



Summer 2019

Progress Towards the Substrate-bound Structure of Streptococcus pneumoniae Sortase A

Nicholas M. Horvath

Western Washington University, nick.horvath8810@gmail.com

Follow this and additional works at: <https://cedar.wvu.edu/wwuet>

 Part of the [Chemistry Commons](#)

Recommended Citation

Horvath, Nicholas M., "Progress Towards the Substrate-bound Structure of Streptococcus pneumoniae Sortase A" (2019). *WWU Graduate School Collection*. 897.
<https://cedar.wvu.edu/wwuet/897>

This Masters Thesis is brought to you for free and open access by the WWU Graduate and Undergraduate Scholarship at Western CEDAR. It has been accepted for inclusion in WWU Graduate School Collection by an authorized administrator of Western CEDAR. For more information, please contact westerncedar@wwu.edu.

Progress Towards the Substrate-bound
Structure of *Streptococcus pneumoniae* Sortase A

By

Nicholas Horvath

Accepted in Partial Completion
of the Requirements for the Degree
Master of Science

ADVISORY COMMITTEE

Dr. John Antos, Chair

Dr. P. Clint Spiegel

Dr. James Vyvyan

GRADUATE SCHOOL

David L. Patrick, Interim Dean

Master's Thesis

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Western Washington University, I grant to Western Washington University the non-exclusive royalty-free right to archive, reproduce, distribute, and display the thesis in any and all forms, including electronic format, via any digital library mechanisms maintained by WWU.

I represent and warrant this is my original work, and does not infringe or violate any rights of others. I warrant that I have obtained written permissions from the owner of any third party copyrighted material included in these files.

I acknowledge that I retain ownership rights to the copyright of this work, including but not limited to the right to use all or part of this work in future works, such as articles or books.

Library users are granted permission for individual, research and non-commercial reproduction of this work for educational purposes only. Any further digital posting of this document requires specific permission from the author.

Any copying or publication of this thesis for commercial purposes, or for financial gain, is not allowed without my written permission.

Nicholas Horvath

8/8/19

Progress Towards the Substrate-bound
Structure of *Streptococcus pneumoniae* Sortase A

A Thesis
Presented to
The Faculty of
Western Washington University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

by
Nicholas Horvath
July 2019

Abstract

Sortases are cysteine transpeptidases found primarily on the cell surface of Gram-positive bacteria. Sortase-mediated ligations have become an attractive option for protein modification chemistry, enabling the synthesis of a wide range of non-natural polypeptide derivatives. Attempts at understanding how these enzymes recognize and bind substrates are integral to furthering their usefulness in protein engineering and, potentially, treatment of bacterial diseases. However, the variable substrate specificity and activity between homologs of these enzymes is not yet fully understood. Of specific interest to us is sortase A from *Streptococcus pneumoniae* (SrtA_{pneu}), as it demonstrates a broad substrate tolerance not observed in other sortase A homologs. Correspondingly, we have made advances towards characterizing a substrate bound structure of SrtA_{pneu} in an effort to further understand its unique substrate promiscuity, deviating from the canonical LPXTG sorting signal. Our strategy initially involved generating a non-cleavable peptide analog capable of docking into the active site, however, synthesis of a ketomethylene-linked dipeptide isostere and its insertion into a peptide via solid phase peptide synthesis proved to be more challenging than we anticipated. We revised our approach by designing a substrate harboring an LPACG sorting motif. Peptide preparations with a thiopyridine leaving group favorably facilitated disulfide bridging between the active site and sorting motif cysteines, allowing for elucidation of a SrtA_{pneu} structure displaying key interactions that allow the enzyme to recognize a wide-variety of substrates. To this end, we have utilized x-ray crystallography and solution NMR in an attempt to characterize SrtA_{pneu} with a bound substrate analog. Although we were unsuccessful, this work has established a foundation for future efforts toward determining the substrate-bound structure of SrtA_{pneu}.

Table of Contents

Abstract.....	iv
List of Figures and Tables.....	vi
List of Abbreviations and Acronyms.....	ix
Chapter 1 – Introduction	
1.1 Protein Engineering – Advantages of Chemoenzymatic Modification.....	1
1.2 Expanded Substrate Tolerance of Sortase Homologs.....	6
1.3 Substrate Binding and Structural Characteristics of SrtA Homologs.....	9
1.4 SrtA _{pneu} Enzyme Activity as a Function of Oligomeric State.....	14
1.5 Overview of Project Goals.....	18
Chapter 2 – Screening Studies for X-ray Crystallization of SrtA_{pneu}	
2.1 Preparation of SrtA _{pneu} for Crystallization.....	19
2.2 Summary of Crystal Screening Efforts.....	22
Chapter 3 – Preparation and NMR Characterization of ¹⁵N-labeled SrtA_{pneu}	
3.1 Preparation of unlabeled TEV-SrtA _{pneu} and 1D ¹ H-NMR Analysis.....	25
3.2 Expression and Purification of ¹⁵ N-labeled TEV-SrtA _{pneu}	30
3.3 2D-HSQC NMR analysis of ¹⁵ N-labeled TEV-SrtA _{pneu}	34
Chapter 4 – Progress Toward Preparation of Substrate-docked derivatives of SrtA_{pneu}	
4.1 – Design and Synthesis of ketomethylene isosteres.....	39
4.2 – Third generation design using disulfide linked analog.....	44
Chapter 5 – Conclusions and Future Directions.....	48
Chapter 6 – Experimental	
6.1 Expression of SrtA _{pneu}	53
6.2 Protein Crystal Preparations & Diffraction.....	57
6.3 NMR sample Preparation & Acquisition.....	59
6.4 Synthesis of ketomethylene isosteres.....	59
6.5 Peptide Synthesis & Analysis.....	62
6.6 Analysis of Enzyme Transpeptidation Activity.....	65
Chapter 7 – Literature Cited.....	65
Chapter 8 – Appendix.....	73

List of Figures and Tables

- Figure 1.** Generic Schematic of Chemoenzymatic Protein Labeling (page 2)
- Figure 2.** Overview of SrtA *in vivo* mechanism on surface of *Staphylococcus aureus* (page 4)
- Figure 3.** Protein modification using model SML (page 5)
- Figure 4.** Comparison of substrate preferences for the 5th position of the sorting motif among SrtA homologs (page 8)
- Figure 5.** (A) Solution NMR structure of SrtA_{staph}. (B) Predicted structure of SrtA_{pneu} based on a one-to-one threading model of SrtA_{pyogenes}. (C) Solution NMR structure of the SrtA_{staph} active site with a bound LPAT* substrate analog (page 11)
- Figure 6.** Sequence alignment of selected SrtA homologs (page 12)
- Figure 7.** (A) The 3D domain swapped dimer of SrtA_{pneu} (PDB ID: 4O8L). (B) A domain swapped monomer from the dimeric structure of SrtA_{pneu}. (C) Predicted structure of SrtA_{pneu} based on a one-to-one threading model of SrtA_{pyogenes} (page 13)
- Figure 8.** RP-HPLC analysis of model SML reactions demonstrating the difference in activity between monomeric and multimeric SrtA_{pneu} preparations (page 15)
- Figure 9.** A comparison of standard vs. refolded SrtA_{pneu} preparations via SE-FPLC (page 17)
- Figure 10.** SDS-PAGE analysis of a SrtA_{pneu} IMAC/refolding purification scheme (page 20)
- Figure 11.** SEC traces of IMAC purified, refolded SrtA_{pneu} (A) and collected monomeric fractions (B). A deconvolved mass spectrum (C) of SrtA_{pneu} generated from the (D) raw ESI-MS spectrum of the purified enzyme (page 21)
- Figure 12.** Crystals formed by screening conditions in Table 2 (page 22)
- Figure 13.** Crystal optimization as a function of pH and PEG 3350 (page 23)

