptomycin (*138*). Combination therapy of phage 9181 with sub-inhibitory concentrations of either of the three drugs had a synergistic effect in killing wild type *Efm* Com12, thus demonstrating the power of phage-antibiotic combination therapy (*135*). We recognize that the narrow host range of 9181 limits its applicability to treat more clinically relevant *Efm* strains, such as VRE. However, these results offer *sagA* mutation and its associated antibiotic fitness cost as a potential avenue for treating VRE.

Given these developments, we sought to evaluate the role SagA PG hydrolase activity – or lack thereof – may play in the phage 9181 resistance and enhanced antibiotic susceptibility observed in the *Efm* Com12 strains harboring *sagA* mutations. We first analyzed the PG profiles of the strains to ascertain whether the antibiotic fitness cost may be attributed to changes in the cell wall upon expression of SagA mutants. LC-MS analysis of soluble muropeptides from log phase (OD_{600} ~0.8) cultures of the phage 9181 *Efm* Com12 strains uncovered that they share similar cell wall composition to the parental *Efm* Com12 (**Fig. 5.2**). Whether alteration to the cell wall composition manifests only during exposure to antibiotics remains to be determined. Next, we used the recombinant *Efm* Com15 SagA-NlpC/p60 construct we developed since it is identical to that of *Efm* Com12 (**Fig 5.1A**) and generated mutants corresponding to each phage 9181 resistant *Efm* Com12 strain (**Fig. 5.3A**) to gauge whether these mutations exhibited impaired the D,L-endopeptidase activity *in vitro*. All mutants were inactive against predigested *Efm* Com12 PG serving as the endogenous substrate (**Fig. 5.3B,C and 5.4**) and predigested *Efs* OG1RF PG (**Fig. 5.5**).

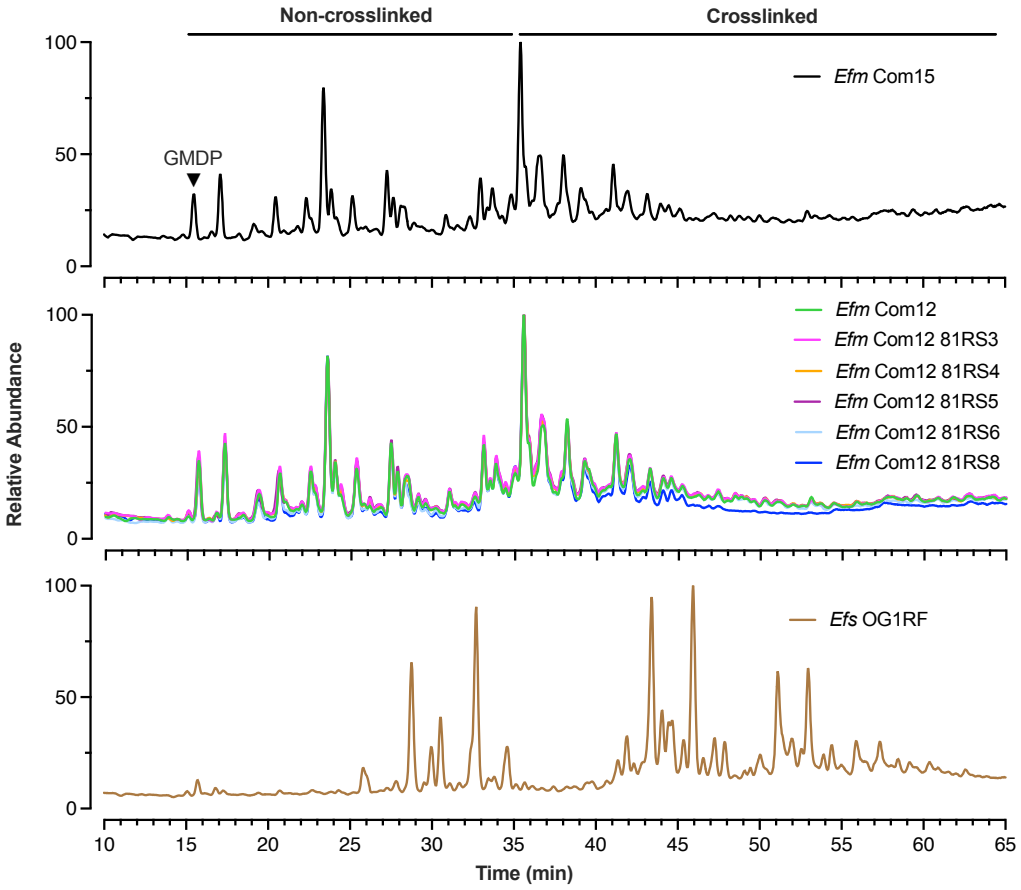


Figure 5.2. Phage 9181 resistant *Efm* Com12 strains share similar PG profiles compared to wild type. LC-MS ion chromatograms of digested PG harvested from cultures of the indicated *Enterococcus* species and strains in Fig 5.1C. Data are shown as relative intensity of base peak ion abundance. GMDP, Non-crosslinked, and crosslinked PG are marked at top.

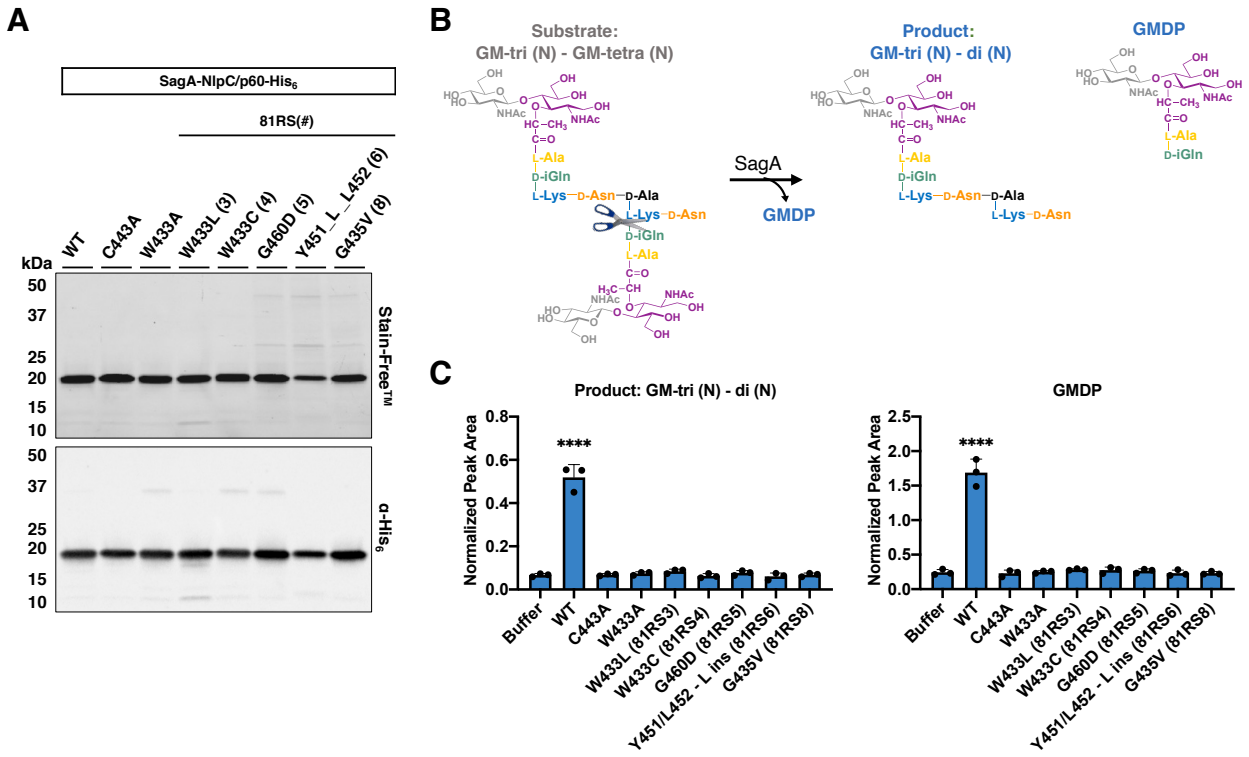


Figure 5.3. SagA-NlpC/p60 mutants demonstrate abrogated D,L-endopeptidase activity towards predigested *Efm* Com12 PG *in vitro*. (A) Stain-Free (Bio-Rad) SDS-PAGE and α-His$_6$ immunoblotting of SagA-NlpC/p60 constructs purified from the cell-free lysate of BL21(DE3)-RIL *Escherichia coli*. The phage 9181 resistant *Efm* Com12 strain corresponding to each mutant is listed in parenthesis. (**B**) Schematic of *in vitro* SagA D,L-endopeptidase activity. Substrate (GM-tri (N) – GM-tetra (N)) represents the most abundant of the soluble, cross-linked muropeptides released following mutanolysin digestion of intact *Efm* Com12 PG. Cleavage of substrate by SagA generates GlcNAc-MurNAc-dipeptide (GMDP) and Product (GM-tri (n) – di (N)). Muropeptides are illustrated with reduced anomeric carbon of MurNAc. (**C**) Quantification of product formation by SagA-NlpC/p60 *in vitro* (n = 3). For each condition, peak areas were recorded from extracted ion chromatograms of substrate and products and normalized to that of the unchanging GM-tri (N). Data were analyzed with one-way ANOVA and Tukey's multiple comparisons post-hoc test. Centerline, average. Error bar, SD ****P < 0.0001. Comparisons with no asterisk had P > 0.05 and were not considered significant.

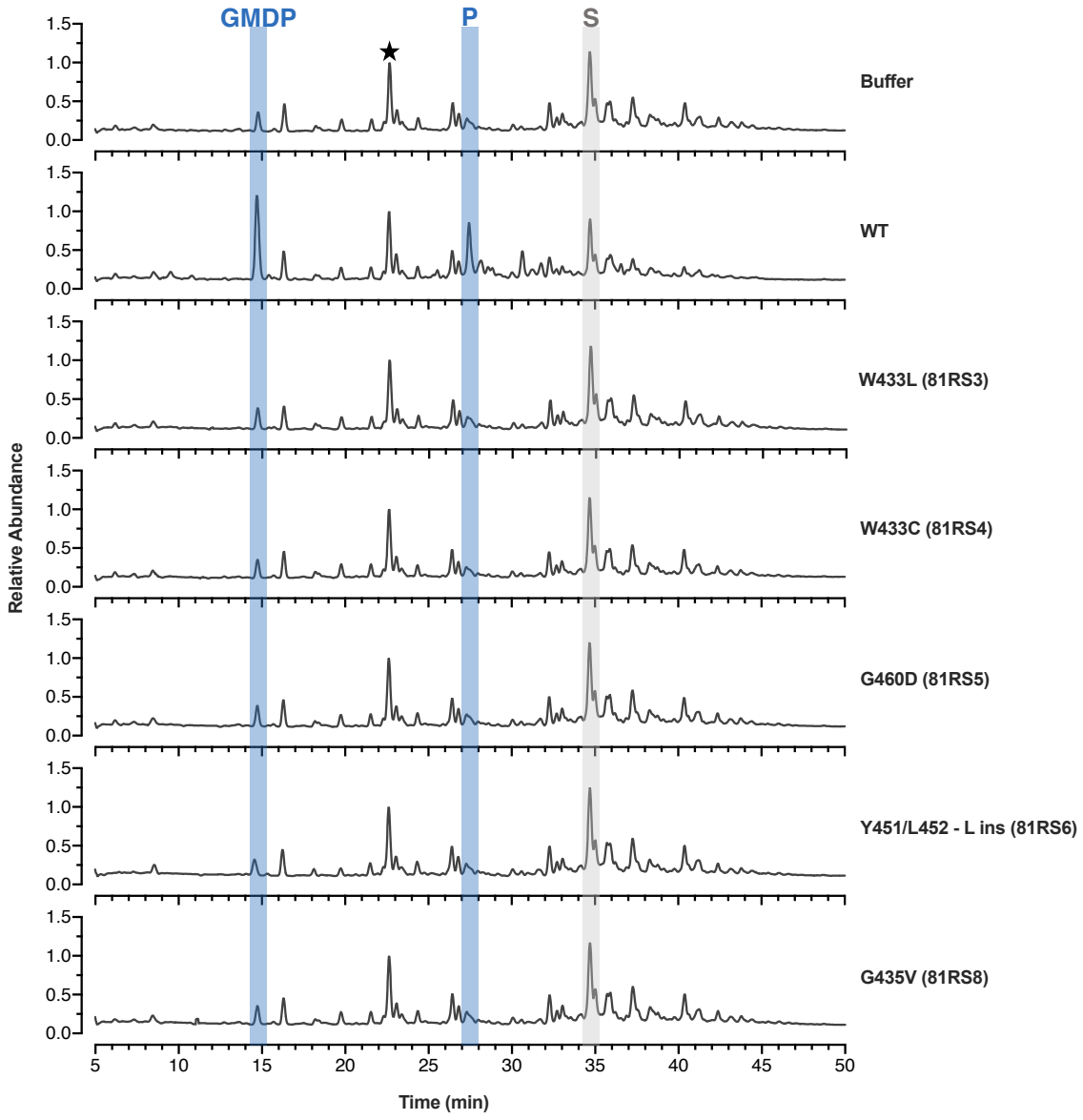
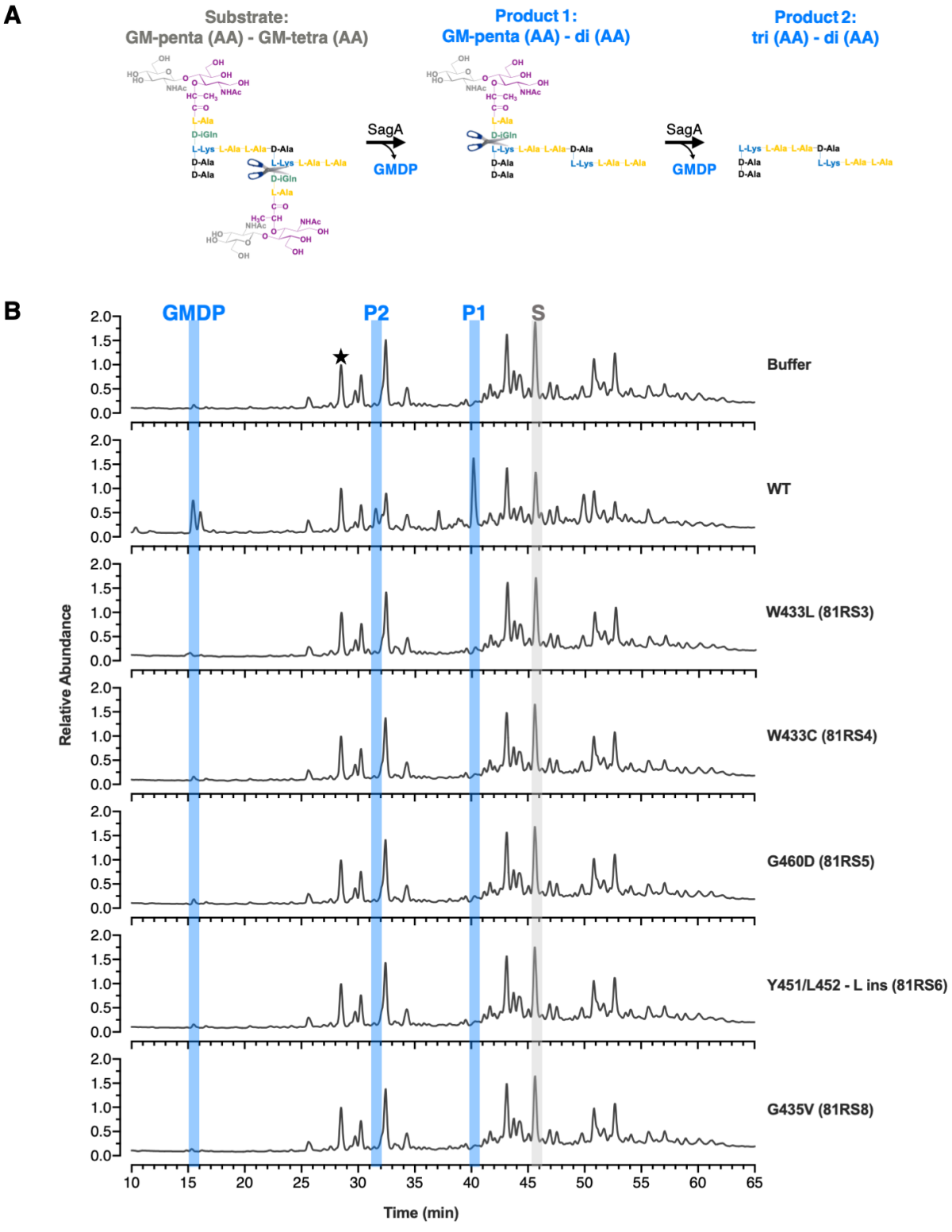


Figure 5.4. Biochemical characterization of *Efm* Com12 SagA-NlpC/p60 mutants incubated with predigested *Efm* Com12 PG. LC-MS analysis of soluble mutanolysin-digested *Efm* Com12 PG incubated with SagA-NlpC/p60 constructs at 37˚C for 16 hr. Data are shown as total ion chromatograms. The intensity of extracted ion abundance was normalized to the unchanging GM-tri (N) peak marked with a star. Colored columns indicate PG substrate (S, grey) and products (GMDP and P, blue) as described in Fig. 5.3B,C.

Figure 5.5. Biochemical characterization of *Efm* Com12 SagA-NlpC/p60 mutants incubated with predigested *Efs* OG1RF PG. (**A**) Schematic of *in vitro* SagA D,L-endopeptidase activity. Substrate (GM-penta (AA) - GM-tetra (AA)) represents the most abundant of the soluble, cross-linked muropeptides released following mutanolysin digestion of intact *Efs* OG1RF PG. Iterative cleavage of Substrate by SagA generates GlcNAc-MurNAc-dipeptide (GMDP) and two fragments, Product 1 (GM-penta (AA) - di (AA)) and Product 2 ((tri (AA) - di (AA)). Substrate and Product 1 are illustrated with reduced anomeric carbon of MurNAc. (**B**) LC-MS analysis of soluble mutanolysin-digested *Efm* Com12 PG incubated with SagA-NlpC/p60 constructs at 37˚C for 16 hr. Data are shown as total ion chromatograms. The intensity of extracted ion abundance was normalized to the unchanging GM-tri (AA) peak marked with a star. Colored columns indicate PG substrate (S, grey) and products (GMDP, P1, and P2 blue).


A
Substrate:
GM-penta (AA) - GM-tetra (AA)
SagA
GMDP
Product 1:
GM-penta (AA) - di (AA)
SagA
GMDP
Product 2:
tri (AA) - di (AA)
B
GMDP
P2
P1
S
Buffer
WT
W433L (81RS3)
W433C (81RS4)
G460D (81RS5)
Y451/L452 - L ins (81RS6)
G435V (81RS8)
Relative Abundance
Time (min)

These results have several implications towards our understanding of SagA PG hydrolase activity in *Efm* function. First, it is possible that the SagA mutants encoded by the phage 9181 resistant *Efm* Com12 strains retain basal PG hydrolase activity that is below the limit of detection in our *in vitro* assay. An alternative hypothesis is that SagA PG hydrolase activity itself is dispensable for viability entirely. In either case, our findings raise the question if compensatory PG hydrolases exist or are expressed under these conditions for proper *Efm* cell division and function. There is evidence for expression of dispensable PG hydrolases in *M. tuberculosis*, and silencing of the major NlpC/p60 hydrolase, RipA, sensitized the pathogen to antibiotics (*23*). Further functional experiments and genomic analyses are required to distinguish between these possibilities in *Efm*, and recent advancements towards gene editing in *Efm* from the Hang laboratory (*139*) offer promise towards elucidating the precise function of SagA PG hydrolase activity in the acquisition of 9181 resistance and antibiotic susceptibility. Nevertheless, this biochemical evidence of impaired SagA NlpC/p60 PG hydrolase activity indicates that *sagA* mutation may be impacting the endogenous PG remodeling in *Efm*, which can motivate future studies on SagA and NlpC/p60 hydrolases as candidate VRE targets.

Acknowledgements

We thank the Duerkop laboratory for the opportunity to collaborate on this exciting project.

Chapter 6. Biochemical characterization of NlpC/p60 hydrolases as potential MDR *E. faecium* targets.