Figure 14. (Left) SDS-PAGE analysis of non-denaturing IMAC purification of TEV-SrtA_{pneu} (page 26)

Figure 15. SEC traces of IMAC purified TEV-SrtA_{pneu} (A) and collected monomeric fractions (B). A deconvolved mass spectrum (C) of TEV-SrtA_{pneu} generated from the corresponding raw ESI-MS spectrum (D) (page 27)

Figure 16. RP-HPLC analysis of model SML reaction using TEV-SrtA_{pneu} (page 28)

Figure 17. 1D ¹H-NMR analysis of monomeric SrtA_{pneu}-TEV over the course of several days (page 29)

Figure 18. An SDS-PAGE analysis of ¹⁵N TEV-SrtA_{pneu} (page 30)

Figure 19. An SDS-PAGE analysis of ¹⁵N SrtA_{pneu}-TEV refolding IMAC purification scheme (page 31)

Figure 20. SEC-FPLC traces of IMAC purified ¹⁵N SrtA_{pneu}-TEV (A) and collected monomeric fractions (B). A deconvolved mass spectrum (C) of ¹⁵N SrtA_{pneu}-TEV generated from the corresponding raw ESI-MS spectrum (D) (page 32)

Figure 21. RP-HPLC analysis of model SML reaction using ¹⁵N SrtA_{pneu}-TEV (page 33)

Figure 22. 2D HSQC NMR acquisitions of ¹⁵N SrtA_{pneu}-TEV (98 μM) at RT and 15 °C (page 35)

Figure 23. 2D HSQC NMR acquisition of ¹⁵N SrtA_{pneu}-TEV (225 μM) at RT (page 36)

Figure 24. 2D HSQC NMR acquisition of ¹⁵N SrtA_{pneu} (171 μM) at RT (page 37)

Figure 25. 2D HSQC NMR acquisition of TEV-cleaved ¹⁵N SrtA_{pneu} (80.5 μM) at RT (page 38)

Figure 26. Structural comparison of a model LPATG substrate to a 1st generation inhibitor substrate (page 40)

Figure 27. Structural comparison of 1st generation and 2nd generation inhibitor substrates. (page 40)

Figure 28. Overview of the proposed synthetic scheme based on procedures from Budnjo et al. and Mathieu et al. using a Boc-protected amino acid starting material (page 41)

Figure 29. Reaction scheme for synthesizing a G[keto]G building block for SPPS (page 43)

Figure 30. Synthesis of 2nd generation ketomethylene substrate analog (page 43)

Figure 31. Structural comparison of a model LPATG substrate to a 3rd generation inhibitor substrate (page 44)

Figure 32. RP-HPLC/LC-ESI-MS analysis for determining the purity and molecular weight of Bz-GLPACGG starting material and modified peptide product (page 46)

Figure 33. Deconvolved mass spectra of unmodified SrtA_{pneu} (A) and substrate bound enzyme (B) (page 47)

Table 1. Substrate specificity of Sortase A mutants and wild-type homologs (page 6)

Table 2. Preliminary screening conditions resulting in crystal formation (page 22)

Table 3. Diffracted crystals formed by conditions A-E (page 24)

Table 4. Summary of Boc-Ketomethylene reaction optimization results (page 42)

Table 5. Reaction conditions for SML (page 65)

List of Abbreviations and Acronyms

Abz	2-Aminobenzoyl
Boc	<i>tert</i> -Butyloxycarbonyl
CDI	Carbonyldiimidazole
DCM	Dichloromethane
DIPEA	Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMSO	Dimethylsulfoxide
Dnp	2,4-Dinitrophenol
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ESI-MS	Electrospray ionization mass spectrometry
EtOAc	Ethyl acetate
FGE	Formylglycine generating enzyme
Fmoc	Fluorenylmethyloxycarbonyl
FPLC	Fast protein liquid chromatography
HBTU	O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluroniumhexafluor
IPTG	Isopropyl- β -D-1-thiogalactopyranoside
LiHMDS	Lithium bis(trimethylsilyl)amide
MBHA	4-Methylbenzhydramine
MeCN	Acetonitrile
Ni-NTA	Nickel nitriloacetic acid
NMP	N-methyl-2-pyrrolidone

NMR	Nuclear magnetic resonance
OSu	N-hydroxysuccinimide ester
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
RP	Reverse phase
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SML	Sortase mediated ligation
SPPS	Solid-phase peptide synthesis
SrtA _{anth}	Bacillus anthracis Sortase A
SrtA _{pneu}	Streptococcus pneumoniae Sortase A
SrtA _{pyogenes}	Streptococcus pyogenes Sortase A
SrtA _{staph}	Staphylococcus aureus Sortase A
<i>t</i> -Bu	<i>tert</i> -Butyl
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
Tris	Tris(hydroxymethyl)aminomethane
UV/Vis	Ultraviolet/visible spectroscopy

Chapter 1 – Introduction

1.1 Protein Engineering - Advantages of Chemoenzymatic Modification

Contemporary pursuits toward endowing proteins with unnatural or non-canonical functionalities, known as protein engineering, has garnered significant attention across academic, industrial and medical applications.¹⁻⁴ Protein engineering efforts were historically limited to molecular biology techniques to install protein modifications genetically, which were applied through single-point mutations as well as frameshift sequence insertions and deletions.⁵ The modern utility of mutagenic techniques, such as directed evolution and unnatural amino acid incorporation, have dramatically broadened the variety of protein modifications, however, these methods continue to be burdened by substantial time and cost investments.⁶⁻¹⁰ In contrast, protein modification through novel direct chemical ligation strategies is a relatively quick and cost-effective way of engineering proteins. This bioconjugation technique takes advantage of the natural reactivity of sterically unencumbered amino acid side chains (i.e. lysine, cysteine, glutamic and aspartic acids), which is optimal for generating proteins with non-natural modifications.¹¹⁻¹⁵ While these contemporary modification strategies have revolutionized protein engineering, continuing to expand the scope and efficacy of direct protein bioconjugation is critical for the advancement of several fields, including fundamental biochemistry, the design of protein therapeutics, and the generation of new biomaterials.