Introduction

Our first foray into studying SagA and NlpC/p60 hydrolase activity in the context of VRE and MDR *Enterococcus* was promising. We established that SagA+ bacteria can have beneficial host-microbe interactions (*40*, *43*, *68*, *69*), but we now have evidence that *sagA* in *Efm* may be exploited to target pathogenic SagA+ bacteria (*135*). This line of reasoning motivated us to investigate whether *Efm* strains encode other NlpC/p60 hydrolases with similar or divergent activity to *Efm* Com15 SagA. Our workflow for the biochemical characterization of NlpC/p60 hydrolases (*41*) may then be applied to identify other catalytically active NlpC/p60 hydrolases as potential MDR *Efm* targets. Subsequent functional studies using RecT-mediated deletion and CRISPRi knockdown of the candidate NlpC/p60 hydrolases using gene editing tools recently developed in our laboratory (*139*) could then validate whether they play a role in MDR *Efm* pathogenicity. Fortunately, several MDR *Efm* strains have robust online genome support that would facilitate the genomic analyses required to begin to address our questions regarding the existence of compensatory or redundant NlpC/p60 hydrolases in MDR *Efm*. To that end, comparative genomics of NlpC/p60 hydrolases in commensal and MDR strains of *Efm* by our laboratory revealed several clusters of conserved NlpC/p60 hydrolases. SagA NlpC/p60 hydrolases were conserved across both commensal and MDR *Efm*, as are previously unidentified putative NlpC/p60 amidases. Interestingly, bifunctional CwlT-like NlpC/p60 hydrolases (*37*, *38*) were identified to be uniquely encoded by readily-available MDR *Efm* strains, such as *Efm* TX0082, which has previously been isolated from endocarditis patients (*140*), *Efm* ERV165, and *Efm* E1162. These strains shared protein expression and cell wall architecture similar to commensal *Efm* Com15 in culture, suggesting that any changes to these markers of PG hydrolase activity may have to be reassessed in cultures grow with antibiotics. To confirm if the NlpC/p60 proteins are functional PG hydrolases, we used recombinant truncation constructs of a putative NlpC/p60 amidase (501F) and CwlT ortholog (51B9 CwlT) from *Efm* TX0082 in our LC-MS *in vitro* activity assay. We found that the NlpC/p60 domains of both encoded D,L-endopeptidase activity dependent on the conserved catalytic cysteine *in vitro*. While the 501F truncation construct exhibited PG hydrolase activity comparable to SagA, full-length 51B9 CwlT fully hydrolyzed all soluble muropeptides in the heterogenous PG mixture used as substrate. Furthermore, the 51B9 CwlT-NlpC/p60 domain alone was sufficient for this robust PG hydrolase activity towards both polymeric and monomeric PG fragments with either tripeptide or pentapeptide stems *in vitro*. Our biochemical studies therefore introduce 501F and 51B9 CwlT in particular as catalytic NlpC/p60 hydrolases that could be explored as MDR *Efm* targets.

Results and Discussion

To identify if other NlpC/p60 hydrolases are conserved in commensal and MDR *Efm*, we performed bioinformatic analysis of the NlpC/p60 family across the genomes of readily-available commensal and clinical isolate, MDR *Efm* strains (**Fig. 6.1A**). Comparative genomics revealed three clusters of highly conserved NlpC/p60 hydrolases. SagA NlpC/p60 hydrolases with virtually identical NlpC/p60 domains were conserved in both commensal and pathogenic *Efm*, as we determined in previous sequence analysis in Chapter 5. A cluster of uncharacterized putative NlpC/p60 amidases with unique domain architecture was also identified in both commensal and MDR *Efm* (**Fig. 6.1A,B**). These putative NlpC/p60 amidases contain a secretion signal peptide

similar to SagA, followed by an extensive, disordered region with Ser/Thr and Asn/Ser/Thr-rich sections, a centrally located NlpC/p60 domain with ~60% sequence identity to SagA-NlpC/p60 (**Fig. 6.1C**), and a C-terminal domain with homology to amidases according to InterPro domain analysis (*141*) that is flanked by another disordered region (**Fig. 6.1B**). The possible bifunctional activity of this class of NlpC/p60 proteins, in conjunction with their availability across all the *Efm* strains in question, signifies that they may have a specific role in *Efm* PG remodeling that is yet to be determined.

Most interesting from the genomic analysis was the observation that CwlT-like NlpC/p60 hydrolases are uniquely conserved in the MDR *Efm* strains (**Fig. 6.1A,B**). Chapter 1 of this thesis describes *B. subtilis* (*Bs*) CwlT as a bifunctional hydrolase that encodes both muramidase and NlpC/p60 domains, which catalyze the hydrolysis of the MurNAc-GlcNAc glycosidic linkage, and *m*DAP-type PG peptide stems in *Bs* PG, respectively (*38*). Structures of *S. aureus* and *C. difficile* CwlT orthologs have been solved to show that the two catalytic domains likely do not interact due to the short linker between them (*37*). Functional studies of CwlT in *Bs* and *B. anthracis* have elucidated the contributions of the CwlT domain architecture in conjugation of the mobile genetic element, ICE*Bs*1, that encodes *Bs cwlT* (*55*). The CwlT transmembrane domain was found to be essential for ICE*Bs*1 conjugation, as was the muramidase activity of CwlT (*55*). The endopeptidase activity of the CwlT-NlpC/p60 domain, however, was found to be important, but not essential for ICE*Bs*1 conjugation (*55*). These findings are of particular to interest to us because bacterial conjugation is a major driver of horizontal gene transfer (HGT) that can lead to the transfer of antibiotic resistance between bacteria (*142*). Bacterial conjugation mediated by CwlT PG hydrolase activity in *B. subtilis* therefore raises the possibility that CwlT can play a similar role in the virulence of MDR *Efm*.

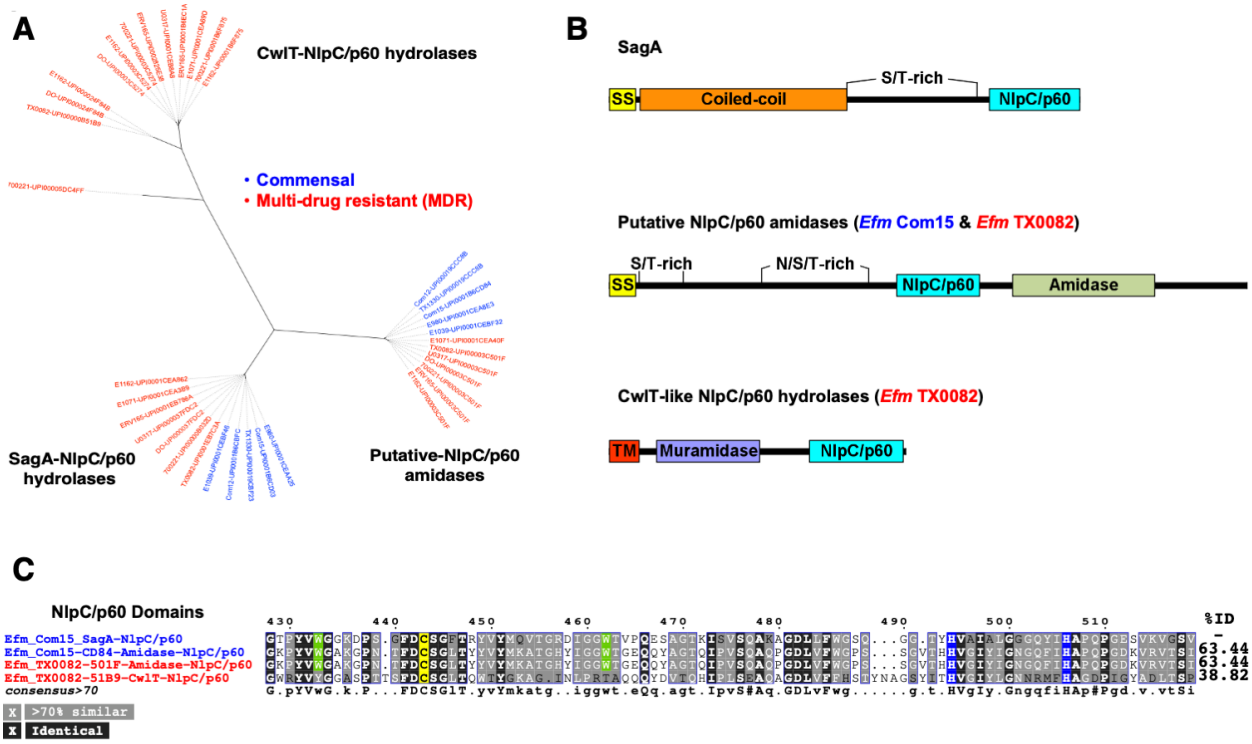


Figure 6.1. Comparative genomics of NlpC/p60 hydrolases in *Efm* strains. (**A**) SagA NlpC/p60 hydrolases and putative-NlpC/p60 amidases are expressed in commensal (blue) and MDR (red) strains of *Efm* strains. In contrast, CwlT-like NlpC/p60 hydrolases are uniquely expressed in MDR strains of *Efm* strains. (**B**) Domain architecture schematic of NlpC/p60 hydrolases. SS, secretion signal; TM, transmembrane helix. (**C**) Multiple sequence analysis of NlpC/p60 domains. NlpC/p60 domains were determined using InterPro. Multiple sequence alignment was performed using ClustalOmega and analyzed with ESPript. Percent identity of residues to *Efm* Com15 SagA-NlpC/p60 listed to the right of each sequence.

Next, we sought to study if the MDR *Efm* displayed differences in the common markers of PG hydrolase activity in *Efm*, specifically SagA expression and secretion and cell wall architecture. The MDR *Efm* abundantly secreted SagA (supernatant) and retained intracellular SagA (pellet) in culture compared to commensal *Efm* Com15 (**Fig 6.2A**). RNA-seq and secretomics analyses are required to confirm expression of the putative NlpC/p60 amidase and CwlT orthologs in the MDR *Efm*, but no obvious differences in the protein expression profiles were observed here for cultures grown to midlog phase (OD_{600} ~0.6). Likewise, the PG profiles of all *Efm* strains were similar (**Fig 6.2B**). We could not deduce from the cell wall profiles of the MDR *Efm* whether the uniquely encoded CwlT was involved in PG remodeling. Given that some VRE *Efm* (*143*) and *Efs* (*144*, *145*) strains have been shown to have modified PG composition when grown in conditions that trigger their antibiotic resistance genes, it is likely that we need to reevaluate these phenotypes in the MDR *Efm* strains using cultures grown with antibiotics.

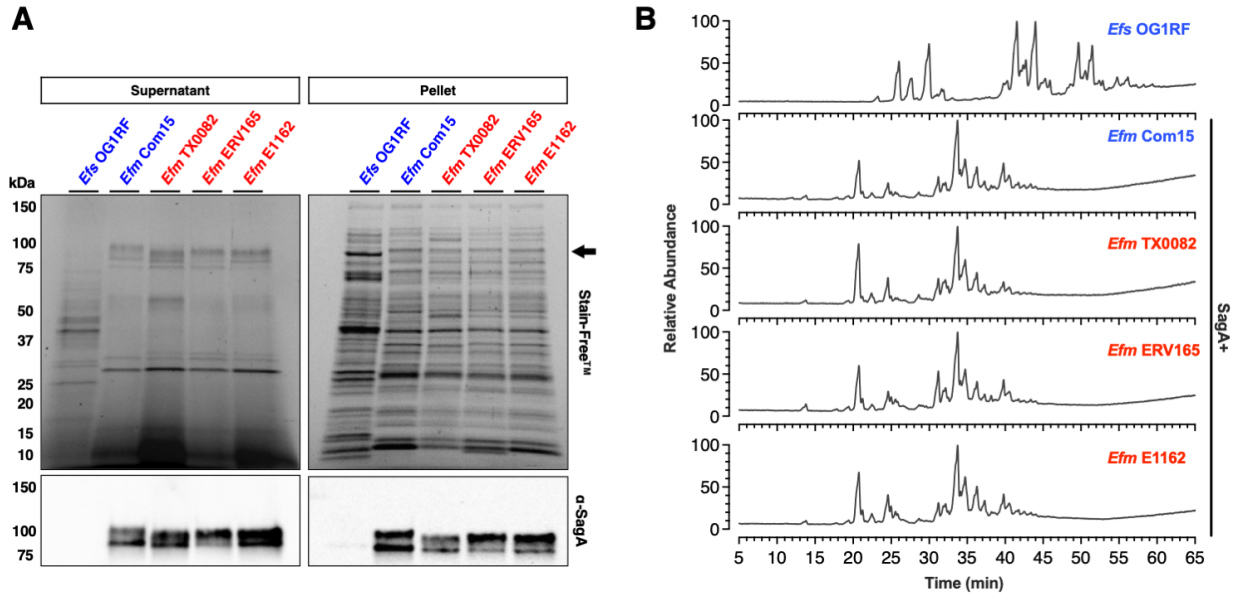


Figure 6.2. MDR *Efm* strains abundantly secrete SagA and share similar PG profiles compared to commensal strains in culture. (A) Expression of SagA in secreted protein (supernatant) and cell pellet (pellet) fractions harvested from midlog phase (OD_{600} 0.6) cultures of the indicated *Enterococcus* species and strains. Top panel (Stain-Free™) represents total protein loading. Bottom panel (α-SagA) represents Western blot detection of SagA using antiserum raised against recombinant *Efm* Com15 SagA. Commensal and MDR *Efm* strains are in blue and red, respectively. Numbers indicate estimated molecular weight (kDa). Arrow denotes migration of SagA band ~75 kDa mark. **(B)** LC-MS ion chromatograms of mutanolysin digested PG harvested from cultures of the indicated *Enterococcus* species and strains in (A). Data are shown as relative intensity of base peak ion abundance.

To determine if the putative NlpC/p60 amidase and CwlT orthologs are functional PG hydrolases, we pursued the purification and *in vitro* PG hydrolase activity of recombinant proteins from the representative *Efm* TX0082 given the high sequence identities for each class of proteins between strains. *Efm* Com15 SagA-NlpC/p60 served as our positive control for PG hydrolase activity (*40*, *42*). The putative NlpC/p60 amidase from *Efm* TX0082, which we will reference henceforth as 501F, did not express in *E. coli* under our typical conditions (*41*). Several 501F truncation constructs were then explored and a construct with truncations to the disordered N- and C-terminal regions (501F ΔN/ΔC) displayed the most robust expression (**Fig. 6.3**). Expression of the 501F NlpC/p60 domain alone was unsuccessful, while the 501F-amidase construct did not show PG hydrolase activity in our *in vitro* assay (data not shown). The cleavage bond specificity of the muramidase (MurNAc-GlcNAc) and NlpC/p60 (γ-D-Glu-mDAP) domains of *Bs* CwlT was previously determined (*38*), but the substrate specificity of the CwlT-NlpC/p60 domain had not been studied thoroughly. Therefore, in addition to full-length *Efm* TX0082 CwlT (referenced hereafter as 51B9 CwlT), we generated the 51B9 CwlT-NlpC/p60 domain alone to dissect the role of each domain in full-length PG hydrolase activity (**Fig. 6.3**). Mutagenesis to alanine of the conserved catalytic cysteine in the NlpC/p60 domains of 501F ΔN/ΔC, 51B9 CwlT, and 51B9 CwlT-NlpC/p60 were produced to discern the precise function/substrate specificity of their NlpC/p60 domains (**Fig. 6.3**).

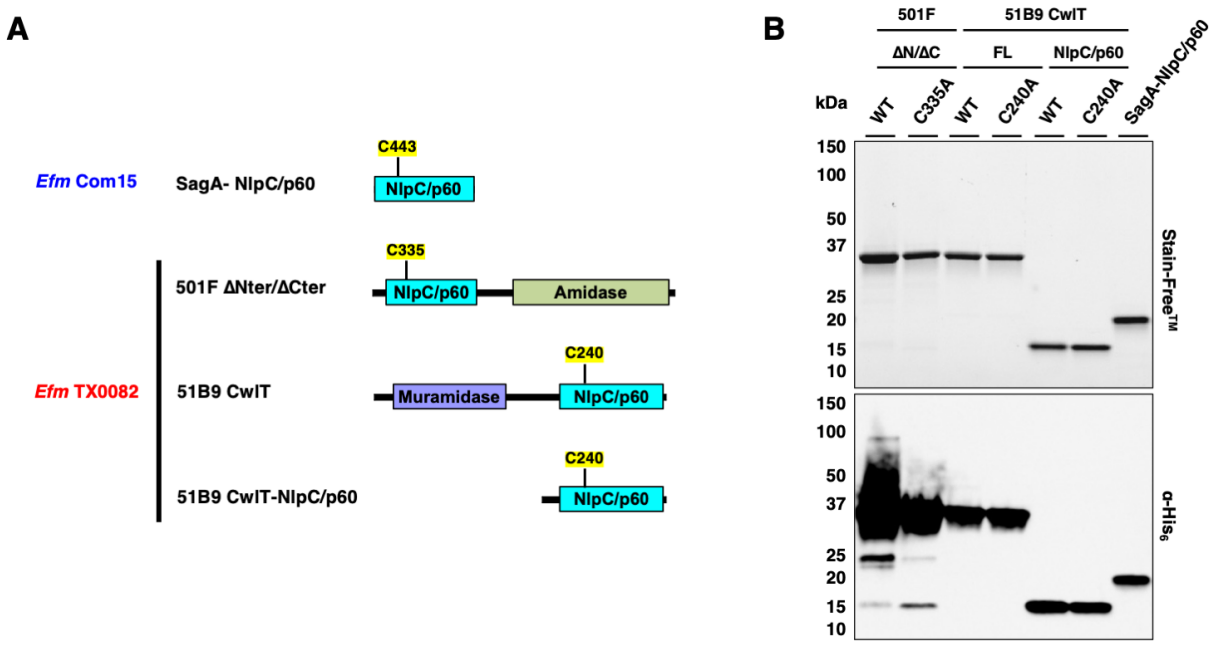


Figure 6.3. Purification of NlpC/p60 hydrolase truncation constructs. (**A**) Domain architecture schematic of NlpC/p60 hydrolase truncation constructs. Catalytic cysteine of NlpC/p60 domain highlighted in yellow. (**B**) Stain-Free (Bio-Rad) SDS-PAGE and α-His_6 immunoblotting of NlpC/p60 hydrolase truncation constructs purified from the cell-free lysate of BL21(DE3)-RIL *Escherichia coli*.

LC-MS analysis of soluble mutanolysin-digested *Efs* OG1RF PG incubated with the NlpC/p60 truncation constructs revealed that they are all D,L-endopeptidases (**Fig. 6.4 and 6.5**). 501F ΔN/ΔC showed similar activity to SagA-NlpC/p60 (**Fig. 6.4 and 6.5**), which was lost with the 501F ΔN/ΔC C335A mutant (data not shown). Whether the 501F-Amidase domain encodes its own PG hydrolase activity that would truly classify 501F as a bifunctional PG hydrolase remains to be determined and may require either different conditions or substrates for *in vitro* activity assays. 51B9 CwlT exhibited more thorough D,L-endopeptidase activity than SagA-NlpC/p60, as it fully hydrolyzed all available muropeptide fragments (**Fig. 6.4 and 6.5**). To elaborate, we have shown that SagA-NlpC/p60 can cleave a monomeric pentapeptide PG fragment and can iteratively cleave cross-linked PG fragments with tetra- and/or pentapeptide stems only (*42*) (**Fig. 6.4**). The resulting D,L-endopeptidase products are the remaining peptide stem backbones, and the peaks for these products were more obvious in the 51B9 CwlT sample (**Fig. 6.4**). Moreover, unlike SagA-NlpC/p60, 51B9 CwlT hydrolyzed PG fragments ranging from monomers to tetramers terminating in tripeptide stems (**Fig. 6.5**), which are a hallmark of mature PG (*100*, *146*). The peptide stem backbone products from the hydrolysis of these substrates were unique to 51B9 CwlT (**Fig. 6.5**). The 51B9 CwlT C240A mutant was inactive (data not shown), indicating that the muramidase domain is not involved in this robust D,L-endopeptidase activity. Accordingly, only wild type 51B9 CwlT-NlpC/p60 (**Fig 6.4 and 6.5**) and not the C240A mutant (data not shown) displayed the same activity as full-length 51B9 CwlT-NlpC/p60, which confirms that the 51B9 CwlT-NlpC/p60 domain alone is sufficient for this broad substrate specificity and complete D,L-endopeptidase activity *in vitro*. *Bs* CwlT-NlpC/p60 PG hydrolase activity was discovered using mDAP-type *Bs* PG, and these results imply that CwlT proteins either have the ability to hydrolyze both mDAP and L-Lys-type PG similar to *M. tuberculosis* (*Mtb*) RipA (*7*, *21*), or that unknown structural differences exist between *Bs* and *Efm* CwlT that control hydrolytic specificity. Our biochemical studies ultimately demonstrate that 501F and 51B9 CwlT are catalytically active PG hydrolases that could serve as targets for functional experiments in commensal (501F) and MDR (501F and 51B9 CwlT) strains of *Efm*.

Figure 6.4. Biochemical characterization of *in vitro* NlpC/p60 D,L-endopeptidase activity towards muropeptides with pentapeptide stems. (**A**) LC-MS analysis of soluble mutanolysin-digested *Efs* OG1RF PG incubated with NlpC/p60 truncation constructs at 37˚C for 16 hr. Data are shown as total ion chromatograms. Colored columns indicate groupings of PG substrates containing pentapeptide stems (S, even numbers) and final cross-linked peptide stem backbone products (P) as in (B-E). Iterative cleavage product GMDP is marked by dashed box. (**B-E**) Schematics of *in vitro* NlpC/p60 D,L-endopeptidase activity. The substrates represent soluble *Efs* OG1RF PG fragments released following mutanolysin digestion of intact *Efs* OG1RF PG. Iterative cleavage of each tetrameric (B), trimeric (C), dimeric (D) and monomeric (E) substrate by NlpC/p60 D,L-endopeptidase activity denoted by scissors generates GlcNAc-MurNAc-dipeptide (GMDP) until only the cross-linked peptide stem backbone remains. All fragments illustrated with reduced anomeric carbon of MurNAc.

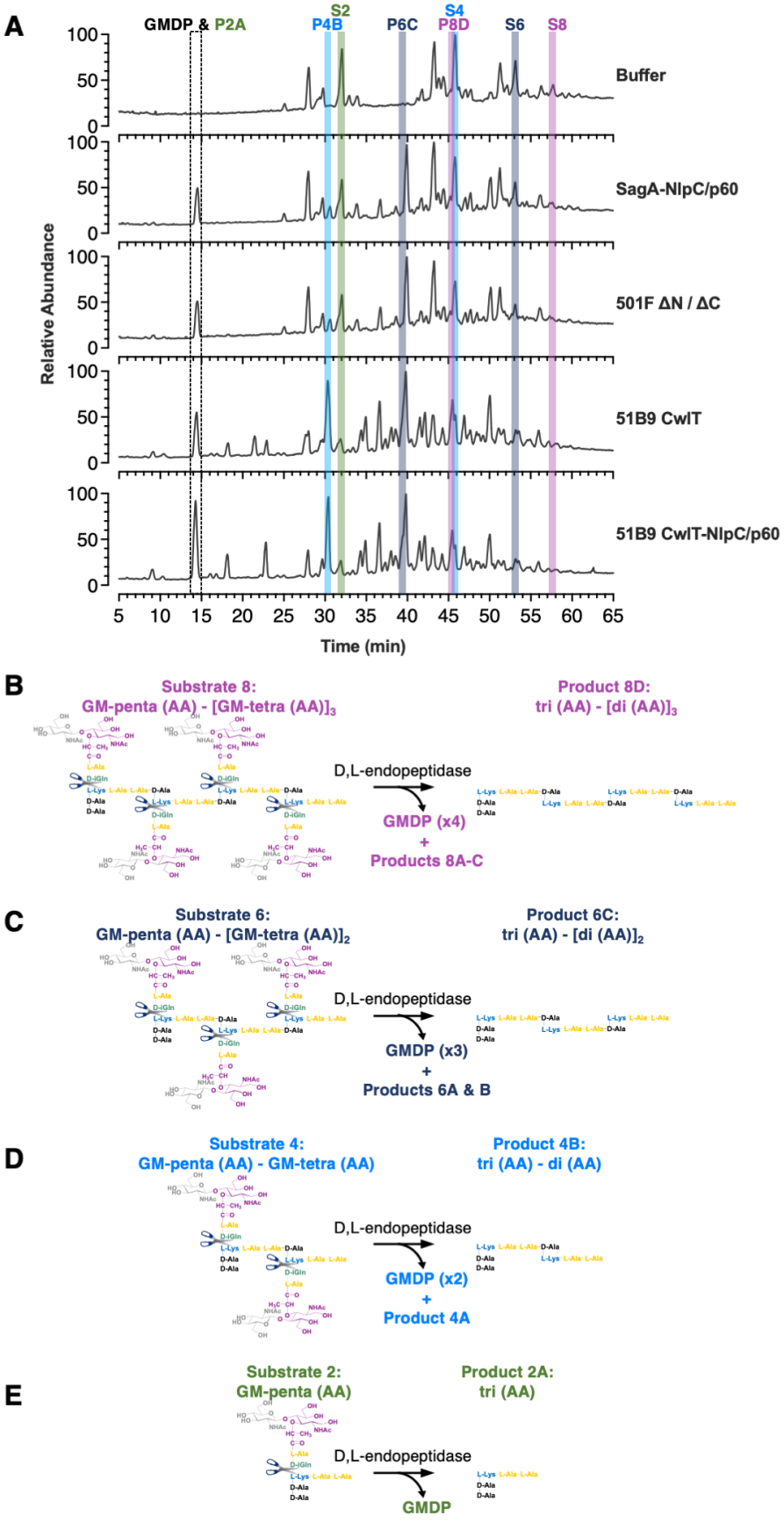
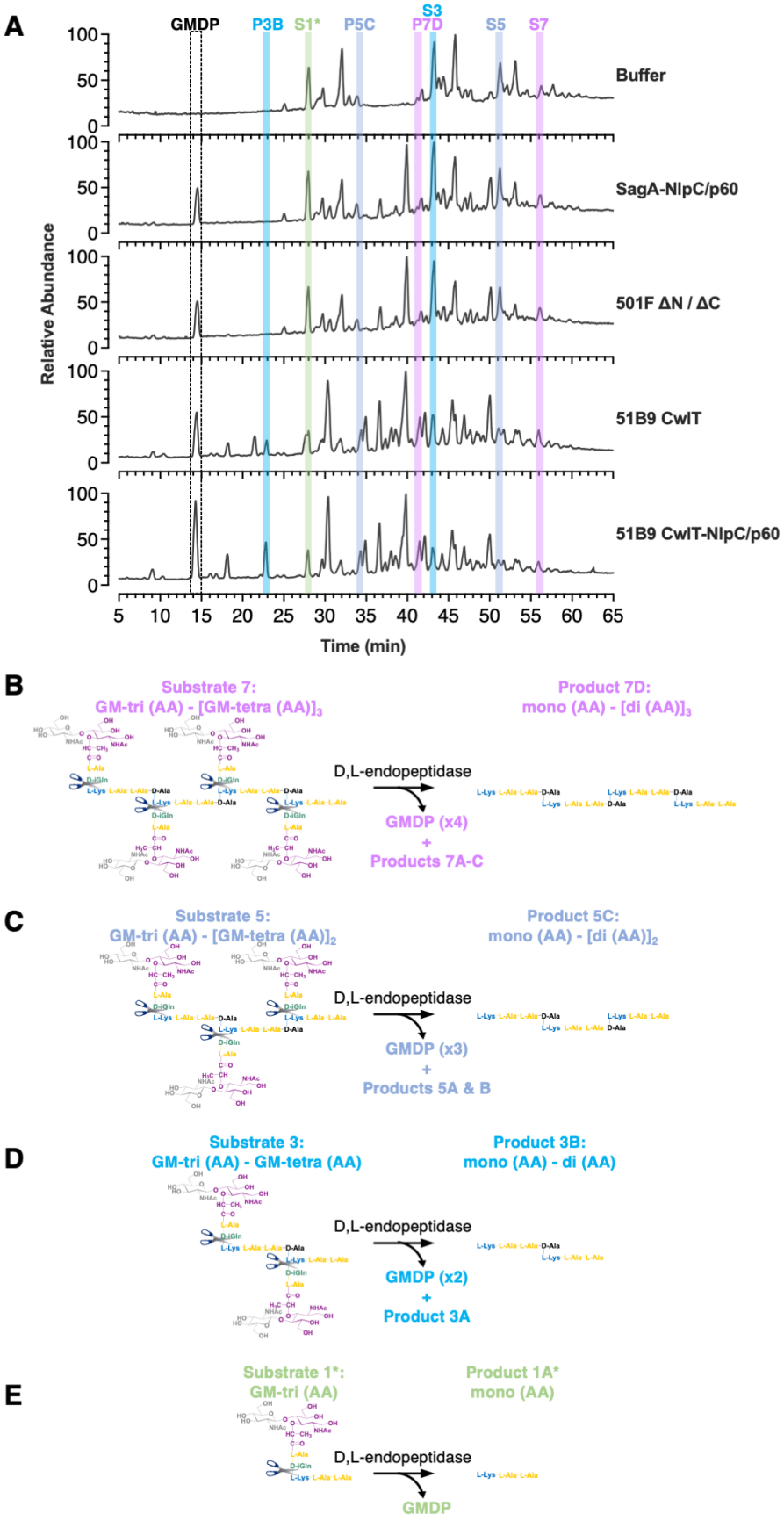

A
GMDP & P2A
S2
P4B
P6C
S4
P8D
S6
S8
Buffer
SagA-NlpC/p60
501F ΔN / ΔC
51B9 CwlT
51B9 CwlT-NlpC/p60
Relative Abundance
Time (min)
B
Substrate 8:
GM-penta (AA) - [GM-tetra (AA)]3
Product 8D:
tri (AA) - [di (AA)]3
D,L-endopeptidase
GMDP (x4)
+
Products 8A-C
C
Substrate 6:
GM-penta (AA) - [GM-tetra (AA)]2
Product 6C:
tri (AA) - [di (AA)]2
D,L-endopeptidase
GMDP (x3)
+
Products 6A & B
D
Substrate 4:
GM-penta (AA) - GM-tetra (AA)
Product 4B:
tri (AA) - di (AA)
D,L-endopeptidase
GMDP (x2)
+
Product 4A
E
Substrate 2:
GM-penta (AA)
Product 2A:
tri (AA)
D,L-endopeptidase
GMDP

Figure 6.5. Biochemical characterization of *in vitro* NlpC/p60 D,L-endopeptidase activity towards muropeptides with tripeptide stems. (**A**) LC-MS analysis of soluble mutanolysin-digested *Efs* OG1RF PG incubated with NlpC/p60 truncation constructs at 37˚C for 16 hr. Data are shown as total ion chromatograms. Colored columns indicate groupings of PG substrates containing tripeptide stems (S, odd numbers) and final cross-linked peptide stem backbone products (P) as in (B-E). Iterative cleavage product GMDP is marked by dashed box. Substrate 1 (GM-tri (AA); panel E, left) marked with star, as the peak abundance decreases in samples incubated with 51B9 CwlT constructs, but Product 1A (mono (AA); panel E, right) cannot be resolved well under our chromatography conditions. (**B-E**) Schematics of *in vitro* NlpC/p60 D,L-endopeptidase activity. The substrates represent soluble *Efs* OG1RF PG fragments released following mutanolysin digestion of intact *Efs* OG1RF PG. Iterative cleavage of each tetrameric (B), trimeric (C), dimeric (D) and monomeric (E) substrate by NlpC/p60 D,L-endopeptidase activity denoted by scissors generates GlcNAc-MurNAc-dipeptide (GMDP) until only the cross-linked peptide stem backbone remains. All fragments illustrated with reduced anomeric carbon of MurNAc.


A
GMDP
P3B
S1*
P5C
S3
P7D
S5
S7
Buffer
SagA-NlpC/p60
501F ΔN / ΔC
51B9 CwlT
51B9 CwlT-NlpC/p60
Relative Abundance
Time (min)
B
Substrate 7:
GM-tri (AA) - [GM-tetra (AA)]3
Product 7D:
mono (AA) - [di (AA)]3
D,L-endopeptidase
GMDP (x4)
+
Products 7A-C
C
Substrate 5:
GM-tri (AA) - [GM-tetra (AA)]2
Product 5C:
mono (AA) - [di (AA)]2
D,L-endopeptidase
GMDP (x3)
+
Products 5A & B
D
Substrate 3:
GM-tri (AA) - GM-tetra (AA)
Product 3B:
mono (AA) - di (AA)
D,L-endopeptidase
GMDP (x2)
+
Product 3A
E
Substrate 1*:
GM-tri (AA)
Product 1A*
mono (AA)
D,L-endopeptidase
GMDP

In conclusion, our preliminary investigation into other NlpC/p60 PG hydrolases in addition to SagA as potential MDR *Efm* targets have yielded two candidates: the putative NlpC/p60 amidase, 501F, conserved in commensal and MDR *Efm* strains, and 51B9 CwlT unique to MDR *Efm*. Protein expression and cell wall profiling of MDR *Efm* were inconclusive in discovering whether the strains have altered profiles compared to commensal *Efm* Com15 that would be indicative of differential PG hydrolase activity. Future studies on this front should include MDR *Efm* grown in the presence of antibiotics to see if unique phenotypes arise. To validate 501F and 51B9 CwlT as catalytically active PG hydrolases that could be targeted in functional studies of MDR *Efm* strains, we purified recombinant truncation constructs and evaluated their PG hydrolase activity *in vitro*. A 501F ΔN/ΔC construct, full-length 51B9 CwlT, and the 51B9 CwlT-NlpC/p60 domain alone demonstrated D,L-endopeptidase activity towards L-Lys-type *Efs* PG *in vitro* dependent on the conserved catalytic cysteine of NlpC/p60 PG hydrolases. Interestingly, 51B9 CwlT was able to fully hydrolyze all muropeptide substrates under our conditions, including PG fragments with tripeptide stems originating from mature PG. The 51B9 CwlT-NlpC/p60 domain alone was sufficient for this activity, raising two immediate possibilities: (1) CwlT orthologs may have broad substrate specificity (i.e., tripeptide and pentapeptide stems) in addition to broad hydrolytic specificity (i.e., mDAP and L-Lys-type PG) similar to *Mtb* RipA, or (2) as of yet undetermined structural feature(s) account for differential activity in CwlT orthologs of different bacteria. Follow up biochemical studies using mDAP-type PG substrates would help distinguish these possibilities, as would the structural determination of 51B9 CwlT-NlpC/p60. Together, our results have set a foundation for the assessment of these catalytically active NlpC/p60 hydrolases in MDR *Efm* function and virulence.

Chapter 7. Summary and Future Outlook.

PG is a vital component of the bacterial cell wall, and its dynamic remodeling by NlpC/p60 hydrolases is crucial for proper cell division and survival among other functions, as described in Chapter 1. Beyond these essential functions, we previously discovered that *Enterococcus* species express and secrete the NlpC/p60 hydrolase SagA, whose catalytic activity can modulate host immune responses in animal models (*68*, *69*). In this thesis, we describe biochemical methods to investigate the involvement of SagA NlpC/p60 hydrolase activity in limiting bacterial pathogenesis, remodeling of *Enterococcus* PG, improving efficacy of cancer immunotherapy, and treatment of pathogenic enterococci.

In Chapter 2, we discuss methods for the biochemical analysis of NlpC/p60 hydrolases. These methodologies include bacterial expression and purification of recombinant NlpC/p60 proteins, large scale isolation of peptidoglycan, and *in vitro* activity assays of NlpC/p60 hydrolase activity (*41*). We applied these optimized methods to study *Efm* Com15 SagA, which demonstrated the ability to cleave L-Lys type PG with varying cross bridges and generate GMDP *in vitro* (*40*). Molecular docking studies using the SagA-NlpC/p60 domain structure revealed conserved Trp residues, Trp433 and Trp462, that may be involved in binding of PG and whose importance for SagA-NlpC/p60 PG hydrolase activity we corroborated using our enzymatic assays (*40*). PG products such as GMDP were isolated from enzymatic reactions with SagA and activated the pattern recognition receptor NOD2 in mammalian cells *ex vivo*, thus directly linking SagA-generated muropeptides in activating host immune pathways (*40*). The probiotic *L. plantarum* was engineered to heterologously express SagA constructs and provided evidence that secretion of catalytically active SagA is required for protection *in vivo* (*40*). Therefore, these results greatly improved our understanding of the role SagA NlpC/p60 hydrolase activity may play in the SagA-mediated protection against enteric pathogens (*40*). In addition, our results served as a reference for subsequent biochemical studies on SagA NlpC/p60 hydrolase activity in *Enterococcus* PG remodeling during cell division.

In Chapter 3, we describe imaging studies of enterococci and SagA as well as additional structure activity studies. Despite lacking loss-of-function studies, the role of SagA in cell division has been suggested due to the essentiality of the *sagA* gene for cell growth (*99*). Imaging studies of SagA+ enterococci and *Efm* Com15 overexpressing SagA and its truncation constructs revealed that SagA contributes to a triseptal structure in growing cells and colocalizes with these active areas of PG turnover via its coiled-coil N-terminus (*42*). These results suggest SagA is involved in PG remodeling during cell division and may contribute to simultaneous cell division and separation (*42*). However, whether the NlpC/p60 domain of SagA plays a structural or catalytic role in this triseptal phenotype remains to be determined (*42*). Further biochemical studies determined that the coiled-coil N-terminus of SagA does not have an autoinhibitory effect observed in the similar PG hydrolase PcsB from *Streptococcus pneumoniae* (*89*); full-length SagA and SagA-NlpC/p60 exhibited comparable D,L-endopeptidase activity towards predigested, heterogenous *Efs* OG1RF PG fragments *in vitro* (*42*). Subsequent *in vitro* and *in silico* alanine scanning of the putative *Efm* Com15 SagA-NlpC/p60 substrate-binding groove revealed additional key residues that may be involved in PG turnover (*42*). Surprisingly, these studies revealed that SagA-NlpC/p60 D,L-endopeptidase activity may occur via a catalytic Cys-His dyad (*42*) as opposed to the putative catalytic Cys-His-His triad in homologous NlpC/p60 hydrolases (*40*). SagA orthologs from other commensal *Enterococcus* species have homologous NlpC/p60 domains to *Efm* Com15 SagA, and Cys-His-Ala mutants of these full-length SagA orthologs were active,

which suggests that the second histidine is not required for the *in vitro* activity of Saga orthologs from enterococci (*42*). Comparative structural analysis of SagA-NlpC/p60 and other biochemically and structurally characterized NlpC/p60-type hydrolases illustrated the challenge of elucidating the exact structural motifs and substrate-protein interactions that govern their diverse enzymatic and hydrolytic specificities. Ultimately, our results highlight how the activity of SagA may be involved in the viability of *Efm* and provides a greater understanding of how SagA generates the defined muropeptides that we have shown to have therapeutic potential (*40*, *68*, *69*).

In Chapter 4, we focus on the role SagA NlpC/p60 hydrolase activity plays in improving efficacy of cancer immunotherapy by checkpoint blockade. Interestingly, SagA is expressed and secreted by other commensal enterococci (*43*) that have been recovered from immunotherapy-responsive cancer patients, such as *Eds*, *Ehe*, and *Emi* (*117*). These SagA+ enterococci improved the efficacy of the immune-checkpoint inhibitor anti–PD-L1 in a mouse tumor model (*43*). Moreover, the PG architecture of these strains was conserved, as was the *in vitro* D,L-endopeptidase activity of the respective SagA orthologs (*43*). Heterologous expression of *sagA* in the nonprotective *Efs* was sufficient to improve the antitumor activity of anti–PD-L1, anti–PD-1, and anti–CTLA-4 via an adaptive immune response in mice, and of anti–PD-1 in a Nod2-dependant manner (*43*). Since *sagA* is proposed to be essential in *Efm* (*99*), and to avoid concerns regarding the acquisition of antibiotic resistance in *Efm* (*120–122*), *Lls* strains with chromosomal *sagA* insertions were used to determine that secretion of catalytically active SagA (i.e., no C443A mutation in the NlpC/p60 domain) is necessary for the enhancement of immunotherapy activity (*43*). Similar to the SagA D,L-endopeptidase product GMDP, the Nod2-active L,D-isomer of MDP significantly improved the antitumoral activity of anti–PDL-L1 (*43*) further supporting that SagA enzymatic activity is required to improve response to immunotherapy. Collectively, these results highlight the therapeutic potential of the PG hydrolase activity of SagA and suggest that other secreted NlpC/p60 hydrolases may deployed in a similar fashion in commensal bacteria to treat bacterial infection or cancer.