Chemoenzymatic modification of proteins has provided an attractive alternative to site-directed mutagenesis and direct chemical modification strategies. This bioconjugation technique has been utilized in a variety of instances, including the production of fluorescently labeled proteins for live-cell trafficking, antibody-drug conjugates for site-specific payload delivery, and

adhering protein to nanoparticles.^{16–18} In chemoenzymatic modification, a sequence of amino acids is recognized by the modification enzyme, which results in the site-specific attachment of the desired moiety (**Figure 1**). If an endogenous protein target does not possess the required recognition sequence, which is often the case, then a recognition site must be added, typically using site-directed mutagenesis.¹⁹ There is a continuously expanding assortment of modifications that can be installed using chemoenzymatic modification strategies, along with a growing catalog of enzymes able to catalyze these processes.

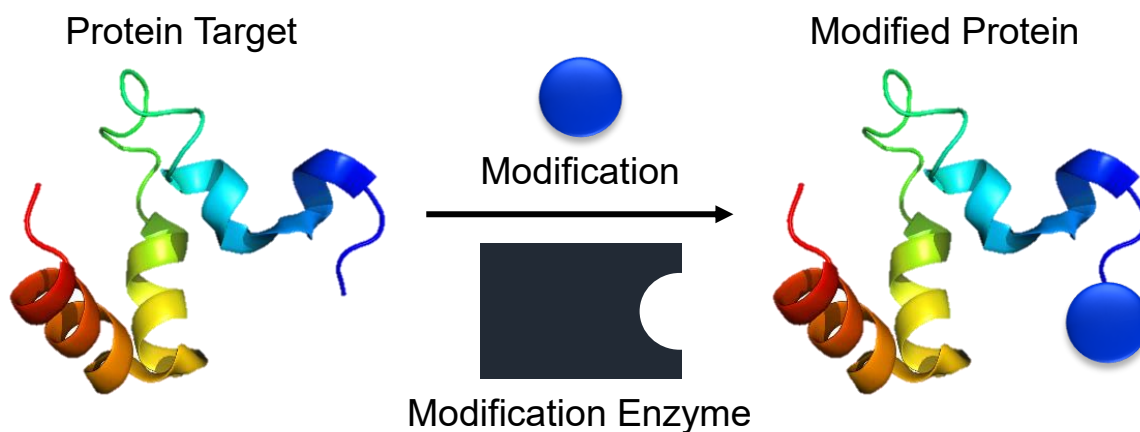


Figure 3. Generic Schematic of Chemoenzymatic Protein Labeling. A protein (left) harboring a binding motif is recognized by the modification enzyme, then the modification is covalently attached to the protein target (right).

A notable example of an enzyme used in chemoenzymatic strategies is formylglycine generating enzyme (FGE), which recognizes a CXPXR sequence of amino acids, then modifies the cysteine residue to a formylglycine reaction handle commonly utilized for generating stable oxime ligation products.^{20–22} Lipoic acid ligase has seen use through similar bioorthogonal ligation approaches, wherein this enzyme canonically functions to adhere lipoic acid to the ϵ -amine of lysine side chains within the primary sequence of its target protein.²³ Interestingly, this promiscuous enzyme has demonstrated the ability to install a diverse variety of substrates, notably azide and alkyne containing click handles, which has significantly broadened the scope of site-

specific modifications that can be installed using lipoic acid ligase.²⁴ Biotin ligase manifests a similar behavior, where it covalently attaches a biotin residue to the ϵ -amine of the lysine side chain within its target recognition sequence.²⁵ Biotin-based protein engineering has attracted attention due to its robust function as a site-specific tag for binding streptavidin or avidin containing biomolecules, including functionalized nano-particles or quantum dots, with exceptional specificity and pico-molar affinity.^{26,27}

The sortase enzyme family has also been extensively studied for its utility in protein modification. Sortases are endogenous to Gram-positive bacteria, where they function as transpeptidases through a catalytic mechanism involving a nucleophilic cysteine within the enzyme's active site.²⁸⁻³⁰ Sortases are separated into distinct classes (A-F) based on their unique contrasting structural and biochemical traits.^{31,32} Class A sortases (SrtA) have demonstrated the most relevance to protein engineering, having been recombinantly expressed with a truncated N-terminus to remove the transmembrane domain, which has resulted in a soluble derivative of SrtA for *in vitro* ligation reactions. *In vivo*, SrtA performs an essential "housekeeping" role in maintaining the extracellular environment by anchoring a variety of proteins to the cell wall.³³⁻³⁶ Proteins appended to the extracellular matrix by SrtA are key virulence factors, including collagen adhesion proteins as well as fibronectin and immunoglobulin binding proteins, that are responsible for bacterial cell colonization and evading host immune detection.³⁷⁻⁴² The *in vivo* function of SrtA has been highlighted as a viable drug target in Gram-positive bacteria, as studies have reported a dramatically reduced virulence of SrtA knockout Gram-positive bacterial strains.⁴³⁻⁴⁷ As the catalog of ineffective antibiotic drugs continues to rapidly expand, it has become imperative to develop an in-depth understanding of sortase structure and enzymology to further the development

of novel therapeutics while simultaneously providing insight into their role in protein engineering.⁴⁸

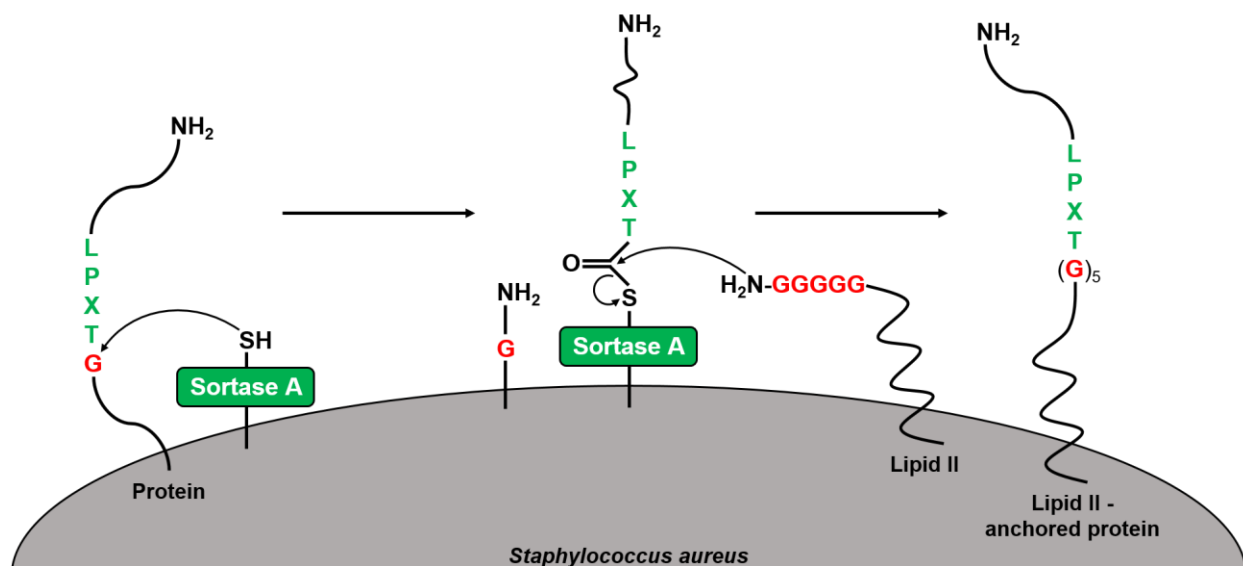


Figure 4. Overview of SrtA *in vivo* mechanism on surface of *Staphylococcus aureus*.