In Chapter 5, we describe our collaboration with the Duerkop laboratory to study the mechanisms underlying lytic bacteriophage infection of *Efm* as a therapeutic alternative to current antibiotics. To that end, the Duerkop lab generated phage resistant mutant strains of commensal *Efm* Com12 and identified mutations mapped to the NlpC/p60 domain of *sagA* among other cell wall-related genes (*135*). These *sagA* mutant strains abundantly secreted SagA and did not appear to have any growth defects in culture despite two of the mutations mapping to Trp433 (W433L and W433C) (*135*), which was previously determined to be required for SagA-NlpC/p60 hydrolase activity *in vitro* (*40*). Predictably, these mutations led to fitness trade-offs; mutation in *sagA* conferred resistance to a specific phage, 9181, but at the expense of enhanced susceptibility to PG-targeting antibiotics (*135*). Moreover, combining phage 9181 with β-lactam and lipopeptide antimicrobials had a synergistic effect in inhibiting *Efm* growth (*135*). Subsequent biochemical studies revealed that although the cell wall profiles of the phage 9181 resistant mutant strains were similar to parental *Efm* Com12, the SagA-NlpC/p60 mutants corresponding to each mutant strain exhibited abrogated D,L-endopeptidase activity *in vitro*. This observation implies that either the phage 9181 resistant *Efm* Com12 strains retain basal SagA NlpC/p60 hydrolase activity that is below the limit of detection in our *in vitro* assay, or that SagA PG hydrolase activity itself is dispensable for viability entirely. Ultimately, these results suggest that targeting the NlpC/p60 domain of SagA at the genome (and perhaps protein) level may sensitize pathogenic, multi-drug resistant strains of *E. faecium* to currently available antimicrobials (*135*).

In Chapter 6, we discuss our exploration into other NlpC/p60 hydrolases as potential MDR *Efm* targets for functional studies and therapeutic development. *Efm* SagA is involved in PG remodeling during cell division of commensal enterococci (*42*), and our recent collaboration with the Duerkop laboratory positions *sagA* as a potential target for the treatment of *Efm* infection (*135*). In this regard, comparative genomics of NlpC/p60 proteins in commensal and clinical isolate, MDR *Efm* strains identified putative NlpC/p60 amidases and CwlT orthologs as candidate targets. Protein expression and cell wall profiling of three readily-available MDR *Efm* strains was similar to commensal *Efm* Com15 under normal growth conditions, implying that future studies may have to include antibiotics to ascertain whether these strains display different profiles when their resistance genes are activated. Our optimized biochemical methods were then employed to validate if proteins from each of the identified clusters are catalytically active PG hydrolases that could be targeted in functional studies. Using *Efm* TX0082 as a representative MDR strain, we purified recombinant constructs of a putative NlpC/p60 amidase (501F) and CwlT ortholog (51B9 CwlT). A 501F ΔN/ΔC construct, full-length 51B9 CwlT, and the 51B9 CwlT-NlpC/p60 domain alone demonstrated D,L-endopeptidase activity towards L-Lys-type *Efs* PG *in vitro* dependent on the conserved catalytic cysteine of NlpC/p60 PG hydrolases. Interestingly, 51B9 CwlT and 51B9 CwlT-NlpC/p60 were able to fully hydrolyze all muropeptide substrates under our conditions, including PG fragments with tripeptide stems originating from mature PG. The structural feature(s) that endow 51B9 CwlT-NlpC/p60 with this robust PG hydrolase activity are of great interest and will be the focus of future studies. Our preliminary results have set a foundation for the assessment of these catalytically active NlpC/p60 hydrolases in MDR *Efm* function and virulence.

Collectively, our results underscore the utility of our methods in studying NlpC/p60 hydrolase activity, and its roles in modulating host immunity. Several challenges remain, however, towards the widespread application of NlpC/p60 hydrolase activity in therapies for human disease. Being that NlpC/p60 hydrolases are ubiquitous in bacteria (*1*), it is imperative that more high throughput methods are established to identify catalytically active NlpC/p60 hydrolases for further biochemical and functional studies. In-gel profiling of ANTS-labeled PG fragments has proven valuable for evaluating the NlpC/p60 hydrolase activity of around a dozen SagA-NlpC/p60 constructs at a time (*41*, *42*), but its low resolution makes it more useful as an intermediary screen of mutants after prior candidate validation. LC-MS analysis of PG fragments has also been instrumental towards characterizing the substrate and cleavage bond specificities of NlpC/p60 hydrolases (*38*, *40–42*, *56*), but it is not time and cost-efficient. Moreover, as in the case of SagA-NlpC/p60 mutants corresponding to the phage 9181 resistant *Efm* Com12 strains (*135*), we cannot discard the possibility that these assays are not sensitive enough to definitively prove if mutations at sites other than the catalytic cysteine renders NlpC/p60 hydrolases completely inactive.

Several NlpC/p60 hydrolases have shown activity towards intact PG or even cells in spectrophotometric assays (*49*, *50*, *56*), which can facilitate biochemical characterization of many proteins in parallel (*49*). However, for proteins like SagA that can only cleave predigested, soluble muropeptides *in vitro* (*40*), alternative PG substrates should considered. Something that can serve as a launching point for the development of new assays for NlpC/p60 hydrolase activity is the fact that we can produce large stocks of pure, heterogenous mixtures of PG fragments (*41*). To that end, our laboratory has recently explored derivatization of crosslinked PG fragments with self-quenching dyes, whose fluorescence may be activated and quantified in fluorescence polarization assays upon hydrolysis and product release from the NlpC/p60 substrate-binding groove. One can envision that these dyes can be linked to the free anomeric carbon of the MurNAc moiety of soluble cross-linked PG fragments This common feature in PG fragments would also broaden the

versatility of these synthetic or semisynthetic fragments, as the derivatization could be applied to PG from a variety of bacteria. Moreover, the peptide stems would be unaltered under this scheme, thus allowing for multiple types of NlpC/p60 hydrolases (i.e., amidases, endopeptidases, and carboxypeptidases) to be evaluated.

The next frontier for NlpC/p60 hydrolases is as targets in MDR *Enterococcus* for functional studies and the downstream development of inhibitors. Unlike the elucidation of RipA function in *Mtb* cell division and virulence (*23–27*), functional studies of NlpC/p60 hydrolases in commensal and pathogenic *Efm* have been hindered by limited gene editing technologies. However, recent developments from our laboratory on this front have afforded new avenues of investigation regarding NlpC/p60 hydrolase activity in *Efm* viability and pathogenicity. RecT-mediated recombineering in *Efm* should enable insertions, deletions and substitutions of genomic DNA in *Efm* (*139*). These gene editing tools will allow us to target *sagA* in *Efm* strains to introduce mutations, knockdown expression, or delete the gene altogether. We could then identify and exploit canonical impaired cell division phenotypes (e.g., altered growth, protein expression, and cell wall composition, and susceptibility to antibacterial agents) as they have for RipA (*23*). Similarly, now that we have identified 501F and 51B9 CwlT as catalytically active NlpC/p60 hydrolases, targeted deletion and knockdown of the conserved genes will allow us to ascertain the function(s) of these proteins in their respective *Efm* strains. 51B9 CwlT is of particular interest given its unique presence in MDR *Efm*, robust *in vitro* D,L-endopeptidase activity, and possible role in bacterial conjugation (*55*). Once we validate these proteins as MDR *Efm* targets, we can develop inhibitors of the conserved catalytic cysteine in their NlpC/p60 domains. Co-crystallization studies using these proteins with PG substrates or related small molecules could then reveal key ligand-protein interactions that can inform rational design of potent inhibitors of NlpC/p60 hydrolase activity. Together, our methods for the biochemical analysis of NlpC/p60 hydrolases and gene editing tools in *Efm* could establish a new workflow for unraveling MDR *Efm* pathogenicity.

Materials and Methods

Cloning and transformations

All cloning was done in chemically competent *E. coli* Dh5α (New England Biolabs, NEB) according to the manufacturer's protocol (pAM401 plasmids = 25 μg mL^{-1} chloramphenicol; pET-21a(+) plasmids = 100 μg mL^{-1} ampicillin). Colonies were picked and verified by Sanger sequencing (Genewiz). Confirmed pAM401 plasmids were transformed into electrocompetent *Efm* Com15 (**Appendix Table 1**) according to a Palmer laboratory *Enterococci* transformation protocol (*147*) for constitutive expression (10 μg mL^{-1} chloramphenicol). Confirmed pET-21a(+) plasmids were transformed into competent *E. coli* BL21-CodonPlus (DE3)-RIL (Agilent) according to the manufacturer's protocol for IPTG inducible expression (100 μg mL^{-1} ampicillin *and* 25 μg mL^{-1} chloramphenicol; see note in **Appendix Table 1**).

Plasmid constructions for protein expression

Chapter 2

Site-directed mutagenesis of full-length SagA and SagA-NlpC/p60 constructs in the pET-21a(+) vector backbone (Millipore Sigma) was performed using Q5 Hot-Start High-Fidelity 2x master mix (NEB) according to the manufacturer's protocol. To generate the C443A mutant of *Efm* Com15 SagA and SagA-NlpC/p60 ΔSS, *Efm*_Com15_SagA_-His_6-pET-21a(+) and *Efm*_Com15_SagA-NlpC/p60_ΔSS-His_6-pET-21a(+) (*40*) were used as templates, respectively (**Appendix Table 1**), along with primers 1 and 2 (**Appendix Table 2**). *Efm*_Com15_SagA-NlpC/p60_ΔSS-His_6-pET-21a(+) served as the template along with the respective primer pairs to generate the single W433A (3 and 4) and W462A (5 and 6) mutants. *Efm*_Com15_SagA-NlpC/p60_ΔSS_W433A-His_6-pET-21a(+) was then used along with primers 5 and 6 to generate the double W433A/W462A mutant. Transformation into *E. coli* Dh5α and subsequent transformation into *E. coli* BL21-CodonPlus (DE3)-RIL performed as described above.

Chapter 3

Site-directed mutagenesis of *Efm* Com15 SagA constructs in the pET-21a(+) vector backbone (Millipore Sigma) was performed using Q5 Hot-Start High-Fidelity 2x master mix (NEB) according to the manufacturer's protocol. To generate individual alanine mutants of *Efm* Com15 SagA-NlpC/p60, *Efm*_Com15_SagA-NlpC/p60_ΔSS-His_6-pET-21a(+)(*40*) was used as a template (**Appendix Table 1**) along with the appropriate primer pairs from primers 13-42 (**Appendix Table 2**). To generate signal sequence deletion mutants of full-length *Efm* Com15 SagA WT and C443A, *Efm*_Com15_SagA-His_6-pET-21a(+) (*69*) and *Efm*_Com15_SagA_C443A-His_6-pET-21a(+)(*40*) were used as templates, respectively, along with primers 43 and 44. To generate the full-length *Efm* Com15 SagA ΔSS H506A mutant, *Efm*_Com15_SagA_ΔSS-His_6-pET-21a(+) was used as a template along with primers 41 and 42. Transformation into *E. coli* Dh5α and subsequent transformation into *E. coli* BL21-CodonPlus (DE3)-RIL performed as described above.

To clone signal sequence deletion mutants of full-length SagA orthologs, genomic DNA from *E. durans* 23C2 and *E. mundtii* NCDO 2375 was purified from 5 mL overnight cultures using the E.Z.N.A. Bacterial DNA Kit (Omega Bio-Tek) according to the manufacturer's protocol. *E. durans* 23C2 *sagA* (NCBI, WP_016177750.1; primers 45 and 46) and *E. mundtii* NCDO 2375

sagA (NCBI, WP_023520500.1; 47 and 48) were amplified as signal sequence deletion mutants from the respective genomic DNA samples using Q5 Hot-Start High-Fidelity 2x master mix according to the manufacturer's protocol. Sequences were inserted into the pET-21a(+) plasmid that was double digested with NdeI (NEB) and XhoI (NEB) using the NEBuilder HiFi DNA Assembly Cloning Kit (NEB). Transformation into *E. coli* Dh5α and subsequent transformation into *E. coli* BL21-CodonPlus (DE3)-RIL performed as described above.

Chapter 4

To clone signal sequence deletion mutants of full-length SagA orthologs and related proteins, genomic DNA from *E. durans* 23C2 , *E. hirae* R, *E. mundtii* NCDO 2375, and *E. faecalis* OG1RF was purified from 5 mL overnight cultures using the E.Z.N.A. Bacterial DNA Kit (Omega Bio-Tek) according to the manufacturer's protocol. *E. durans* 23C2 *sagA* (primers 45 and 46 in **Appendix Table 2**), *E. hirae* R *sagA* (53 and 54), *E. mundtii* NCDO 2375 *sagA* (47 and 48), and *E. faecalis* OG1RF *salA* (55 and 56) and *salB* (57 and 58) were amplified as signal sequence deletion mutants from the respective genomic DNA samples using Q5 Hot-Start High-Fidelity 2x master mix according to the manufacturer's protocol. Sequences were inserted into the pET-21a(+) plasmid that was double digested with NdeI (NEB) and XhoI (NEB) using the NEBuilder HiFi DNA Assembly Cloning Kit (NEB). Transformation into *E. coli* Dh5α and subsequent transformation into *E. coli* BL21-CodonPlus (DE3)-RIL performed as described above.

Chapter 5

Site-directed mutagenesis of phage 9181 resistant *Efm* Com12 SagA-NlpC/p60 mutants in the pET-21a(+) vector backbone (Millipore Sigma) was performed using Q5 Hot-Start High-Fidelity 2x master mix (NEB) according to the manufacturer's protocol. To generate mutants corresponding to each phage 9181 resistant *Efm* Com12 strain, *Efm*_Com15_SagA-NlpC/p60_ΔSS-His_6-pET-21a(+) (*40*) was used as a template (**Appendix Table 1**) along with the appropriate primer pairs from primers 59-68 (**Appendix Table 2**). Transformation into *E. coli* Dh5α and subsequent transformation into *E. coli* BL21-CodonPlus (DE3)-RIL performed as described above.

Chapter 6

To generate signal sequence deletion mutants *E. faecium* TX0082 501F and 51B9 CwlT constructs, genomic DNA from *E. faecium* was purified from 5 mL overnight cultures using the E.Z.N.A. Bacterial DNA Kit (Omega Bio-Tek) according to the manufacturer's protocol. The genes encoding *E. faecium* TX0082 501F (primers 69 and 70 in **Appendix Table 2**) and 51B9 CwlT (71 and 72) were amplified as signal sequence deletion mutants from the *E. faecium* TX0082 genomic DNA using Q5 Hot-Start High-Fidelity 2x master mix according to the manufacturer's protocol. Sequences were inserted into the pET-21a(+) plasmid that was double digested with NdeI (NEB) and XhoI (NEB) using the NEBuilder HiFi DNA Assembly Cloning Kit (NEB). The truncation constructs and cysteine to alanine mutants described below were prepared using Q5 Hot-Start High-Fidelity 2x master mix (NEB) according to the manufacturer's protocol. Transformations into *E. coli* Dh5α and subsequent transformations into *E. coli* BL21-CodonPlus (DE3)-RIL performed as described above.

To generate the *E. faecium* TX0082 501F ΔN/ΔC ΔSS construct, the 501F N-terminus was first deleted from the *Efm*_TX0082_501F_ΔSS-His_6-pET-21a(+) plasmid using primers 73 and 74 to give *Efm*_TX0082_501F_ΔNter_ΔSS-His_6-pET-21a(+). The 501F C-terminus was deleted from

this plasmid using primers 75 and 76 to finally give the *Efm*_TX0082_501F_ΔNter_ΔCter_ΔSS-His_6 plasmid. Primers 79 and 80 were then used to generate the inactive C335A mutant.

To generate the *E. faecium* TX0082 51B9 CwlT ΔSS construct, the 51B9 CwlT muramidase domain was deleted from *Efm*_TX0082_51B9_CwlT_ΔSS-His_6 using primers 77 and 78 to give *Efm*_TX0082_51B9_CwlT-NlpC/p60_ΔSS-His_6. Primers 81 and 82 were used to generate the C240A mutant in both of these plasmids.

Plasmid constructions for fluorescent microscopy

The Tetracysteine (Cys_4) Tag (*104*, *105*) was inserted into pAM401 plasmids containing *Efm* Com15 *sagA* (NCBI, WP_127808574.1) and its variants (*69*) by using the QuikChange XL site-directed mutagenesis kit (Stratagene) according to the manufacturer's protocol. To create *SagA-Cys_4-His_6*-pAM401, *SagA-His_6*-pAM401 was used as the template (**Appendix Table 1**) along with primers 7 and 8 (**Appendix Table 2**). To create $SagA^{\Delta C}$-*Cys_4-His_6*-pAM401, $SagA^{\Delta C}$-*His_6*-pAM401 was used as the template along with primers 9 and 10. To create $SagA^{Nter}$-*Cys_4-His_6*-pAM401, $SagA^{Nter}$-*His_6*-pAM401 was used as the template along with primers 11 and 12. Transformation into *E. coli* Dh5α and subsequent transformation into *Efm* Com15 performed as described above.

Overexpression and purification of SagA constructs

For *E. coli* BL21-RIL (DE3), 250 mL – 1 L LB cultures were inoculated with overnight cultures with appropriate antibiotics, grown for 2 hr or until OD_{600} ~0.5, induced with 1 mM isopropyl-*D*-thiogalactopyranoside (IPTG), then grown for an additional 2 hr. Cells were collected and resuspended in 20 ml of lysis buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% SDS, 0.025 U/ml benzonase, and 1x protease inhibitor cocktail). After 15 min of sonication followed by centrifugation at 30,000 x *g* for 30 min, the supernatant containing the soluble target protein was collected and loaded onto a His60 Ni Superflow resin (Takara Bio) equilibrated with the binding buffer (PBS buffer) for a 1 hr incubation. The target protein was eluted with 300 mM imidazole. Semi-purified protein was dialyzed into PBS buffer at 4°C overnight using 10K MWCO Slide-A-Lyzer MINI dialysis devices (Thermo Fisher Scientific). Protein was further purified by loading onto a Superdex 200 (for proteins > 30 kDa) or Superdex 75 (proteins < 30 kDa) Increase 10/300 GL column (GE Healthcare) pre-equilibrated with PBS. Fractions containing the target protein were combined and concentrated. Protein concentration was estimated by BCA assay (Pierce Protein Biology) and protein was stored at −80°C in PBS buffer.

For *E. coli* BL21-RIL (DE3) encoding full length SagA-His_6 constructs, 300 mL LB culture media was collected after 2 hr of IPTG-induced expression and secretion of the target protein. The culture supernatant containing the secreted target protein was vacuum filtered through 0.2 μm membrane and then added to Jumbosep centrifugal devices with 10K MWCO membrane inserts (Pall Life Sciences) for fivefold concentration and buffer exchange to PBS. The protein concentrate was collected and loaded onto a His60 Ni Superflow resin pre-equilibrated with PBS for overnight incubation and subsequent affinity chromatography and dialysis as described above. The protein was further purified by loading onto a Superdex 200 Increase 10/300 GL column (GE Healthcare) pre-equilibrated with PBS. Fractions containing the target protein were combined and concentrated to 2 mg/ml for *in vitro* activity assays. Protein concentration estimation and protein storage were performed as described above.

Western blot analysis of bacterial supernatants and pellets, and purified proteins

Western blot analysis of bacterial supernatants and pellets, and purified proteins performed as described previously (*40*, *41*). Briefly, proteins were separated by SDS-PAGE on 4–20% Criterion TGX precast gels (Bio-Rad), then transferred to nitrocellulose membrane (0.2 μM, BioTrace NT Nitrocellulose Transfer Membranes, Pall Laboratory). HRP conjugated polyclonal anti-His_6 (abcam, ab1187) and polyclonal SagA serum as a primary antibody/HRP conjugated anti-Rabbit IgG (GE Healthcare, NA 934V) as a secondary antibody were used for His_6 and SagA blots, respectively. Polyclonal SagA primary antibodies were used at a 1:50,000 dilution and secondary antibody at 1:10,000, unless otherwise stated. Membranes were blocked for 1 hr in 1% milk, incubated with primary antibody in 1% milk for 1 hr, washed 5 × with TBS-T (Tris-buffered saline, 0.1% Tween 20), incubated with secondary antibody, and washed with 4 × TBS-T. Protein detection was performed with ECL detection reagent (GE Healthcare) on a Bio-Rad ChemiDoc MP Imaging System.