Sortase A enzymes share a common mechanism of action involving the recognition of a five amino acid sequence, which will hereafter be referred to as a “sorting motif”. The most common sorting motif for SrtA enzymes is the LPXTG sequence, where X is any amino acid.^{31,32} However, it is now known that different SrtA homologs can recognize a number of variations of the standard LPXTG sequence. Sortase A from *Staphylococcus aureus* (SrtA_{staph}) recognizes a protein substrate harboring an LPXTG sorting motif (**Figure 2**).³⁶ Next, the active site cysteine cleaves the amide bond between threonine and glycine, which releases the excised C-terminal fragment from the substrate. A transient acyl enzyme intermediate is formed through this process, and the scissile thioester linkage is subsequently intercepted by nucleophilic attack of the N-terminal amine of a pentaglycine peptide of lipid II anchored to the peptidoglycan matrix. Reconstitution of the amide bond linkage fuses the protein to the peptidoglycan, at which point the protein substrate is released from the active site and the enzymatic potency of SrtA_{staph} is restored for additional catalytic cycles.

In vitro, SrtA enzymes have been utilized extensively in protein engineering chemistry due to their ability to catalyze site-specific modifications at the sorting motif. Recent efforts that have used this approach include conjugating proteins and peptides to fluorophores, nanoparticle solid-supports, synthetic peptides, surfaces of live cells and other proteins (**Figure 3**).^{49–53}

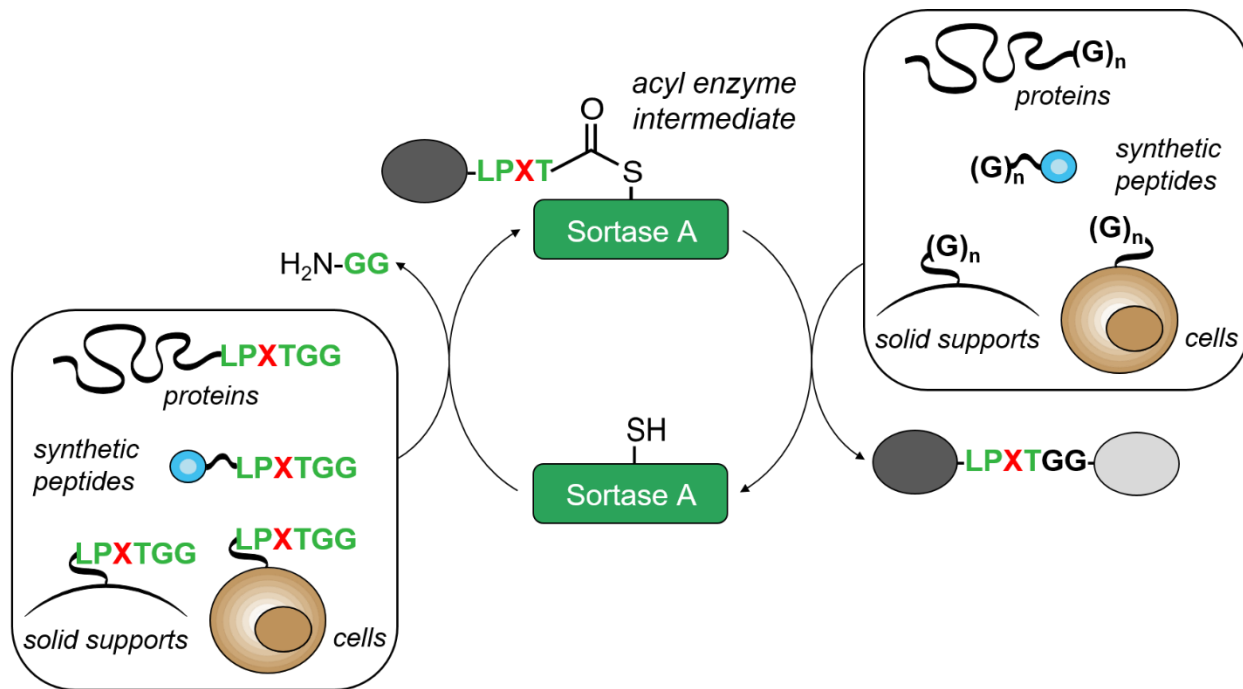


Figure 3. Protein modification using model SML.

This by no means encapsulates the full breadth of modifications that can be achieved using sortase-mediated ligation (SML), and we refer the reader to other excellent reviews for more comprehensive discussions of SML applications.^{54–58} One of the key factors in the versatility of SML is the ability of users to control which reaction partner is functioning as the LPXTG-substrate and which is serving as the reaction nucleophile.^{59–61} In doing so, one is able to use SML for appending modifications to exposed C- and N-termini, and in some cases sterically unencumbered secondary structures. To date, the majority of these SML studies have utilized wild-type SrtA_{staph}. However, over the past decade a number of efforts to improve the properties of SrtA_{staph} have

resulted in evolved mutants demonstrating improved reaction rates, non-canonical substrate tolerances, and Ca²⁺ cofactor independence.^{62–65} These evolved variants now provide a range of sortase derivatives that can be selected for specific applications.

1.2 Expanded Substrate Tolerance of Sortase Homologs

As noted above, there exists a diverse archive of published protein engineering applications utilizing sortase-mediated ligation, which is continually expanding along with contemporary efforts to circumvent limitations historically associated with this technique. Notably, issues associated with SML include the slow reactions rates of ligations using SrtA_{staph}, reaction reversibility, strict substrate specificity, and a narrow scope of compatible amine nucleophiles.^{66–}

⁶⁸ Of relevance to this thesis, there have now been reported multiple studies on expanding the substrate scope of SML using either SrtA_{staph} mutants or other naturally occurring sortase homologs.^{63,64,69–71}

Table 2. Substrate specificity of Sortase A mutants and wild-type homologs.

Species	Motif
Staphylococcus aureus (Wild-Type)	LPXXG
Staphylococcus aureus (Mutant)	XPKTG, LAXTG, LPEXG, LPXSG, FPXTG, APXTG
Streptococcus pyogenes	LPXTG, LPXTA, LPXLG
Streptococcus pneumoniae	LPATX

Bioinformatic investigations of SrtA homolog specificities through the CW-PRED2 genome alignment algorithm has revealed a universal preference for LPXTG motifs.^{65,72} *In vitro* analysis of SrtA preferences with computationally derived peptide substrates have revealed

discrepancies between actual and predicted SML compatible sorting sequences (**Table 1**). Notably, Kruger et al. have experimentally demonstrated that SrtA_{staph} tolerates LPXXG substrates, and exhibits a preference for glycine in the 6th position, outside of the canonical LPXTG sorting motif.⁷³ Phage and yeast display directed evolution studies have generated evolved variants of SrtA_{staph} with alternative substrate preferences.^{64,69} These SrtA_{staph} mutants have exhibited a relaxed substrate tolerance for residues in 1st, 2nd and 4th position.

Randomization of the $\beta 6/\beta 7$ loop among SrtA_{staph} mutants revealed evolved variants selective for FPXTG or APXTG motifs.⁶⁴ Rather than genetically modifying the substrate preferences of SrtA_{staph}, others have taken the approach of exploiting the natural reactivity and specificity of different SrtA homologs.^{70,71} The model enzyme for SML has historically been wild-type SrtA_{staph}, however, a notable endogenous SrtA homolog from streptococcus pyogenes (SrtA_{pyogenes}) has revealed advantages beyond the utility of SrtA_{staph}. SrtA_{pyogenes} is capable of recognizing a diverse catalog of substrates, which has enabled a multifaceted approach to site-specifically modify different regions within a single protein target.¹⁹ Furthermore, SrtA_{pyogenes} is capable of generating isopeptide bonds by accepting ϵ -amine of lysine, as well as processing the N-terminal amines of glycine, serine, and even D-asparagine residues.⁷⁴ In general, streptococcal sortases have exhibited promiscuous substrate preferences, and an unprecedented tolerance for LPXLG motifs, which may be useful for SML reactions.⁷¹ To date, sortase A homologs employed for SML reactions represent only a fraction of the thousands of sortase genes encoded by genomes across the bacterial kingdom.^{75,76} Therefore, there exists an untapped potential for harnessing the reactivity of sortase A homologs to broaden the scope of applicable SML substrates. Our lab has highlighted this concept by determining the substrate preferences of eight naturally occurring SrtA

homologs, each exhibiting a preference for residues across each position along the sorting motif.⁷⁰ Positions 4 and 5 displayed significant deviation from the canonical LPXTG sorting motif, where many SrtA homologs preferred substrates with several different amino acids in the 5th position. Notably, sortase A from *Streptococcus pneumoniae* (SrtA_{pneu}) demonstrated the broadest substrate tolerance of non-canonical amino acids in the 5th position (**Figure 4**).

Substrate													Product	Excised Fragment		
Abz-LPATXG-K(Dnp)		Sortase A H ₂ N-OH excess											Abz-LPAT-NHOH	+ XG-K(Dnp)		
Sortase A homolog	X =	V	Y	S	W	L	Nle	G	A	N	F	Q	C			
<i>S. aureus</i>		2	1	1	<1	<1	<1	85	2	<1	1	2	7			
<i>S. suis</i>		1	4	69	<1	<1	3	78	72	58	3	3	71			
<i>S. oralis</i>		2	2	53	<1	1	2	75	55	3	2	6	43			
<i>S. pneumoniae</i>		32	24	84	11	38	14	84	92	48	42	12	71			
<i>L. monocytogenes</i>		3	45	42	6	2	2	88	17	35	31	2	43			
<i>E. faecalis</i>		1	2	8	<1	1	1	14	5	6	3	2	10			
<i>L. lactis</i>		1	3	39	<1	2	1	82	58	28	2	1	37			
<i>B. anthracis</i>		2	2	11	<1	1	1	42	19	17	1	<1	15			

Figure 4. Comparison of substrate preferences for the 5th position of the sorting motif among SrtA homologs.⁷⁰ These values represent % conversion of substrate to excised fragment. No cleavage was observed for X = P, T, I, D, E, R, K, H.

Although SrtA_{pneu} recognizes a variety of sorting motifs, high performance liquid chromatography (HPLC) and mass spectrometry (MS) analyses of product conversion from model SML reactions revealed a strong preference for an LPATA substrate *in vitro*, which is strikingly different than the LPETG preference for SrtA_{staph}. The SrtA reactivity and specificity trials we've published may persuade the reader to believe that SrtA is not appropriate for SML applications based on the suboptimal product conversion of various substrates. However, the data presented is reflective of unoptimized reactions, where even the lowest substrate conversion can be drastically

improved by redesigning the reaction conditions. The utility of SrtA_{pneu} has provided an opportunity to bolster the applicability of SML protein engineering by broadening the scope of substrate targets. As a result, SrtA_{pneu} has potentially reduced the necessity to mutagenically implement a sorting motif into protein targets, which improves the compatibility of SML for endogenous proteins. The unique substrate promiscuity of SrtA_{pneu} has potentiated an interest toward elucidating active site residues responsible for dictating the mechanism of recognition. A deep understanding of these interactions on a molecular level would likely provide insight to this phenomenon, which will advance our understanding of sortase enzymology and sortase-mediated chemistry. The structure of SrtA_{pneu} has not been published, which has compelled our efforts toward an in-depth analysis of SrtA_{pneu} substrate recognition through structural characterization.

1.3 Substrate Binding and Structural Characteristics of SrtA Homologs

Multiple structures of various SrtA homologs have been published over the last decade, which were either characterized by X-ray crystallography or solution nuclear magnetic resonance (NMR) spectroscopy. Published structures of SrtA_{staph} have aided in developing an improved understanding of enzyme-substrate binding interactions with residues in the sorting motif, and insight into the reverse protonation mechanism instigating transpeptidation reactions.^{77,78} Published structures of SrtA homologs manifest an 8-stranded β -barrel fold, which is a conserved feature across SrtA enzymes.³⁰ Strands comprising the β -barrel are flanked by a series of alpha and 3_{10} helices, as well as disordered loops varying in size and position among homologs.^{31,33,79–82} In general, SrtA enzymes share an evolutionarily conserved active site housing three catalytic residues; a cysteine to establish a transient thio-acyl linkage, a histidine to facilitate thiolate formation, and an arginine thought to provide hydrogen bonding to stabilize active site residues

essential for catalysis.^{33,83} In the case of SrtA_{staph}, the floor of the active site, or binding groove, is formed by β 4 and β 7 loops and the adjacent walls are composed of conjoining loops, and helices. The binding pocket adopts the appearance of a bent “L” shape, which may justify the necessity for proline in the 2nd position of the sorting motif as it situates the amide bond linking 4th and 5th position residues towards the active site cysteine.⁷⁸ The non-polar carbon fork of the 1st position leucine establishes hydrophobic contacts with residues in the β 6/ β 7 loops, and the 2nd position proline is buried within a hydrophobic cleft formed by residues in β 4 and β 7 strands. The 3rd position alanine maintains distant hydrophobic interactions with the H1 helix, which may provide a rationale for the indiscriminate preference for residues in this position as there exists ample space for cumbersome side chains. The 4th position threonine pushes a nearby tryptophan residue (Trp194) out of the active site, which situates the active site cysteine in proximity to the scissile peptide bond. Preservation of threonine is critical for this mechanism, as substrates substituting glycine in the 4th position are unreactive. The 5th position glycine is predicted to associate with the β 7/ β 8 loop, which hypothetically undergoes a distinct transition to a structurally ordered conformation upon substrate docking.^{78,84–86} Preferential recognition of the 5th position residue is anticipated to be partially dependent on the length of the β 7/ β 8 loop. SrtA_{staph} has a relatively large β 7/ β 8 loop compared to other homologs, which may condone the stringent selectivity for glycine in the 5th position of the sorting motif. After substrate docking is facilitated, nucleophilic attack of the scissile peptide bond by the active site cysteine repositions the β 7/ β 8 loop further from the binding pocket, revealing a sterically unencumbered site for incoming nucleophiles.⁷⁸