Purification and mutanolysin-digestion of PG

Enterococcus were grown in fresh BHI medium with shaking at 37°C to log-phase (OD_{600} of 0.6). Engineered *Lactococcus lactis* strains were grown in fresh M17 medium supplemented with 20 μg mL-1 thymidine without antibiotics and without shaking at 30°C to log-phase (OD_{600} of 0.6). PG was extracted by resuspending the bacterial cell pellet in 0.25% SDS solution in 0.1 M Tris-HCl, pH 6.8 and boiling the suspension for 20 min at 100°C in a heating block as previously described (*96*). The resulting insoluble cell wall preparation was washed with distilled water six times until free of SDS. The cell wall was purified by treatment with benzonase followed by trypsin digestion. Then, insoluble cell wall was recovered by centrifugation (16,000 x g, 10 min, 4°C), and washed once in distilled water. To obtain pure PG, cell wall was then suspended in 1 M HCl and incubated for 4 hr at 37°C in a shaker to remove wall teichoic acid. The insoluble material was collected by centrifugation (16,000 x g, 10 min, 4°C) and washed with distilled water repeatedly until the pH was 5–6. The final PG was lyophilized and stored at −20°C. For muropeptide analysis, 2 mg purified PG was digested with 1 mg/mL mutanolysin from *Streptomyces globisporus* (Sigma, 10 KU/ml of mutanolysin in ddH_2O) in 10 mM sodium phosphate buffer, pH 4.9 for 16 hr at 37°C. The enzyme reaction was stopped by incubating at 95°C for 5 min. The resulting soluble muropeptide mixture was then analyzed by LC-MS for PG profiling or used as substrate for *in vitro* activity assays, as described below.

***In vitro* assays for NlpC/p60 PG hydrolase activity**

To characterize NlpC/p60 PG hydrolase activity, 100-200 μg of mutanolysin-digested PG was incubated with 10 μM of purified protein in 50 mM Bis-tris, pH 5.5 for 16 hr at 37°C. Reactions were quenched by boiling at 95˚C for 5 min and centrifuged at 16,000 x g for 5 min at room temperature to remove precipitated protein. Supernatants containing soluble muropeptides were then isolated and lyophilized before processing for ANTS labeling or LC-MS analysis.

ANTS-labeling for in-gel fluorescence profiling of PG fragments

ANTS labeling was performed as described (*97*). 10 μl of ANTS reaction mix was added to each tube of dried material (1:1 mixture of 0.2 M ANTS (in 3:17 acetic acid:water): 1 M $NaCNBH_3$ (in DMSO)). Reactions were incubated overnight at 37°C. 0.5–3.5 μL of the ANTS labeled mixtures were mixed 1:1 with 50% glycerol and samples were separated by native PAGE on a hand-cast 37–40% Tris-glycine acrylamide gel (19:1 polyacrylamide:bisacrylamide, with a

20% acrylamide stack) at 100 V for ~4 hr. ANTS-labeled synthetic fragments MDP, GlcNAc, MurNAc, and MurNAc-L-Ala were run for comparison. A sugar-less pentapeptide Ala-D-γ-Glu-Lys-D-Ala-D-Ala (Sigma) was run to show specificity of the UV signal and empty lanes adjacent to sample lanes were loaded with samples of ANTS labeled Ala-D-γ-Glu-Lys-D-Ala-D-Ala to prevent lane warping. Remaining lanes were loaded with 20% glycerol. Gels were imaged on the ChemiDoc MP system (Bio-Rad) using the Sybr-safe UV imaging setting.

LC-MS analysis of PG fragments

Sample preparation for LC-MS analysis. Lyophilized muropeptides were resuspended in 20 µL of 250 mM NaH_2BO_3 pH 9.0. and treated with ~1 mg $NaBH_4$. Tubes were shaken until completely dissolved and then incubated for 1 hr at room temperature. Reactions were quenched by adding 10 µL of 20% H_3PO_4 and incubating for 1 hr at room temperature. Reduced samples were centrifuged at 16,000 x *g* for 10 min, and supernatants were transferred to LC-MS vials (*41*)

Liquid chromatography (LC). For each sample, 10 µL was injected onto a C18 reversed phase column [Acclaim 120 C18 3um 2.1 x 150mm (DX059130)] at 176 µL/min and kept at 52°C using an isocratic gradient. The molecules were eluted using a gradient increasing from 0% B / 100% A to 100% B / 0% A in 59 minutes (Buffer A: 0.1% TFA, Buffer B: 30% Methanol/0.1% TFA). Solvent composition was held at 100% B / 0% A for 5 minutes where after the column was conditioned for 5 minutes at 0% B / 100% A. All solvents were LCMS grade.

Mass spectrometry (MS). Reversed phase separated molecules were measured in MS mode using an Orbitrap XL operated at 60,000 resolution (Auto Gain Control of 1-2e5 and maximum injection time of 500 ms). Collision Induced Dissociation MS/MS spectra were acquired in ion trap mode (AGC of 1e4 and maximum injection time of 25 ms) fragmenting 1 or 2 of the most abundant ions per cycle.

Data analysis. Total ion and extracted ion chromatograms were analyzed using Xcalibur and Skyline. Molecular masses of PG fragments were determined using ChemDraw.

Intact protein analysis by Matrix-assisted Laser Desorption Time-of-Flight Mass Spectrometry (MALDI-TOF)

1 µL of the sample was mixed with 9 µL of matrix consisting of a saturated solution of α-cyano-4-hydroxycinnamic acid (4-HCCA) in a 1:3:2 (v/v/v) mixture of formic acid/water/isopropanol (FWI). An aliquot of 0.5 – 1 µL of this protein-matrix solution was spotted onto a MALDI plate precoated with an ultrathin layer of 4-HCCA matrix (*148*, *149*). The sample spots were then briefly washed for a few seconds with 2 µL of cold 0.1% aqueous trifluoroacetic acid (TFA) solution and vacuum suction. MALDI spectra were acquired in linear, delayed extraction mode using a Spiral TOF JMS-S3000 (JEOL, Tokyo, Japan). The instrument is equipped with a Nd:YLF laser, delivering 10-Hz pulses at 349 nm. Delayed extraction time was set at 0.75 – 1 µs and acquisition was performed with a sampling rate of 2 ns. The MALDI spectrum consists of an average of 500 scans. Horse myoglobin was used as a mass calibrant with a technique of pseudo-internal calibration wherein a few laser shots on a calibrant spot near a sample spot were collected and averaged with the sample shots into a single spectrum. The spectra were processed and analyzed using MoverZ (Proteometrics, LLC).

Crystallization, X-ray data collection, structure determination and refinement

These experiments were performed by Dr. Byungchul Kim. The initial crystallization conditions for SagA-NlpC/p60 were identified using commercial screen solutions (Molecular

Dimensions) by the sitting-drop vapor-diffusion method at 18°C. The final optimized crystals were obtained using a precipitant solution consisting of 2 M ammonium sulfate and 0.1 M Bis-Tris, pH 5.5, again using the sitting-drop vapor-diffusion method at 18°C. The crystals were flash-cooled in liquid nitrogen using crystallization mother liquor without additional cryoprotectant. X-ray diffraction data were collected from a single crystal on beamline AMX at NSLS-II beamline (Brookhaven National Laboratory) to 2.4 Å resolution. The data were processed with the HKL2000 program suite (*150*). Initial phase estimates and electron-density maps were obtained by molecular replacement with Phaser (*151*) using the C-terminus of CwlT (PDB: 4FDY) as an initial search model in Phenix (*152*). Then the structure was improved by building in COOT and refinement in Phenix iteratively (*153*). Comprehensive model validation was performed with MolProbity (*154*) with 97.4/2.6% of residues falling within the favored and allowed region of the Ramachandran plot, respectively. (data-collection and refinement statistics are summarized in Supplementary file 5 here (*40*)). All molecular graphics were prepared with PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.). Atomic coordinates and experimental structure factors have been deposited in the PDB under accession code 6B8C.

Secondary structure homology modeling

The secondary structure homology of the *Efm* Com15 SagA N-terminus (residues 27-268) was evaluated by submitting the sequence to Phyre2 (*106*). The confidence is reported, or the probability that the match between the *Efm* Com15 SagA N-terminus sequence and the top hit, *Streptococcus pneumoniae* PcsB (*89*), is based on predicted secondary structure and sequence identity.

Molecular docking

Chapter 2

This study was performed by Dr. Byungchul Kim. The Grid-based Ligand Docking with Energetics (GLIDE) module of the Schrödinger suite was used to generate a list of poses with highest scores (*108*, *109*). The docking search space was specified by setting the ligand diameter midpoint cubic box to $15 \times 15 \times 15$ Å^3, which broadly covers the whole region of the SagA-NlpC/p60 protein. The key ligand binding cysteine residues identified from previous biochemistry studies on *E. faecium*_/ *E. faecalis* PG digested by SagA-NlpC/p60, which showed a loss of function on mutation, were used to define the ligand recognition sites. Standard parameters were applied including van der Waals (vdW) scaling of nonpolar atoms (by 0.7) to include modest 'induced fit' effect of ligand. The molecular docking of SagA-NlpC/p60 with PG ligands was done in a flexible and non-constrained manner, allowing the ligand to move freely over the entire volume of the grid box. All other settings were kept as default and docking simulations were performed in two steps which included initial validation of the standard precision (SP) docking algorithms to predict the most accurate binding of ligand with apo SagA-NlpC/p60 structure, then used molecular mechanics generalized born surface area (MM-GBSA) calculations of binding free energies of top-ranked ligand poses (*155*).

Chapter 3

The Grid-based Ligand Docking with Energetics (GLIDE) module of the Schrödinger suite was used to generate a list of poses with highest scores (*108*, *109*) similar to our previously described docking studies (*40*). The raw SagA-NlpC/p60 structure was first prepared for the

docking study using the Schrödinger Preparation Wizard and then aligned to the NlpC/p60 domain of the *Bacillus cereus* YkfC structure (PDB: 3H41). The ligand recognition site in SagA-NlpC/p60 was centered on the L-Ala-γ-D-isoGlu dipeptide bound to BcYkfC-NlpC/p60. The docking search space was specified by setting the ligand diameter midpoint cubic box to 10 x 10 x 15 $Å^3$, which broadly covers the solvent exposed groove of SagA-NlpC/p60. Standard parameters were applied including van der Waals (vdW) scaling of nonpolar atoms (by 0.7) to include modest 'induced fit' effect of ligand. The molecular docking of SagA-NlpC/p60 with the GlcNAc-MurNAc-L-Ala-D-isoGln-L-Lys-D-Ala-D-Ala ligand was done in a flexible, non-constrained manner and restricted to the reference position of the L-Ala-γ-D-isoGlu dipeptide bound to BcYkfC-NlpC/p60; the inclusion of this restriction allowed the GlcNAc-MurNAc and D-Ala-D-Ala groups to rotate freely within the volume of the grid box. All other settings were kept as default. Docking simulations were performed in two steps, which included initial validation of the standard precision (SP) docking algorithms to predict the most accurate binding of ligand with apo SagA-NlpC/p60 structure, and then molecular mechanics generalized born surface area (MM-GBSA) calculations of binding free energies of top-ranked ligand poses (**Appendix 3.3**) (*155*).

Residue scanning

The residue scanning panel was used in Schrödinger to produce the following mutated structures: SagA-NlpC/p60 Y493A, SagA-NlpC/p60 H506A, and SagA-NlpC/p60 E512A. A Prime side-chain prediction with backbone sampling was performed to refine the mutated residue and nearby residues in each mutant structure.

Structure alignment

The *align* function was used in PyMOL v2.3.5 (Schrödinger) to perform a sequence alignment followed by structural superposition of NlpC/p60 orthologs.

Sequence analysis and phylogeny

When possible, sequences that were directly PCR amplified from genomic DNA were used. Annotated *Enterococcus* proteome sequences were obtained from UniProt. Proteomes were searched for sequences containing InterPro annotations for PG hydrolase domains (*44*). Proteins were removed if they contained either InterPro annotations for phage proteins or if they were identified as a prophage sequence by the PHASTER tool (*156*). Proteins were then aligned by Clustal Omega and output in FASTA format. FASTA files were uploaded to the IQ-TREE tool using default settings with the auto substitution model, and 10,000 bootstrap alignments were calculated under the ultrafast setting (*157–159*). The resulting phylogenetic tree files were visualized using iTOL (*160*). Domain regions were predicted using InterPro analysis. Sequence identity was calculated using the SIM tool via ExPASy. Alignments were visualized using ESPript 3.0 (*161*).

Experiments related to mouse tumor models (Chapter 4)

Dr. Matthew Griffin performed the following, as described in the bioRxiv submission (*43*): animal care and experiments; antibiotic pre-treatment; bacterial administration; tumor inoculation; tumor measurements; antibody administration; and colony forming unit (CFU) analysis.

Rise Therapeutics derived the *Lactococcus lactis* strains, as described in the bioRxiv submission (*43*).

Experiments related to phage 9181 *E. faecium* Com12 strains (Chapter 5)

Refer to the following publication (*135*) for materials and methods used by the Duerkop laboratory concerning: bacteriophage isolation and purification; whole-genome sequence analysis of phage-resistant bacteria; phage susceptibility assays; isolation of phage-resistant *E. faecium* strains; determination of phage host range; and phage-antibiotic synergy assays.

Appendix

Appendix Table 1. Bacterial strains used in this thesis.

Organism	Description	Plasmid Backbone	Abx Resistance
Enterococcus faecium Com15	WT	None	None
Enterococcus faecium Com15	SagA-His_6	pAM401	Cam
Enterococcus faecium Com15	SagA-Cys_4-His_6	pAM401	Cam
Enterococcus faecium Com15	$SagA^{\Delta C}$-His_6	pAM401	Cam
Enterococcus faecium Com15	$SagA^{\Delta C}$-Cys_4-His_6	pAM401	Cam
Enterococcus faecium Com15	$SagA^{Nter}$-His_6	pAM401	Cam
Enterococcus faecium Com15	$SagA^{Nter}$-Cys_4-His_6	pAM401	Cam
Enterococcus faecium Com12	WT	None	None
Enterococcus faecium Com12 81RS3	SagA W433L	None	None
Enterococcus faecium Com12 81RS4	SagA W433C	None	None
Enterococcus faecium Com12 81RS5	SagA G460D	None	None
Enterococcus faecium Com12 81RS6	SagA Y451_L452_L insertion	None	None
Enterococcus faecium Com12 81RS8	SagA G435V	None	None
Enterococcus faecium DO	WT	None	None
Enterococcus faecium 7171	WT	None	None
Enterococcus faecium TX0082	WT	None	Multi-drug
Enterococcus faecium ERV165	WT	None	Multi-drug
Enterococcus faecium E1162	WT	None	Multi-drug
Enterococcus durans 23C2	WT	None	None
Enterococcus hirae R	WT	None	None
Enterococcus mundtii NCDO 2312	WT	None	None
Enterococcus gallinarum NCDO 2375	WT	None	None
Enterococcus casseliflavus 20	WT	None	None
Enterococcus faecalis OG1RF	WT	None	None
Enterococcus faecalis OG1RF	SagA-His_6	None	None
Enterococcus faecalis 775	WT	None	None
Enterococcus faecalis V583	WT	None	None

Lactococcus lactis MG1363::Δ*thyA*	Null	None	None / *thyA* auxotroph
Lactococcus lactis MG1363::Δ*thyA*	SagA WT	None	None / *thyA* auxotroph
Lactococcus lactis MG1363::Δ*thyA*	SagA C443A	None	None / *thyA* auxotroph
Lactococcus lactis MG1363::Δ*thyA*	SagA ΔSS	None	None / *thyA* auxotroph
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA_C443A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA_ΔSS-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA_ΔSS_C443A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA_ΔSS_H506A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_Y431A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_W433A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_W433L-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_W433C-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_G435V-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_K436A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_D442A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_C443A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_S444A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_R448A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_G460D-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_W462A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_T463A-His_6	pET-21a(+)	Amp, Cam[a]

Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_Y493A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_H494A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_H506A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_E512A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_K515A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_V516A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_W521A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_F522A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_ W433A_W62A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_Com15_SagA-NlpC/p60_ΔSS_Y451_L452_insL-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Eds*_23C2_SagA_ΔSS -His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Eds*_23C2_SagA_ΔSS_H503A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Emi*_2312_SagA_ΔSS-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Emi*_2312_SagA_ΔSS_H473A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Ehe*_R_SagA_ΔSS-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efs*_OG1RF_SalA_ΔSS-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efs*_OG1RF_SalB_ΔSS-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_TX0082_501F_ΔSS-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_TX0082_501F_ΔNter_ΔSS-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_TX0082_501F_ΔNter_ΔCter_ΔSS-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_TX0082_501F_ΔNter_ΔCter_ΔSS_C335A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_TX0082_51B9_CwlT_ΔSS-His_6	pET-21a(+)	Amp, Cam[a]

Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_TX0082_51B9_CwlT_ΔSS_C240A-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_TX0082_51B9_CwlT-NlpC/p60_ΔSS-His_6	pET-21a(+)	Amp, Cam[a]
Escherichia coli BL21-CodonPlus (DE3)-RIL	*Efm*_TX0082_51B9_CwlT-NlpC/p60_ΔSS_C240A-His_6	pET-21a(+)	Amp, Cam[a]

[a]*Escherichia coli* BL21-CodonPlus (DE3)-RIL cells contain extra copies of the argU, ileY, and leuW tRNA genes in a plasmid encoding Cam resistance

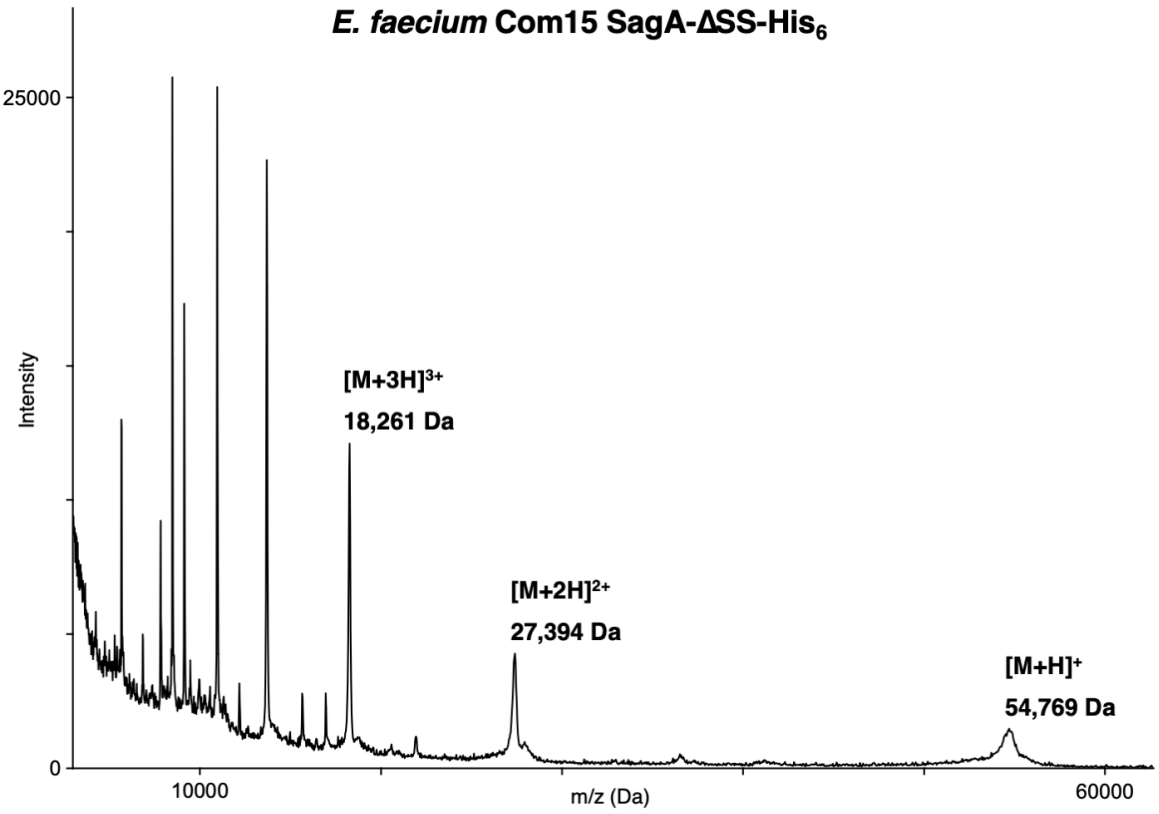
Appendix Table 2. Primers used for cloning and mutagenesis in this thesis.