The substrate bound SrtA_{staph} complex published by Suree et al. has constructed a framework for interpreting the role of each residue positioned along the LPAT substrate analog and deciphering crucial interactions with residues housed within the active site (**Figure 5**).⁷⁸

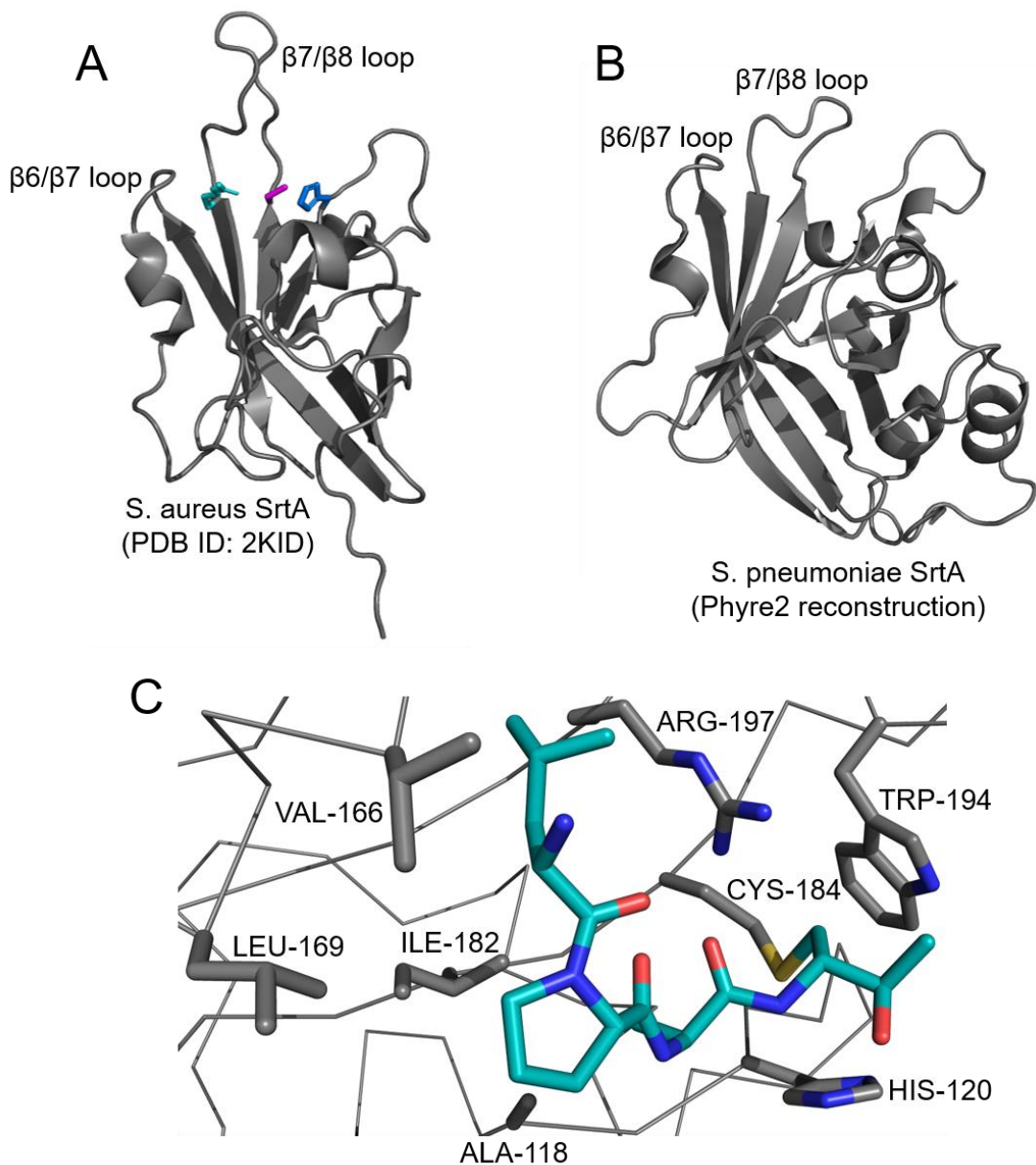


Figure 5. (A) Solution NMR structure of SrtA_{staph}. Arginine (cyan), cysteine (magenta), and histidine (blue) stick structures represent catalytic residues in the enzyme active site. (B) Predicted structure of SrtA_{pneu} based on a one-to-one threading model of SrtA_{pyogenes} (PDB ID: 3NF7) from the Phyre2 structural prediction server. (C) Solution NMR structure of the SrtA_{staph} active site with a bound LPAT* substrate analog (PDB ID: 2KID). Side chains of residues comprised within the active site are shown as stick structures, highlighting several hydrophobic interactions stabilizing the substrate-bound state.

Unfortunately, this systematic model fails to provide a holistic depiction of SrtA substrate recognition and binding since the substrate analog only contains the first four amino acids (LPAT) in the sorting sequence. As a result, identifying active site residues that interact with the 5th position residue of a target sorting motif is challenging based on this model. However, Suree and coworkers illuminated regions in SrtA_{staph} thought to be responsible for recognizing the 5th position residue in the sorting motif as well as coordinating entry of the incoming nucleophile.