#	Primer Name	Sequence (5' to 3')	Ta (˚C)
1	Efm SagA-NlpC/p60 C443A F	TGGATTTGACgccTCAGGATTCACACG	61
2	Efm SagA-NlpC/p60 C443A R	CTTGGATCTTTACCGCCC	
3	Efm SagA-NlpC/p60 W433A F	TCCTTATGTTgcgGGCGGTAAAG	59
4	Efm SagA-NlpC/p60 W433A R	GTACCAATATATTTGTAAGCTTC	
5	Efm SagA-NlpC/p60 W462A F	CATTGGTGGTgcgACAGTTCCTC	59
6	Efm SagA-NlpC/p60 W462A R	TCACGACCAGTTACTTGC	
7	SagA-Cys_4-His_6_S	CTTTGCTGTCAGCATGTGCTGCCCTGGATGCTGCCTCGAGCACCACCACC	57
8	SagA-Cys_4-His_6_AS	GGTGGTGGTGCTCGAGGCAGCATCCAGGGCAGCACATGCTGACAGCAAAG	
9	$SagA^{\Delta C}$-Cys_4-His_6_S	AGTGTGGATCCCATGTGCTGCCCTGGATGCTGCCTCGAGCACCACCAC	56
10	$SagA^{\Delta C}$-Cys_4-His_6_AS	GTGGTGGTGCTCGAGGCAGCATCCAGGGCAGCACATGGGATCCACACT	
11	$SagA^{Nter}$-Cys_4-His_6_S	TGAACGTACTGAAAACTTGCTGCCCTGGATGCTGCCTCGAGCACCACCACC	56
12	$SagA^{Nter}$-Cys_4-His_6_AS	GGTGGTGGTGCTCGAGGCAGCATCCAGGGCAGCAAGTTTTCAGTACGTTCA	
13	Efm SagA-NlpC/p60 Y431A F	TGGTACTCCTgctGTTTGGGGCG	60
14	Efm SagA-NlpC/p60 Y431A R	ATATATTTGTAAGCTTCTGCTAC	
15	Efm SagA-NlpC/p60 K436A F	TTGGGGCGGTgcaGATCCAAGTG	62
16	Efm SagA-NlpC/p60 K436A R	ACATAAGGAGTACCAATATATTTGTAAG	
17	Efm SagA-NlpC/p60 R448A F	AGGATTCACAgccTATGTTTACTTGCAAG	58
18	Efm SagA-NlpC/p60 R448A R	GAGCAGTCAAATCCACTTG	
19	Efm SagA-NlpC/p60 T463A F	TGGTGGTTGGgctGTTCCTCAAG	63
20	Efm SagA-NlpC/p60 T463A R	ATGTCACGACCAGTTACTTG	
21	Efm SagA-NlpC/p60 Y493A F	AGGCGGAACTgctCACGTAGCGA	62

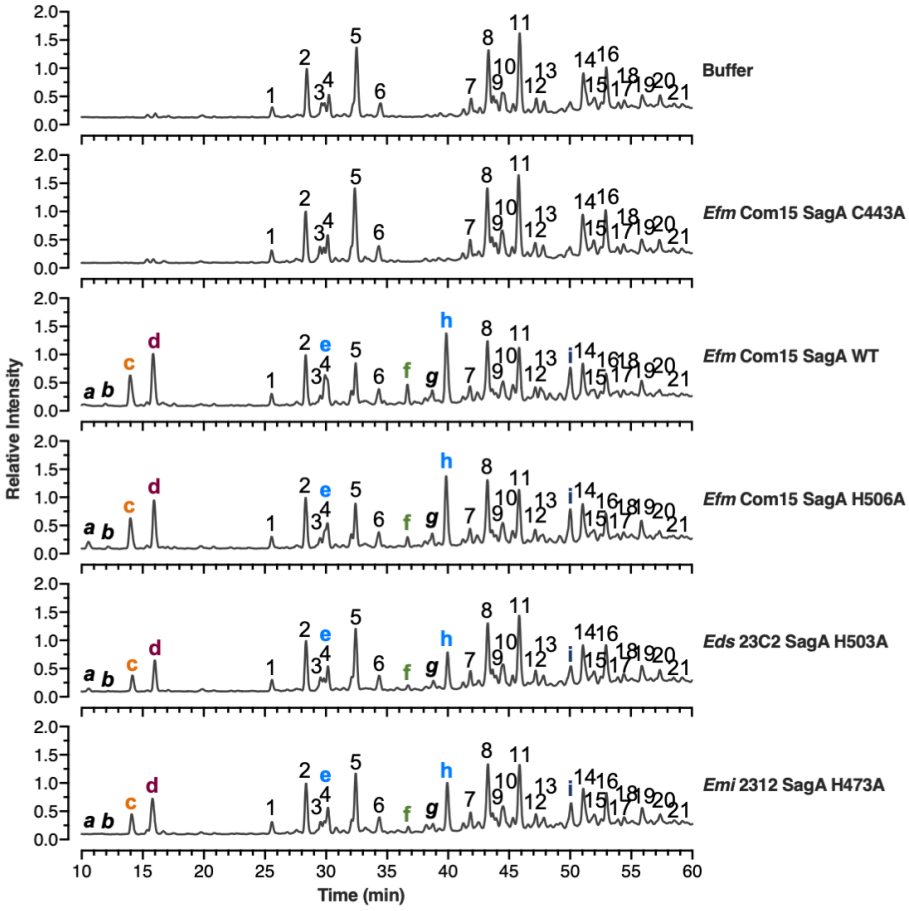
22	Efm SagA-NlpC/p60 Y493A R	GGTGAACCCCAGAATAATAAG	
23	Efm SagA-NlpC/p60 E512A F	TCAACCAGGCgctAGCGTAAAAG	58
24	Efm SagA-NlpC/p60 E512A R	GGAGCATGGATATATTGTC	
25	Efm SagA-NlpC/p60 K515A F	CGAAAGCGTAgcaGTTGGTTCAGTTC	62
26	Efm SagA-NlpC/p60 K515A R	CCTGGTTGAGGAGCATGG	
27	Efm SagA-NlpC/p60 V516A F	AAGCGTAAAAgccGGTTCAGTTCAATGGTTTGC	66
28	Efm SagA-NlpC/p60 V516A R	TCGCCTGGTTGAGGAGCA	
29	Efm SagA-NlpC/p60 W521A F	TTCAGTTCAAgcaTTTGCACCTGACTTTGC	58
30	Efm SagA-NlpC/p60 W521A R	CCAACTTTTACGCTTTCG	
31	Efm SagA-NlpC/p60 F522A F	AGTTCAATGGgctGCACCTGACTTTG	60
32	Efm SagA-NlpC/p60 F522A R	GAACCAACTTTTACGCTTTC	
33	Efm SagA-NlpC/p60 D442A F	AAGTGGATTTgcgTGCTCAGGATTCACAC	61
34	Efm SagA-NlpC/p60 D442A R	GGATCTTTACCGCCCCAA	
35	Efm SagA-NlpC/p60 C443A F	TGGATTTGACgccTCAGGATTCACACG	61
36	Efm SagA-NlpC/p60 C443A R	CTTGGATCTTTACCGCCC	
37	Efm SagA-NlpC/p60 S444A F	ATTTGACTGCgcgGGATTCACAC	61
38	Efm SagA-NlpC/p60 S444A R	CCACTTGGATCTTTACCG	
39	Efm SagA-NlpC/p60 H494A F	CGGAACTTACgcgGTAGCGATTGCTTTAGGC	61
40	Efm SagA-NlpC/p60 H494A R	CCTGGTGAACCCCAGAAT	
41	Efm SagA-NlpC/p60 H506A F	ACAATATATCgctGCTCCTCAACCAGGC	61
42	Efm SagA-NlpC/p60 H506A R	CCTCCGCCTAAAGCAATC	
43	ΔSS F	GACGATTTTGATTCTCAGATAC	57
44	ΔSS R	CATATGTATATCTCCTTCTTAAAG	

45	Eds SagA F NdeI ΔSS	tgtttaactttaagaaggagatatacatatgGCGGACGATTTTGATTCA	61
46	Eds SagA R XhoI ΔSS	gatctcagtggtggtggtggtggtgctcgagCATACTTACTGCAAAATCAGGT	
47	Emi SagA F NdeI ΔSS	tgtttaactttaagaaggagatatacatatgGCGGAAGATTTTGATTCTCAAATAC	63
48	Emi SagA R XhoI ΔSS	gatctcagtggtggtggtggtggtgctcgagCATACTTACAGCAAAGTCAGGTG	
49	Eds SagA H506A F	ACAATATATCgccGCACCTCAACCAGG	61
50	Eds SagA H506A R	CCGCCACCTAAAGCAATT	
51	Emi SagA H506A F	ACAATACATCgctGCACCACAACCG	59
52	Emi SagA H506A R	CCGCCACCTAAAGAAATTG	
53	Ehe SagA F NdeI ΔSS	tgtttaactttaagaaggagatatacatatgGCGGACGATTTTGATTCT	61
54	Ehe SagA R XhoI ΔSS	gatctcagtggtggtggtggtggtgctcgagCATACTTACTGCAAAATCAGGT	
55	Efs SalA F NdeI ΔSS	tgtttaactttaagaaggagatatacatatgGATGAATACGATACAAAGATTCAACAAC	63
56	Efs SalA R XhoI ΔSS	gatctcagtggtggtggtggtggtgctcgagGCCTGGAAAACGCCATACAAC	
57	Efs SalB F NdeI ΔSS	actttaagaaggagatatacatatgGACAATGTTGATAAAAAAATTGAAGAAAAAAATCA	63
58	Efs SalB R XhoI ΔSS	gatctcagtggtggtggtggtggtgctcgagGGCTGAGTGTCCTACGATTGT	
59	Efm SagA-NlpC/p60 W433L F 59	TCCTTATGTTttgGGCGGTAAAG	59
60	Efm SagA-NlpC/p60 W433L R 59	GTACCAATATATTTGTAAGCTTC	
61	Efm SagA-NlpC/p60 W433C F 59	TCCTTATGTTtgcGGCGGTAAAG	59
62	Efm SagA-NlpC/p60 W433C R 59	GTACCAATATATTTGTAAGCTTC	
63	Efm SagA-NlpC/p60 G460D F 60	TCGTGACATTgacGGTTGGACAG	60
64	Efm SagA-NlpC/p60 G460D R 60	CCAGTTACTTGCAAGTAAAC	
65	Efm SagA-NlpC/p60 Y451_L452_insL F 60	ttgTTGCAAGTAACTGGTCGTG	60
66	Efm SagA-NlpC/p60 Y451_L452_insL R 60	GTAAACATAGCGTGTGAATC	
67	Efm SagA-NlpC/p60 G435V F 60	TGTTTGGGGCgtcAAAGATCCAAG	60
68	Efm SagA-NlpC/p60 G435V R 60	TAAGGAGTACCAATATATTTGTAAG	

69	TX0082-501F-Amidase F NdeI dSS 58	ctttaagaaggagatatacatatgGTAGATAATCAAATAGCTGATCAAAC	58
70	TX0082-501F-Amidase R XhoI dSS 58	agtggtggtggtggtggtgctcgagTTTCTTTAAAGGCAGATCAAG	
71	TX0082-51B9-CwlT F NdeI dSS 64	ctttaagaaggagatatacatatgGATGAACAGGACAGCGGAATTTC	64
72	TX0082-51B9-CwlT R XhoI dSS 64	agtggtggtggtggtggtgctcgagTTGTTTGATTCGTCCTGCTC	
73	TX0082-501F dNter dSS F 57	AATGTCCAACAAGCAATC	57
74	TX0082-501F dNter dSS R 57	CATATGTATATCTCCTTCTTAAAG	
75	TX0082-501F dCter dSS F 65	CTCGAGCACCACCACCAC	65
76	TX0082-501F dCter dSS R 65	TGCTTCTTTTTTCGGTTCATCATATTTTG	
77	TX0082-51B9 NlpCp60 Domain F 57	TTTGATGATGATACGGTAC	57
78	TX0082-51B9 NlpCp60 Domain R 57	CATATGTATATCTCCTTCTTAAAG	
79	TX0082-501F NlpCp60 C335A F 60	TACTTTTGACgctTCTGGTTTGACTTATTACGTATATATG	60
80	TX0082-501F NlpCp60 C335A R 60	TTTGGACCTTTTGCTCCC	
81	TX0082-51B9 NlpCp60 C240A F 64	TTCTTTTGATgctAGCGGACTGACACAATGGAC	64
82	TX0082-51B9 NlpCp60 C240A R 64	GTAGTCGGGGAAGCTCCA	



Appendix 3.1. Intact MALDI-TOF mass spectrometry analysis of recombinant full-length *E. faecium* Com15 SagA-ΔSS-His$_6$.



Appendix 3.2. Total ion chromatograms corresponding to Figure 3.10. LC-MS analysis of soluble mutanolysin-digested *E. faecalis* OG1RF PG incubated with buffer or full-length SagA orthologs at 37˚C for 16 hr. Data are shown as total ion chromatograms corresponding to Figure 5C. The peak labels and identities are indicated in Table 3.1 and 3.2 and formatted as described for Figure 3.6. The intensity of total ion abundance was normalized to the unchanging GM-tripeptide (AA) peak (2). *Efm* Com15, *E. faecium* Com15; *Eds* 23C2, *E. durans* 23C2; *Ehe* R, *E. hirae* R; *Emi* 2312, *E. mundtii* NCDO 2312; *Egm* 2375, *E. gallinarum* NCDO 2375; *Ecs* 20, *E. casseliflavus* 20.

Appendix 3.3. Table of predicted binding free energies of highest-scoring poses of docked GlcNAc-MurNAc-L-Ala-D-isoGln-L-Lys-D-Ala-D-Ala as generated with MM-GBSA[a]

Pose	Predicted ΔG (kcal/mol)
1	*-26.32*
2	-38.80
3	-32.97
4	-33.26
5	-33.22
6	-50.61
7	-35.90
8	-21.33
9	-30.66
10	-24.08
11	-9.56
12	-15.24
13	-37.39
14	-26.27
15	-35.90
16	-34.03
17	-21.33
18	-50.61
19	-15.47
20	12.45

a. MM-GBSA calculations were carried out using the Prime_MM-GBSA utility. Pose 1 was selected for analysis after manual inspection revealed no steric clashes between the ligand and the SagA-NlpC/p60 structure.

References

1. V. Anantharaman, L. Aravind, Evolutionary history, structural features and biochemical diversity of the NlpC/P60 superfamily of enzymes. *Genome Biol*. **4**, R11 (2003).

2. A. Bateman, N. D. Rawlings, The CHAP domain: a large family of amidases including GSP amidase and peptidoglycan hydrolases. *Trends Biochem Sci*. **28**, 234–237 (2003).

3. R. Ohnishi, S. Ishikawa, J. Sekiguchi, Peptidoglycan hydrolase LytF plays a role in cell separation with CwlF during vegetative growth of Bacillus subtilis. *J Bacteriol*. **181**, 3178–84 (1999).

4. P. Margot, M. Wahlen, A. Gholamhoseinian, P. Piggot, D. Karamata, A. Gholamhuseinian, The lytE Gene of Bacillus subtilis 168 Encodes a Cell Wall Hydrolase. *J Bacteriol*. **180**, 749–752 (1998).

5. T. J. Smith, S. A. Blackman, S. J. Foster, Autolysins of Bacillus subtilis: multiple enzymes with multiple functions. *Microbiology (Reading, England)*. **146 (Pt 2)**, 249–62 (2000).

6. K. Schubert, A. M. Bichlmaier, E. Mager, K. Wolff, G. Ruhland, F. Fiedler, P45, an extracellular 45 kDa protein of Listeria monocytogenes with similarity to protein p60 and exhibiting peptidoglycan lytic activity. *Arch Microbiol*. **173**, 21–28 (2000).

7. A. Ruggiero, D. Marasco, F. Squeglia, S. Soldini, E. Pedone, C. Pedone, R. Berisio, Structure and Functional Regulation of RipA, a Mycobacterial Enzyme Essential for Daughter Cell Separation. *Structure*. **18**, 1184–1190 (2010).

8. J. Pinheiro, J. Biboy, W. Vollmer, R. P. Hirt, J. R. Keown, A. Artuyants, M. M. Black, D. C. Goldstone, A. Simoes-Barbosa, The Protozoan Trichomonas vaginalis Targets Bacteria with Laterally Acquired NlpC/P60 Peptidoglycan Hydrolases. *Mbio*. **9**, e01784-18 (2018).

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

10. A. Typas, M. Banzhaf, C. A. Gross, W. Vollmer, From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nature Reviews Microbiology*. **10**, nrmicro2677 (2011).

11. M. Wuenscher, S. Köhler, A. Bubert, U. Gerike, W. Goebel, The iap gene of Listeria monocytogenes is essential for cell viability, and its gene product, p60, has bacteriolytic activity. *J Bacteriol*. **175**, 3491–3501 (1993).

12. S. Pilgrim, A. Kolb-Maurer, I. Gentschev, W. Goebel, M. Kuhn, Deletion of the Gene Encoding p60 in Listeria monocytogenes Leads to Abnormal Cell Division and Loss of Actin-Based Motility. *Infect Immun*. **71**, 3473–3484 (2003).

13. Y. Tsuge, H. Ogino, H. Teramoto, M. Inui, H. Yukawa, Deletion of cgR_1596 and cgR_2070, Encoding NlpC/P60 Proteins, Causes a Defect in Cell Separation in Corynebacterium glutamicum R. *J Bacteriol*. **190**, 8204–8214 (2008).

14. M.-C. Duchêne, T. Rolain, A. Knoops, P. Courtin, M.-P. Chapot-Chartier, Y. F. Dufrêne, B. F. Hallet, P. Hols, Distinct and Specific Role of NlpC/P60 Endopeptidases LytA and LytB in Cell Elongation and Division of Lactobacillus plantarum. *Frontiers in Microbiology*. **10**, 713 (2019).

15. S. Ishikawa, Y. Hara, R. Ohnishi, J. Sekiguchi, Regulation of a new cell wall hydrolase gene, cwlF, which affects cell separation in Bacillus subtilis. *J Bacteriol*. **180**, 2549–55 (1998).

16. H. Yamamoto, S. -i Kurosawa, J. Sekiguchi, Localization of the Vegetative Cell Wall Hydrolases LytC, LytE, and LytF on the Bacillus subtilis Cell Surface and Stability of These Enzymes to Cell Wall-Bound or Extracellular Proteases. *J Bacteriol*. **185**, 6666–6677 (2003).

17. T. Fukushima, A. Afkham, -i S Kurosawa, T. Tanabe, H. Yamamoto, J. Sekiguchi, A New D,L-Endopeptidase Gene Product, YojL (Renamed CwlS), Plays a Role in Cell Separation with LytE and LytF in Bacillus subtilis. *J Bacteriol*. **188**, 5541–5550 (2006).

18. R. Carballido-López, A. Formstone, Y. Li, S. D. Ehrlich, P. Noirot, J. Errington, Actin Homolog MreBH Governs Cell Morphogenesis by Localization of the Cell Wall Hydrolase LytE. *Dev Cell*. **11**, 399–409 (2006).

19. E. C. Hett, E. J. Rubin, Bacterial growth and cell division: a mycobacterial perspective. *Microbiol Mol Biology Rev Mmbr*. **72**, 126–56, table of contents (2008).

20. K. J. Kieser, E. J. Rubin, How sisters grow apart: mycobacterial growth and division. *Nat Rev Microbiol*. **12**, 550–562 (2014).

21. F. Squeglia, M. Moreira, A. Ruggiero, R. Berisio, The Cell Wall Hydrolytic NlpC/P60 Endopeptidases in Mycobacterial Cytokinesis: A Structural Perspective. *Cells*. **8**, 609 (2019).

22. D. J. Martinelli, M. S. Pavelka, The RipA and RipB Peptidoglycan Endopeptidases Are Individually Nonessential to Mycobacterium smegmatis. *Journal of Bacteriology*. **198**, 1464–1475 (2016).

23. C. Healy, A. Gouzy, S. Ehrt, Peptidoglycan Hydrolases RipA and Ami1 Are Critical for Replication and Persistence of Mycobacterium tuberculosis in the Host. *Mbio*. **11**, e03315-19 (2020).

24. E. C. Hett, M. C. Chao, A. J. Steyn, S. M. Fortune, L. L. Deng, E. J. Rubin, A partner for the resuscitation-promoting factors of Mycobacterium tuberculosis. *Molecular Microbiology*. **66**, 658–668 (2007).

25. E. C. Hett, M. C. Chao, L. L. Deng, E. J. Rubin, A Mycobacterial Enzyme Essential for Cell Division Synergizes with Resuscitation-Promoting Factor. *Plos Pathog*. **4**, e1000001 (2008).

26. M. C. Chao, K. J. Kieser, S. Minami, D. Mavrici, B. B. Aldridge, S. M. Fortune, T. Alber, E. J. Rubin, Protein complexes and proteolytic activation of the cell wall hydrolase RipA regulate septal resolution in mycobacteria. *Plos Pathog*. **9**, e1003197 (2013).

27. H. Botella, J. Vaubourgeix, M. H. Lee, N. Song, W. Xu, H. Makinoshima, M. S. Glickman, S. Ehrt, Mycobacterium tuberculosis protease MarP activates a peptidoglycan hydrolase during acid stress. *Embo J*. **36**, 536–548 (2017).

28. D. Mavrici, M. J. Marakalala, J. M. Holton, D. M. Prigozhin, C. L. Gee, Y. J. Zhang, E. J. Rubin, T. Alber, Mycobacterium tuberculosis FtsX extracellular domain activates the peptidoglycan hydrolase, RipC. *Proceedings of the National Academy of Sciences*. **111**, 8037–8042 (2014).

29. A. C. Storer, R. Ménard, Methods in Enzymology. *Methods Enzymol*. **244**, 486–500 (1994).

30. H.-H. Otto, T. Schirmeister, Cysteine Proteases and Their Inhibitors. *Chem Rev*. **97**, 133–172 (1997).

31. M. E. McGrath, THE LYSOSOMAL CYSTEINE PROTEASES. *Annu Rev Bioph Biom*. **28**, 181–204 (1999).

32. J. M. Aramini, P. Rossi, Y. J. Huang, L. Zhao, M. Jiang, M. Maglaqui, R. Xiao, J. Locke, R. Nair, B. Rost, T. B. Acton, M. Inouye, G. T. Montelione, Solution NMR structure of the NlpC/P60 domain of lipoprotein Spr from Escherichia coli: structural evidence for a novel cysteine peptidase catalytic triad. *Biochemistry*. **47**, 9715–7 (2008).

33. Q. Xu, S. Sudek, D. McMullan, M. D. Miller, B. Geierstanger, D. H. Jones, S. S. Krishna, G. Spraggon, B. Bursalay, P. Abdubek, C. Acosta, E. Ambing, T. Astakhova, H. L. Axelrod, D. Carlton, J. Caruthers, H.-J. Chiu, T. Clayton, M. C. Deller, L. Duan, Y. Elias, M.-A. Elsliger, J. Feuerhelm, S. K. Grzechnik, J. Hale, G. Han, J. Haugen, L. Jaroszewski, K. K. Jin, H. E. Klock, M. W. Knuth, P. Kozbial, A. Kumar, D. Marciano, A. T. Morse, E. Nigoghossian, L. Okach, S. Oommachen, J. Paulsen, R. Reyes, C. L. Rife, C. V. Trout, H. van den Bedem, D. Weekes, A. White, G. Wolf, C. Zubieta, K. O. Hodgson, J. Wooley, A. M. Deacon, A. Godzik, S. A. Lesley, I. A. Wilson, Structural Basis of Murein Peptide Specificity of a γ-D-Glutamyl-L-Diamino Acid Endopeptidase. *Structure*. **17**, 303–313 (2009).

34. N. D. Rawlings, A. J. Barrett, (2013), *Handbook of Proteolytic Enzymes*.

35. B. Elsässer, F. B. Zauner, J. Messner, W. Soh, E. Dall, H. Brandstetter, Distinct Roles of Catalytic Cysteine and Histidine in the Protease and Ligase Mechanisms of Human Legumain As Revealed by DFT-Based QM/MM Simulations. *Acs Catal*. **7**, 5585–5593 (2017).

36. F. Squeglia, A. Ruggiero, M. Romano, L. Vitagliano, R. Berisio, Mutational and structural study of RipA, a key enzyme in Mycobacterium tuberculosis cell division: evidence for the l-to-d inversion of configuration of the catalytic cysteine. *Acta Crystallographica Section D: Biological Crystallography*. **70**, 2295–300 (2014).

37. Q. Xu, H.-J. Chiu, C. L. Farr, L. Jaroszewski, M. W. Knuth, M. D. Miller, S. A. Lesley, A. Godzik, M.-A. Elsliger, A. M. Deacon, I. A. Wilson, Structures of a Bifunctional Cell Wall Hydrolase CwlT Containing a Novel Bacterial Lysozyme and an NlpC/P60 dl-Endopeptidase. *Journal of Molecular Biology*. **426**, 169–184 (2014).

38. T. Fukushima, T. Kitajima, H. Yamaguchi, Q. Ouyang, K. Furuhata, H. Yamamoto, T. Shida, J. Sekiguchi, Identification and Characterization of Novel Cell Wall Hydrolase CwlT A TWO-DOMAIN AUTOLYSIN EXHIBITING N-ACETYLMURAMIDASE AND dl-ENDOPEPTIDASE ACTIVITIES. *Journal of Biological Chemistry*. **283**, 11117–11125 (2008).

39. D. Böth, E. Steiner, A. Izumi, G. Schneider, R. Schnell, RipD (Rv1566c) from Mycobacterium tuberculosis: adaptation of an NlpC/p60 domain to a non-catalytic peptidoglycan-binding function. *Biochem J*. **457**, 33–41 (2014).

40. B. Kim, Y.-C. Wang, C. W. Hespen, J. Espinosa, J. Salje, K. J. Rangan, D. A. Oren, J. Kang, V. A. Pedicord, H. C. Hang, Enterococcus faecium secreted antigen A generates muropeptides to enhance host immunity and limit bacterial pathogenesis. *eLife*. **8**, e45343 (2019).

41. B. Kim, J. Espinosa, H. C. Hang, Biochemical analysis of NlpC/p60 peptidoglycan hydrolase activity. *Methods Enzymol* (2020), doi:10.1016/bs.mie.2020.02.017.

42. J. Espinosa, T.-Y. Lin, Y. Estrella, B. Kim, H. Molina, H. C. Hang, Enterococcus NlpC/p60 Peptidoglycan Hydrolase SagA Localizes to Sites of Cell Division and Requires Only a Catalytic Dyad for Protease Activity. *Biochemistry-us* (2020), doi:10.1021/acs.biochem.0c00755.

43. M. E. Griffin, J. Espinosa, J. L. Becker, J. K. Jha, G. R. Fanger, H. C. Hang, *Biorxiv*, in press, doi:10.1101/2020.08.20.256263.

44. A. Vermassen, S. Leroy, R. Talon, C. Provot, M. Popowska, M. Desvaux, Cell Wall Hydrolases in Bacteria: Insight on the Diversity of Cell Wall Amidases, Glycosidases and Peptidases Toward Peptidoglycan. *Frontiers in Microbiology*. **10**, 331 (2019).

45. W. Vollmer, B. Joris, P. Charlier, S. Foster, Bacterial peptidoglycan (murein) hydrolases. *FEMS Microbiology Reviews*. **32**, 259–286 (2008).

46. Y. Redko, P. Courtin, C. Mezange, C. Huard, -P M Chapot-Chartier, Lactococcus lactis Gene yjgB Encodes a -D-Glutaminyl-L-Lysyl- Endopeptidase Which Hydrolyzes Peptidoglycan. *Appl Environ Microb*. **73**, 5825–5831 (2007).

47. P. Mitkowski, E. Jagielska, E. Nowak, J. M. Bujnicki, F. Stefaniak, D. Niedziałek, M. Bochtler, I. Sabała, Structural bases of peptidoglycan recognition by lysostaphin SH3b domain. *Sci Rep-uk*. **9**, 5965 (2019).

48. J. Kurushima, I. Hayashi, M. Sugai, H. Tomita, Bacteriocin protein BacL1 of Enterococcus faecalis is a peptidoglycan D-isoglutamyl-L-lysine endopeptidase. *J Biological Chem*. **288**, 36915–25 (2013).

49. H. Sekiya, E. Tamai, J. Kawasaki, K. Murakami, S. Kamitori, Structural and biochemical characterizations of the novel autolysin Acd24020 from Clostridioides difficile and its full-function catalytic domain as a lytic enzyme. *Mol Microbiol* (2020), doi:10.1111/mmi.14636.

50. J. Lisboa, C. Pereira, A. Rifflet, J. Ayala, M. S. Terceti, A. V. Barca, I. Rodrigues, P. J. B. Pereira, C. R. Osorio, F. G. Portillo, I. G. Boneca, A. do Vale, N. M. S. dos Santos, A Secreted NlpC/P60 Endopeptidase from Photobacterium damselae subsp. piscicida Cleaves the Peptidoglycan of Potentially Competing Bacteria. *Msphere*. **6** (2021), doi:10.1128/msphere.00736-20.

51. J. Wong, H. Alsarraf, J. Kaspersen, J. Pedersen, J. Stougaard, S. Thirup, M. Blaise, Cooperative binding of LysM domains determines the carbohydrate affinity of a bacterial endopeptidase protein. *FEBS Journal*. **281**, 1196–1208 (2014).

52. J. E. M. M. Wong, S. R. Midtgaard, K. Gysel, M. B. Thygesen, K. K. Sørensen, K. J. Jensen, J. Stougaard, S. Thirup, M. Blaise, An intermolecular binding mechanism involving multiple LysM domains mediates carbohydrate recognition by an endopeptidase. *Acta Crystallographica Section D: Biological Crystallography*. **71**, 592–605 (2015).

53. J. B. Raymond, S. Mahapatra, D. C. Crick, M. S. Pavelka, Identification of the namH Gene, Encoding the Hydroxylase Responsible for the N-Glycolylation of the Mycobacterial Peptidoglycan*. *J Biol Chem*. **280**, 326–333 (2005).

54. D. Böth, G. Schneider, R. Schnell, Peptidoglycan Remodeling in Mycobacterium tuberculosis: Comparison of Structures and Catalytic Activities of RipA and RipB. *Journal of Molecular Biology*. **413**, 247–260 (2011).

55. T. DeWitt, A. D. Grossman, The Bifunctional Cell Wall Hydrolase CwlT Is Needed for Conjugation of the Integrative and Conjugative Element ICEBs1 in Bacillus subtilis and B. anthracis. *J Bacteriol*. **196**, 1588–1596 (2014).

56. S. Singh, L. SaiSree, R. N. Amrutha, M. Reddy, Three redundant murein endopeptidases catalyse an essential cleavage step in peptidoglycan synthesis of Escherichia coli K12. *Mol Microbiol*. **86**, 1036–1051 (2012).

57. Q. Xu, D. Mengin-Lecreulx, D. Patin, J. C. Grant, H.-J. Chiu, L. Jaroszewski, M. W. Knuth, A. Godzik, S. A. Lesley, M.-A. Elsliger, A. M. Deacon, I. A. Wilson, Structure-Guided

Functional Characterization of DUF1460 Reveals a Highly Specific NlpC/P60 Amidase Family. *Structure*. **22**, 1799–1809 (2014).

58. Q. Xu, D. Mengin-Lecreulx, X. W. Liu, D. Patin, C. L. Farr, J. C. Grant, H.-J. Chiu, L. Jaroszewski, M. W. Knuth, A. Godzik, S. A. Lesley, M.-A. Elsliger, A. M. Deacon, I. A. Wilson, Insights into Substrate Specificity of NlpC/P60 Cell Wall Hydrolases Containing Bacterial SH3 Domains. *mBio*. **6**, e02327-14 (2015).

59. J. P. Bannantine, C. K. Lingle, P. R. Adam, K. X. Ramyar, W. J. McWhorter, J. R. Stabel, W. D. Picking, B. V. Geisbrecht, NlpC/P60 domain-containing proteins of Mycobacterium avium subspecies paratuberculosis that differentially bind and hydrolyze peptidoglycan. *Protein Sci*. **25**, 840–851 (2016).

60. Q. Xu, N. D. Rawlings, H.-J. Chiu, L. Jaroszewski, H. E. Klock, M. W. Knuth, M. D. Miller, M.-A. Elsliger, A. M. Deacon, A. Godzik, S. A. Lesley, I. A. Wilson, Structural Analysis of Papain-Like NlpC/P60 Superfamily Enzymes with a Circularly Permuted Topology Reveals Potential Lipid Binding Sites. *PLoS ONE*. **6**, e22013 (2011).

61. Q. Xu, P. Abdubek, T. Astakhova, H. L. Axelrod, C. Bakolitsa, X. Cai, D. Carlton, C. Chen, H.-J. Chiu, M. Chiu, T. Clayton, D. Das, M. C. Deller, L. Duan, K. Ellrott, C. L. Farr, J. Feuerhelm, J. C. Grant, A. Grzechnik, G. Han, L. Jaroszewski, K. K. Jin, H. E. Klock, M. W. Knuth, P. Kozbial, S. S. Krishna, A. Kumar, W. W. Lam, D. Marciano, M. D. Miller, A. T. Morse, E. Nigoghossian, A. Nopakun, L. Okach, C. Puckett, R. Reyes, H. J. Tien, C. B. Trame, H. van den Bedem, D. Weekes, T. Wooten, A. Yeh, K. O. Hodgson, J. Wooley, M.-A. Elsliger, A. M. Deacon, A. Godzik, S. A. Lesley, I. A. Wilson, Structure of the γ-d-glutamyl-l-diamino acid endopeptidase YkfC from Bacillus cereus in complex with l-Ala-γ-d-Glu: insights into substrate recognition by NlpC/P60 cysteine peptidases. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*. **66**, 1354–1364 (2010).

62. M. Marino, M. Banerjee, R. Jonquières, P. Cossart, P. Ghosh, GW domains of the Listeria monocytogenes invasion protein InlB are SH3-like and mediate binding to host ligands. *Embo J*. **21**, 5623–5634 (2002).

63. J. Z. Lu, T. Fujiwara, H. Komatsuzawa, M. Sugai, J. Sakon, Cell Wall-targeting Domain of Glycylglycine Endopeptidase Distinguishes among Peptidoglycan Cross-bridges*. *J Biol Chem*. **281**, 549–558 (2006).

64. T. J. Smith, S. A. Blackman, S. J. Foster, Autolysins of Bacillus subtilis: multiple enzymes with multiple functions. *Microbiology+*. **146**, 249–262 (2000).

65. L. L. Lenz, S. Mohammadi, A. Geissler, D. A. Portnoy, SecA2-dependent secretion of autolytic enzymes promotes Listeria monocytogenes pathogenesis. *Proc National Acad Sci*. **100**, 12432–12437 (2003).

66. T. G. Senkevich, L. S. Wyatt, A. S. Weisberg, E. V. Koonin, B. Moss, A conserved poxvirus NlpC/P60 superfamily protein contributes to vaccinia virus virulence in mice but not to replication in cell culture. *Virology*. **374**, 506–14 (2008).

67. G. Parthasarathy, S. Lun, H. Guo, N. C. Ammerman, D. E. Geiman, W. R. Bishai, Rv2190c, an NlpC/P60 family protein, is required for full virulence of Mycobacterium tuberculosis. *Plos One*. **7**, e43429 (2012).

68. V. A. Pedicord, A. A. Lockhart, K. J. Rangan, J. W. Craig, J. Loschko, A. Rogoz, H. C. Hang, D. Mucida, Exploiting a host-commensal interaction to promote intestinal barrier function and enteric pathogen tolerance. *Science Immunology*. **1** (2016), doi:10.1126/sciimmunol.aai7732.

69. K. J. Rangan, V. A. Pedicord, Y.-C. Wang, B. Kim, Y. Lu, S. Shaham, D. Mucida, H. C. Hang, A secreted bacterial peptidoglycan hydrolase enhances tolerance to enteric pathogens. *Science*. **353**, 1434–1437 (2016).

70. M. Kuhn, W. Goebel, Identification of an extracellular protein of Listeria monocytogenes possibly involved in intracellular uptake by mammalian cells. *Infect Immun*. **57**, 55–61 (1989).

71. R. L. Schmidt, H. C. Filak, J. D. Lemon, T. A. Potter, L. L. Lenz, A LysM and SH3-Domain Containing Region of the Listeria monocytogenes p60 Protein Stimulates Accessory Cells to Promote Activation of Host NK Cells. *Plos Pathog*. **7**, e1002368 (2011).

72. L.-Y. Gao, M. Pak, R. Kish, K. Kajihara, E. J. Brown, A Mycobacterial Operon Essential for Virulence In Vivo and Invasion and Intracellular Persistence in Macrophages. *Infect Immun*. **74**, 1757–1767 (2006).

73. A. E. Sears, K. Palczewski, Lecithin:Retinol Acyltransferase: A Key Enzyme Involved in the Retinoid (visual) Cycle. *Biochemistry-us*. **55**, 3082–3091 (2016).

74. K. A. Estes, R. Kalamegham, W. Hanna-Rose, Membrane localization of the NlpC/P60 family protein EGL-26 correlates with regulation of vulval cell morphogenesis in Caenorhabditis elegans. *Dev Biol*. **308**, 196–205 (2007).

75. X. Ren, J. Lin, C. Jin, B. Xia, Solution structure of the N-terminal catalytic domain of human H-REV107--a novel circular permutated NlpC/P60 domain. *Febs Lett*. **584**, 4222–6 (2010).

76. V. H. Tang, G. A. Stewart, B. J. Chang, House dust mites possess a polymorphic, single domain putative peptidoglycan d,l endopeptidase belonging to the NlpC/P60 Superfamily. *Febs Open Bio*. **5**, 813–23 (2015).

77. V. H. Tang, G. A. Stewart, B. J. Chang, Dermatophagoides pteronyssinus lytFM encoding an NlpC/P60 endopeptidase is also present in mite-associated bacteria that express LytFM variants. *Febs Open Bio*. **7**, 1267–1280 (2017).

78. L. T. Mathaba, C. H. Pope, J. Lenzo, M. Hartofillis, H. Peake, R. L. Moritz, R. J. Simpson, A. Bubert, P. J. Thompson, G. A. Stewart, Isolation and characterisation of a 13.8-kDa bacteriolytic enzyme from house dust mite extracts: homology with prokaryotic proteins suggests that the enzyme could be bacterially derived. *Fems Immunol Medical Microbiol*. **33**, 77–88 (2002).

79. V. H. Tang, B. J. Chang, A. Srinivasan, L. T. Mathaba, G. B. Harnett, G. A. Stewart, Skin-associated Bacillus, staphylococcal and micrococcal species from the house dust mite, Dermatophagoides pteronyssinus and bacteriolytic enzymes. *Exp Appl Acarol*. **61**, 431–447 (2013).

80. C. Manichanh, N. Borruel, F. Casellas, F. Guarner, *Nature Reviews Gastroenterology and Hepatology*, in press, doi:10.1038/nrgastro.2012.152.

81. G. M. de Waal, W. J. S. de Villiers, T. Forgan, T. Roberts, E. Pretorius, Colorectal cancer is associated with increased circulating lipopolysaccharide, inflammation and hypercoagulability. *Sci Rep-uk*. **10**, 8777 (2020).

82. H. Zhuang, L. Cheng, Y. Wang, Y.-K. Zhang, M.-F. Zhao, G.-D. Liang, M.-C. Zhang, Y.-G. Li, J.-B. Zhao, Y.-N. Gao, Y.-J. Zhou, S.-L. Liu, Dysbiosis of the Gut Microbiome in Lung Cancer. *Front Cell Infect Mi*. **9**, 112 (2019).

83. M. T. Sorbara, D. J. Philpott, Peptidoglycan: a critical activator of the mammalian immune system during infection and homeostasis. *Immunological Reviews*. **243**, 40–60 (2011).

84. J. Royet, R. Dziarski, Peptidoglycan recognition proteins: pleiotropic sensors and effectors of antimicrobial defences. *Nature Reviews Microbiology*. **5**, nrmicro1620 (2007).

85. M. A. Boudreau, J. F. Fisher, S. Mobashery, Messenger Functions of the Bacterial Cell Wall-derived Muropeptides. *Biochemistry*. **51**, 2974–2990 (2012).

86. A. J. Wolf, D. M. Underhill, Peptidoglycan recognition by the innate immune system. *Nature reviews. Immunology* (2018), doi:10.1038/nri.2017.136.

87. M. E. Griffin, C. W. Hespen, Y. Wang, H. C. Hang, Translation of peptidoglycan metabolites into immunotherapeutics. *Clin Transl Immunol*. **8**, e1095 (2019).

88. T. J. Wyckoff, J. A. Taylor, N. R. Salama, Beyond growth: novel functions for bacterial cell wall hydrolases. *Trends in Microbiology*. **20**, 540–547 (2012).