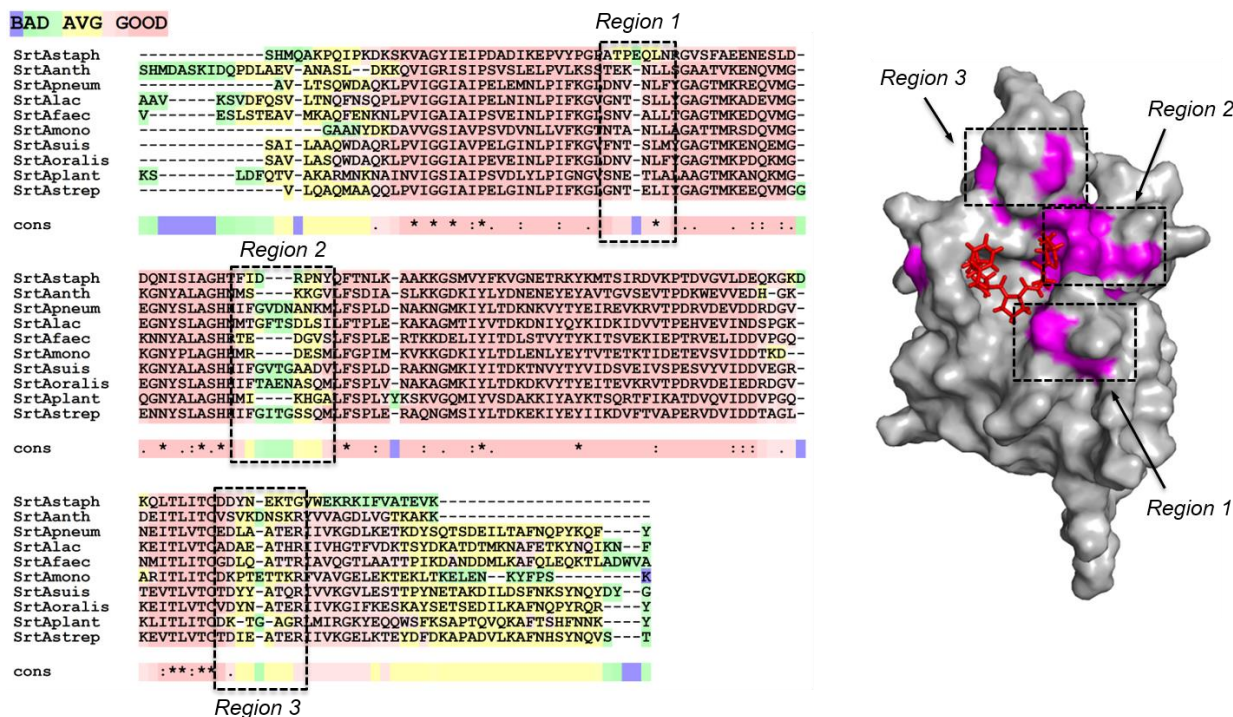


Figure 6. Sequence alignment of selected SrtA homologs. The regions of greatest difference (boxed) correspond to the regions highlighted in the structure (right), indicating the least sequence homology on the structural features predicted to interact with the 5th position of the sorting sequence. The structure (right) is a surface representation of SrtA_{staph} bound to a substrate analog LPAT* (PDB ID: 2KID). Residues highlighted in magenta were determined to interact with the incoming nucleophile by analysis of peak perturbation during an ¹⁵N-HSQC monitored titration of SrtA_{staph} with triglycine. These residues are primarily situated around the region of the binding pocket and are predicted to interact with the C-terminus of the sorting signal.

Three regions in SrtA_{staph} displayed significant alterations in their backbone resonances in the presence of a triglycine nucleophile, which was monitored using ¹⁵N heteronuclear single quantum coherence (HSQC) NMR (Figure 6).⁸⁷ Residues harbored within these regions are likely

contacting the 5th position residue of the sorting motif, which suggests that they are critical for substrate recognition. Sequence alignments of SrtA homologs indicated distinct differences in primary structure within these regions, which provides a rationale for the diverse tolerance of various residues in the 5th position. This evidence coincides with our experimental findings, where SrtA homologs demonstrated a variety of preferences for the 5th position residue. Supplementary assessments are necessary to establish a more thorough understanding of SrtA substrate recognition.

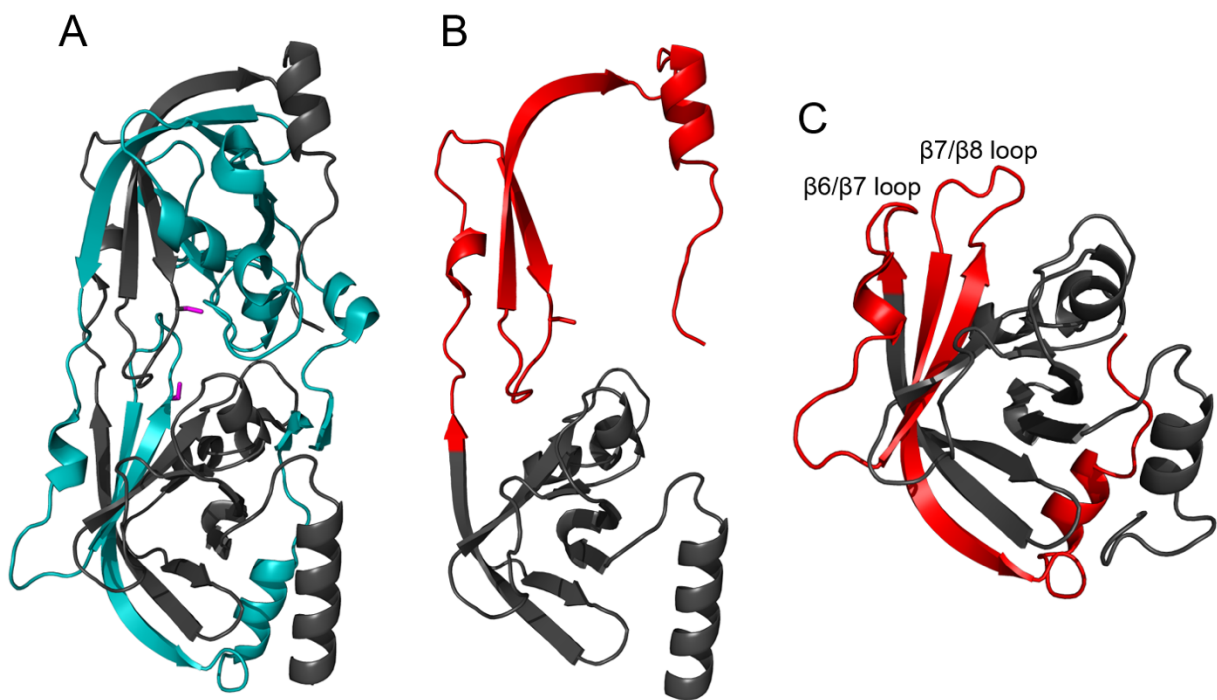


Figure 7. (A) The 3D domain swapped dimer of SrtA_{pneu} (PDB ID: 4O8L). (B) A domain swapped monomer from the dimeric structure of SrtA_{pneu}. (C) Predicted structure of SrtA_{pneu} based on a one-to-one threading model of SrtA_{pyogenes} (PDB ID: 3FN7) from the Phyre2⁸⁸ structural prediction server. In both (B) and (C) structures, the red colored regions resemble the domain swapped portion of the structure shown in (A).

A 3D domain swapped structure of SrtA_{pneu} has been deposited in the Protein Data Bank (PDB ID: 4O8L), but monomeric enzyme with (or without) bound substrate has not yet been characterized (**Figure 7**). A structure of monomeric SrtA_{pneu} bound to a substrate may distinguish

novel features that are responsible for its unique substrate tolerance, similar to aforementioned efforts toward determining which active site residues contribute to substrate recognition in SrtA_{staph}. Phyre2 structural prediction algorithms have served as a preliminary means of determining which residues perpetuate the promiscuous substrate tolerance of SrtA_{pneu}.⁸⁸ Structure predictions suggest that SrtA_{pneu} has a smaller $\beta 7/\beta 8$ loop than SrtA_{staph}, which may confer a broader substrate tolerance, as this loop region is thought to be important for recognizing the 5th position residue of the sorting motif.