89. S. G. Bartual, D. Straume, G. Stamsås, I. G. Muñoz, C. Alfonso, M. Martínez-Ripoll, L. Håvarstein, J. A. Hermoso, Structural basis of PcsB-mediated cell separation in Streptococcus pneumoniae. *Nature Communications*. **5**, 3842 (2014).

90. I. J. J. Claes, G. Schoofs, K. Regulski, P. Courtin, M.-P. Chapot-Chartier, T. Rolain, P. Hols, I. von Ossowski, J. Reunanen, W. M. de Vos, A. Palva, J. Vanderleyden, S. C. J. D.

Keersmaecker, S. Lebeer, Genetic and Biochemical Characterization of the Cell Wall Hydrolase Activity of the Major Secreted Protein of Lactobacillus rhamnosus GG. *Plos One*. **7**, e31588 (2012).

91. T. Fukushima, N. Uchida, M. Ide, T. Kodama, J. Sekiguchi, DL-endopeptidases function as both cell wall hydrolases and poly-γ-glutamic acid hydrolases. *Microbiology*. **164**, 277–286 (2018).

92. B. Glauner, J. V. Holtje, U. Schwarz, The composition of the murein of Escherichia coli. *The Journal of Biological Chemistry*. **263** (1988).

93. R. S. Rosenthal, R. Dziarski, [20] Isolation of peptidoglycan and soluble peptidoglycan fragments. *Methods in Enzymology*. **235**, 253–285 (1994).

94. S. M. Desmarais, M. A. D. Pedro, F. Cava, K. C. Huang, Peptidoglycan at its peaks: how chromatographic analyses can reveal bacterial cell wall structure and assembly. *Mol Microbiol*. **89**, 1–13 (2013).

95. S. M. Desmarais, F. Cava, M. A. de Pedro, K. Huang, Isolation and preparation of bacterial cell walls for compositional analysis by ultra performance liquid chromatography. *Journal of visualized experiments : JoVE*, e51183 (2014).

96. D. Kühner, M. Stahl, D. D. Demircioglu, U. Bertsche, From cells to muropeptide structures in 24 h: Peptidoglycan mapping by UPLC-MS. *Scientific Reports*. **4**, 7494 (2014).

97. P. Jackson, The use of polyacrylamide-gel electrophoresis for the high-resolution separation of reducing saccharides labelled with the fluorophore 8-aminonaphthalene-1,3,6-trisulphonic acid. Detection of picomolar quantities by an imaging system based on a cooled charge-coupled device. *Biochem J*. **270**, 705–713 (1990).

98. G. Mugunthan, D. Sriram, P. Yogeeswari, K. P. R. Kartha, Synthesis and biological evaluation of sugar-derived chiral nitroimidazoles as potential antimycobacterial agents. *Carbohyd Res*. **346**, 1760–1766 (2011).

99. F. Teng, M. Kawalec, G. M. Weinstock, W. Hryniewicz, B. E. Murray, An Enterococcus faecium Secreted Antigen, SagA, Exhibits Broad-Spectrum Binding to Extracellular Matrix Proteins and Appears Essential for E. faecium Growth. *Infect Immun*. **71**, 5033–5041 (2003).

100. G. J. Patti, S. J. Kim, J. Schaefer, Characterization of the Peptidoglycan of Vancomycin-Susceptible Enterococcus faecium †. *Biochemistry-us*. **47**, 8378–8385 (2008).

101. G. J. Patti, J. Chen, M. L. Gross, Method Revealing Bacterial Cell-Wall Architecture by Time-Dependent Isotope Labeling and Quantitative Liquid Chromatography/Mass Spectrometry. *Analytical Chemistry*. **81**, 2437–2445 (2009).

102. G. J. Patti, J. Chen, J. Schaefer, M. L. Gross, Characterization of Structural Variations in the Peptidoglycan of Vancomycin-Susceptible Enterococcus faecium: Understanding Glycopeptide–Antibiotic Binding Sites Using Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*. **19**, 1467–1475 (2008).

103. M. F. L. Parker, R. R. Flavell, J. Luu, O. S. Rosenberg, M. A. Ohliger, D. M. Wilson, Small molecule sensors targeting the bacterial cell wall. *Acs Infect Dis* (2020), doi:10.1021/acsinfecdis.9b00515.

104. S. R. Adams, R. E. Campbell, L. A. Gross, B. R. Martin, G. K. Walkup, Y. Yao, J. Llopis, R. Y. Tsien, New Biarsenical Ligands and Tetracysteine Motifs for Protein Labeling in Vitro and in Vivo: Synthesis and Biological Applications. *J Am Chem Soc*. **124**, 6063–6076 (2002).

105. M. F. Copeland, S. T. Flickinger, H. H. Tuson, D. B. Weibel, Studying the Dynamics of Flagella in Multicellular Communities of Escherichia coli by Using Biarsenical Dyes. *Appl Environ Microb*. **76**, 1241–1250 (2009).

106. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. **10**, 845–858 (2015).

107. W.-L. Ng, K. M. Kazmierczak, M. E. Winkler, Defective cell wall synthesis in Streptococcus pneumoniae R6 depleted for the essential PcsB putative murein hydrolase or the VicR (YycF) response regulator: Defective cell wall synthesis in pcsB mutants. *Mol Microbiol*. **53**, 1161–1175 (2004).

108. T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J Med Chem*. **47**, 1750–1759 (2004).

109. R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J Med Chem*. **47**, 1739–1749 (2004).

110. A. Emirian, S. Fromentin, C. Eckert, F. Chau, L. Dubost, M. Delepierre, L. Gutmann, M. Arthur, S. Mesnage, Impact of peptidoglycan O-acetylation on autolytic activities of the Enterococcus faecalis N-acetylglucosaminidase AtlA and N-acetylmuramidase AtlB. *FEBS Letters*. **583**, 3033–3038 (2009).

111. F. Lebreton, A. L. Manson, J. T. Saavedra, T. J. Straub, A. M. Earl, M. S. Gilmore, Tracing the Enterococci from Paleozoic Origins to the Hospital. *Cell*. **169**, 849-861.e13 (2017).

112. G. S. Chilambi, H. R. Nordstrom, D. R. Evans, J. A. Ferrolino, R. T. Hayden, G. M. Marón, A. N. Vo, M. S. Gilmore, J. Wolf, J. W. Rosch, D. V. Tyne, Evolution of vancomycin-resistant Enterococcus faecium during colonization and infection in immunocompromised pediatric patients. *Proc National Acad Sci*. **117**, 11703–11714 (2020).

113. H. Hanchi, W. Mottawea, K. Sebei, R. Hammami, The Genus Enterococcus: Between Probiotic Potential and Safety Concerns—An Update. *Front Microbiol*. **9**, 1791 (2018).

114. S. M. Vieira, M. Hiltensperger, V. Kumar, D. Zegarra-Ruiz, C. Dehner, N. Khan, F. R. C. Costa, E. Tiniakou, T. Greiling, W. Ruff, A. Barbieri, C. Kriegel, S. S. Mehta, J. R. Knight, D. Jain, A. L. Goodman, M. A. Kriegel, Translocation of a gut pathobiont drives autoimmunity in mice and humans. *Science*. **359**, 1156–1161 (2018).

115. C. K. Stein-Thoeringer, K. B. Nichols, A. Lazrak, M. D. Docampo, A. E. Slingerland, J. B. Slingerland, A. G. Clurman, G. Armijo, A. L. C. Gomes, Y. Shono, A. Staffas, M. B. da Silva, S. M. Devlin, K. A. Markey, D. Bajic, R. Pinedo, A. Tsakmaklis, E. R. Littmann, A. Pastore, Y. Taur, S. Monette, M. E. Arcila, A. J. Pickard, M. Maloy, R. J. Wright, L. A. Amoretti, E. Fontana, D. Pham, M. A. Jamal, D. Weber, A. D. Sung, D. Hashimoto, C. Scheid, J. B. Xavier, J. A. Messina, K. Romero, M. Lew, A. Bush, L. Bohannon, K. Hayasaka, Y. Hasegawa, M. J. G. T. Vehreschild, J. R. Cross, D. M. Ponce, M. A. Perales, S. A. Giralt, R. R. Jenq, T. Teshima, E. Holler, N. J. Chao, E. G. Pamer, J. U. Peled, M. R. M. van den Brink, Lactose drives Enterococcus expansion to promote graft-versus-host disease. *Science*. **366**, 1143–1149 (2019).

116. A. Ribas, J. D. Wolchok, Cancer immunotherapy using checkpoint blockade. *Science*. **359**, 1350–1355 (2018).

117. B. Routy, E. L. Chatelier, L. Derosa, C. P. M. Duong, M. T. Alou, R. Daillère, A. Fluckiger, M. Messaoudene, C. Rauber, M. P. Roberti, M. Fidelle, C. Flament, V. Poirier-Colame, P. Opolon, C. Klein, K. Iribarren, L. Mondragón, N. Jacquelot, B. Qu, G. Ferrere, C. Clémenson, L. Mezquita, J. R. Masip, C. Naltet, S. Brosseau, C. Kaderbhai, C. Richard, H. Rizvi, F. Levenez, N. Galleron, B. Quinquis, N. Pons, B. Ryffel, V. Minard-Colin, P. Gonin, J.-C. Soria, E. Deutsch, Y. Loriot, F. Ghiringhelli, G. Zalcman, F. Goldwasser, B. Escudier, M. D. Hellmann, A. Eggermont, D. Raoult, L. Albiges, G. Kroemer, L. Zitvogel, Gut microbiome influences efficacy of PD-1–based immunotherapy against epithelial tumors. *Science*. **359**, 91–97 (2017).

118. V. Matson, J. Fessler, R. Bao, T. Chongsuwat, Y. Zha, M.-L. Alegre, J. J. Luke, T. F. Gajewski, The commensal microbiome is associated with anti–PD-1 efficacy in metastatic melanoma patients. *Science*. **359**, 104–108 (2018).

119. J. A. Mohamed, F. Teng, S. R. Nallapareddy, B. E. Murray, Pleiotrophic Effects of 2 Enterococcus faecalis sagA –Like Genes, salA and salB, Which Encode Proteins That Are Antigenic during Human Infection, on Biofilm Formation and Binding to Collagen Type I and Fibronectin. *J Infect Dis*. **193**, 231–240 (2006).

120. C. A. Arias, B. E. Murray, The rise of the Enterococcus: beyond vancomycin resistance. *Nat Rev Microbiol*. **10**, 266–78 (2012).

121. W. Gao, B. P. Howden, T. P. Stinear, Evolution of virulence in Enterococcus faecium, a hospital-adapted opportunistic pathogen. *Current Opinion in Microbiology*. **41**, 76–82 (2018).

122. M. García-Solache, L. B. Rice, The Enterococcus: a Model of Adaptability to Its Environment. *Clin Microbiol Rev*. **32** (2019), doi:10.1128/cmr.00058-18.

123. J. M. Wells, A. Mercenier, Mucosal delivery of therapeutic and prophylactic molecules using lactic acid bacteria. *Nature Reviews Microbiology*. **6**, nrmicro1840 (2008).

124. K. J. Chua, W. C. Kwok, N. Aggarwal, T. Sun, M. W. Chang, Designer probiotics for the prevention and treatment of human diseases. *Curr Opin Chem Biol*. **40**, 8–16 (2017).

125. P. A. Bron, M. Kleerebezem, Lactic Acid Bacteria for Delivery of Endogenous or Engineered Therapeutic Molecules. *Front Microbiol*. **9**, 1821 (2018).

126. E. Davis, L. Hicks, I. Ali, E. Salzman, J. Wang, E. Snitkin, K. Gibson, M. Cassone, L. Mody, B. Foxman, Epidemiology of Vancomycin-Resistant Enterococcus faecium and Enterococcus faecalis Colonization in Nursing Facilities. *Open Forum Infect Dis*. **7**, ofz553 (2020).

127. M. J. Rybak, B. J. McGrath, Combination Antimicrobial Therapy for Bacterial Infections. *Drugs*. **52**, 390–405 (1996).

128. M. Beganovic, M. K. Luther, L. B. Rice, C. A. Arias, M. J. Rybak, K. L. LaPlante, A Review of Combination Antimicrobial Therapy for Enterococcus Faecalis Bloodstream Infections and Infective Endocarditis. *Clin Infect Dis Official Publ Infect Dis Soc Am*. **67**, 303–309 (2018).

129. M. J. Matar, J. Tarrand, I. Raad, K. V. I. Rolston, Colonization and infection with vancomycin-resistant enterococcus among patients with cancer. *Am J Infect Control*. **34**, 534–536 (2006).

130. R. T. Schooley, B. Biswas, J. J. Gill, A. Hernandez-Morales, J. Lancaster, L. Lessor, J. J. Barr, S. L. Reed, F. Rohwer, S. Benler, A. M. Segall, R. Taplitz, D. M. Smith, K. Kerr, M. Kumaraswamy, V. Nizet, L. Lin, M. D. McCauley, S. A. Strathdee, C. A. Benson, R. K. Pope, B. M. Leroux, A. C. Picel, A. J. Mateczun, K. E. Cilwa, J. M. Regeimbal, L. A. Estrella, D. M. Wolfe, M. S. Henry, J. Quinones, S. Salka, K. A. Bishop-Lilly, R. Young, T. Hamilton, Development and Use of Personalized Bacteriophage-Based Therapeutic Cocktails To Treat a Patient with a Disseminated Resistant Acinetobacter baumannii Infection. *Antimicrob Agents Ch*. **61**, e00954-17 (2017).

131. B. K. Chan, P. E. Turner, S. Kim, H. R. Mojibian, J. A. Elefteriades, D. Narayan, Phage treatment of an aortic graft infected with Pseudomonas aeruginosa. *Evol Medicine Public Heal*. **2018**, 60–66 (2018).

132. R. Fish, E. Kutter, D. Bryan, G. Wheat, S. Kuhl, Resolving Digital Staphylococcal Osteomyelitis Using Bacteriophage—A Case Report. *Antibiotics*. **7**, 87 (2018).

133. E. J. Cano, K. M. Caflisch, P. L. Bollyky, J. D. V. Belleghem, R. Patel, J. Fackler, M. J. Brownstein, B. Horne, B. Biswas, M. Henry, F. Malagon, D. G. Lewallen, G. A. Suh, Phage Therapy for Limb-threatening Prosthetic Knee Klebsiella pneumoniae Infection: Case Report and In Vitro Characterization of Anti-biofilm Activity. *Clin Infect Dis* (2020), doi:10.1093/cid/ciaa705.

134. T. Morrisette, R. Kebriaei, K. L. Lev, S. Morales, M. J. Rybak, Bacteriophage Therapeutics: A Primer for Clinicians on Phage-Antibiotic Combinations. *Pharmacotherapy*. **40**, 153–168 (2019).

135. G. S. Canfield, A. Chatterjee, J. Espinosa, M. R. Mangalea, E. K. Sheriff, M. Keidan, S. W. McBride, B. D. McCollister, H. C. Hang, B. A. Duerkop, Lytic bacteriophages facilitate antibiotic sensitization of Enterococcus faecium. *Antimicrob Agents Ch* (2021), doi:10.1128/aac.00143-21.

136. K. L. Palmer, P. Godfrey, A. Griggs, V. N. Kos, J. Zucker, C. Desjardins, G. Cerqueira, D. Gevers, S. Walker, J. Wortman, M. Feldgarden, B. Haas, B. Birren, M. S. Gilmore, Comparative Genomics of Enterococci: Variation in Enterococcus faecalis, Clade Structure in E. faecium, and Defining Characteristics of E. gallinarum and E. casseliflavus. *Mbio*. **3**, e00318-11 (2012).

137. M. de Been, W. van Schaik, L. Cheng, J. Corander, R. J. Willems, Recent Recombination Events in the Core Genome Are Associated with Adaptive Evolution in Enterococcus faecium. *Genome Biol Evol*. **5**, 1524–1535 (2013).

138. G. S. Canfield, A. Chatterjee, J. Espinosa, M. R. Mangalea, E. K. Sheriff, M. Keidan, S. W. McBride, B. D. McCollister, H. C. Hang, B. A. Duerkop, *Biorxiv*, in press, doi:10.1101/2020.09.22.309401.

139. V. Chen, M. G. Griffin, H. C. Hang, *Biorxiv*, in press, doi:10.1101/2020.09.01.278044.

140. S. R. Nallapareddy, K. V. Singh, B. E. Murray, Contribution of the Collagen Adhesin Acm to Pathogenesis of Enterococcus faecium in Experimental Endocarditis ▿ . *Infect Immun*. **76**, 4120–4128 (2008).

141. A. L. Mitchell, T. K. Attwood, P. C. Babbitt, M. Blum, P. Bork, A. Bridge, S. D. Brown, H.-Y. Chang, S. El-Gebali, M. I. Fraser, J. Gough, D. R. Haft, H. Huang, I. Letunic, R. Lopez, A. Luciani, F. Madeira, A. Marchler-Bauer, H. Mi, D. A. Natale, M. Necci, G. Nuka, C. Orengo, A. P. Pandurangan, T. Paysan-Lafosse, S. Pesseat, S. C. Potter, M. A. Qureshi, N. D. Rawlings, N. Redaschi, L. J. Richardson, C. Rivoire, G. A. Salazar, A. Sangrador-Vegas, C. J. A. Sigrist, I. Sillitoe, G. G. Sutton, N. Thanki, P. D. Thomas, S. C. E. Tosatto, S.-Y. Yong, R. D. Finn, InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res*. **47**, gky1100- (2018).

142. J. Guglielmini, L. Quintais, M. P. Garcillán-Barcia, F. de la Cruz, E. P. C. Rocha, The Repertoire of ICE in Prokaryotes Underscores the Unity, Diversity, and Ubiquity of Conjugation. *Plos Genet*. **7**, e1002222 (2011).

143. D. BILLOT-KLEIN, D. SHLAES, D. BRYANT, D. BELL, J. van HEIJENOORT, L. GUTMANN, Peptidoglycan structure of Enterococcus faecium expressing vancomycin resistance of the VanB type. *Biochemical Journal*. **313**, 711–715 (1996).

144. B. L. de Jonge, S. Handwerger, D. Gage, Altered peptidoglycan composition in vancomycin-resistant Enterococcus faecalis. *Antimicrob Agents Ch*. **40**, 863–869 (1996).

145. J. D. Chang, E. E. Foster, A. G. Wallace, S. Kim, Peptidoglycan O-acetylation increases in response to vancomycin treatment in vancomycin-resistant Enterococcus faecalis. *Scientific Reports*. **7**, 46500 (2017).

146. G. J. Patti, J. Chen, M. L. Gross, Method Revealing Bacterial Cell-Wall Architecture by Time-Dependent Isotope Labeling and Quantitative Liquid Chromatography/Mass Spectrometry. *Anal Chem*. **81**, 2437–2445 (2009).

147. K. Hullahalli, M. Rodrigues, K. L. Palmer, Exploiting CRISPR-Cas to manipulate Enterococcus faecalis populations. *Elife*. **6**, e26664 (2017).

148. D. Fenyo, Q. Wang, J. A. DeGrasse, J. C. Padovan, M. Cadene, B. T. Chait, MALDI sample preparation: the ultra thin layer method. *J Vis Exp Jove*, 192 (2007).

149. M. Cadene, B. T. Chait, A Robust, Detergent-Friendly Method for Mass Spectrometric Analysis of Integral Membrane Proteins. *Anal Chem*. **72**, 5655–5658 (2000).

150. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. **276**, 307–26 (1997).

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

152. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, N. Echols, J. J. Headd, L.-W. Hung, S. Jain, G. J. Kapral, R. W. G. Kunstleve, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, The Phenix software for automated determination of macromolecular structures. *Methods*. **55**, 94–106 (2011).

153. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr Sect D Biological Crystallogr*. **60**, 2126–2132 (2004).

154. V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sect D Biological Crystallogr*. **66**, 12–21 (2010).

155. P. D. Lyne, M. L. Lamb, J. C. Saeh, Accurate Prediction of the Relative Potencies of Members of a Series of Kinase Inhibitors Using Molecular Docking and MM-GBSA Scoring. *J Med Chem*. **49**, 4805–4808 (2006).

156. D. Arndt, J. R. Grant, A. Marcu, T. Sajed, A. Pon, Y. Liang, D. S. Wishart, PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res*. **44**, W16–W21 (2016).

157. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol*. **32**, 268–274 (2015).

158. J. Trifinopoulos, L.-T. Nguyen, A. von Haeseler, B. Q. Minh, W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res*. **44**, W232–W235 (2016).

159. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol*. **35**, 518–522 (2017).

160. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res*. **47**, gkz239- (2019).

161. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. **42**, W320–W324 (2014).