1.4 SrtA_{pneu} Enzyme Activity as a Function of Oligomeric State

The structure of monomeric SrtA_{pneu} has yet to be determined, however, the dimeric form of the enzyme has piqued our interests toward understanding the mechanism of assembly *in vivo* and *in vitro*, as well as its role on enzyme activity. In previous SrtA_{pneu} studies, we sought to evaluate the catalytic activity of both dimeric and monomeric forms of the enzyme.⁷⁰ These multimers were identified in an IMAC elution of purified SrtA_{pneu} by native-PAGE, where numerous distinct bands were observed, as opposed to SDS-PAGE analysis displaying a single band. In model SML studies, we hypothesized that monomeric SrtA_{pneu} was an active form of the enzyme, whereas multimeric forms were thought to be an inactive form, based on their activity in the presence of an Abz-LPATAG-K(Dnp) peptide substrate and a potent hydroxylamine (NH₂OH) nucleophile (**Figure 8**). Model SML reactions utilizing multimeric SrtA_{pneu} rapidly plateaued with minimal product formation (21% product conversion). Conversely, identical reactions involving monomeric SrtA_{pneu} achieved significantly higher product formation (95% product conversion). To confirm our suspicions, size exclusion fast protein liquid chromatography (SE-FPLC) was utilized to confirm the presence of all SrtA_{pneu} assemblies. The spectra revealed multiple species

varying in molecular weight, and native-PAGE bands of respective fractions travelled identical to native-PAGE bands of a SrtA_{pneu} heterogeneous mixture.

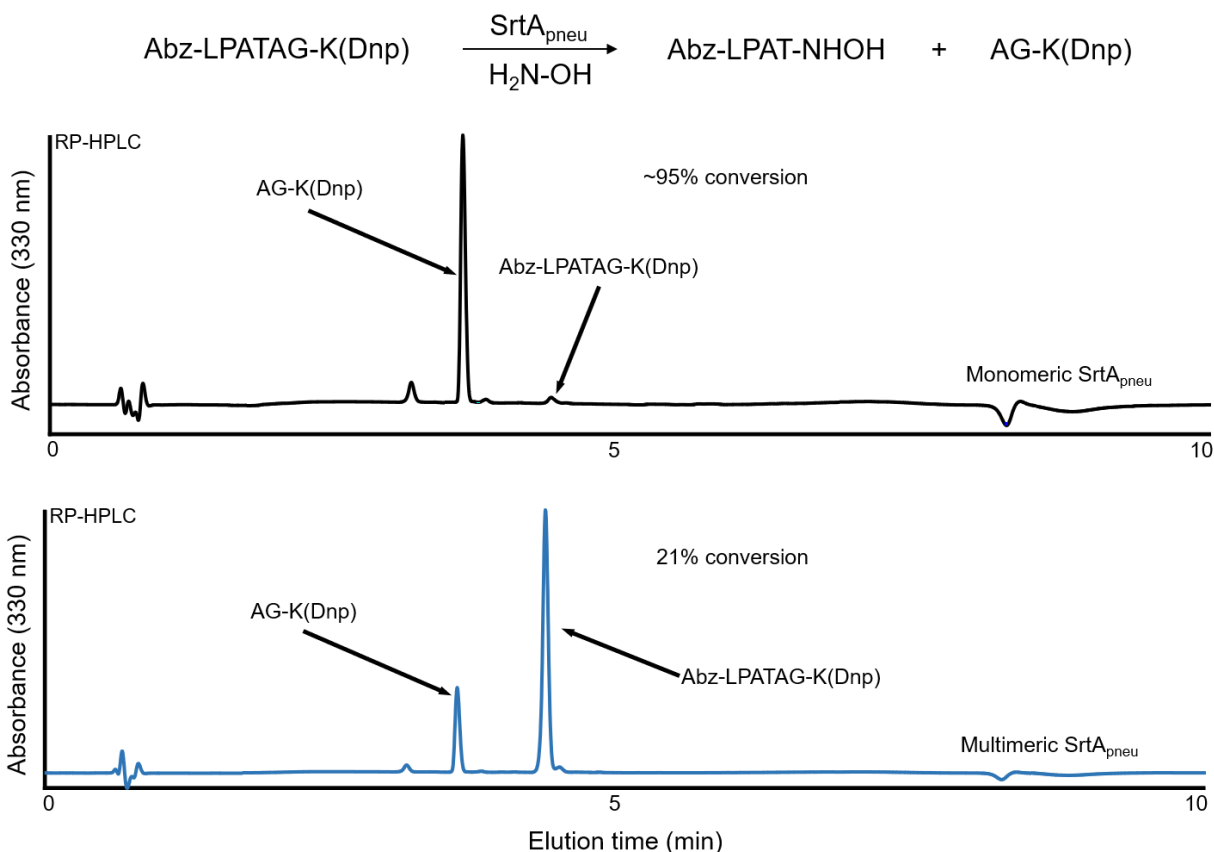


Figure 8. RP-HPLC analysis of model SML reactions demonstrating the difference in activity between monomeric (top) and multimeric (bottom) SrtA_{pneu} preparations.⁷⁰

Bacterial expression of a truncated SrtA_{pneu} clone ($\Delta 80$) generated substantial amounts of inactive enzyme, thus it became imperative to explore options allowing us to regenerate fully active monomeric SrtA_{pneu}.⁷⁰ We initially proposed the idea of subjecting SrtA_{pneu} to conditions eliciting the disassembly of all SrtA_{pneu} multimers, followed by the refolding of denatured SrtA_{pneu} to monomeric enzyme. Our original hypothesis speculated that SrtA_{pneu} dimerization is an equilibrium driven process, where acute concentrations of enzyme may shift the equilibrium toward an energetically favorable dimeric fold. Attempts to incubate serial diluted samples at room temperature sought to evaluate this possibility. Contrary to our rationale, the intensity of a dimeric

SrtA_{pneu} band in a native-PAGE gel remained unaffected. We proposed the possibility that SrtA_{pneu} dimerization is co-translational phenomenon, where artificially elevated concentrations in-vivo may instigate a dimeric fold. If true, we anticipated that dismantling SrtA_{pneu} assemblies during purification, followed by refolding under native conditions, may provide an opportunity for SrtA_{pneu} to reassemble into monomeric enzyme. A denaturing agent capable of disrupting any intermolecular interactions, namely domain swapping contacts, was implemented in our IMAC purification buffers. Specifically, we employed an initial denaturing IMAC purification using 8 M urea, where *E. coli* cells were lysed under denaturing conditions, the protein was purified from clarified lysate in denaturing buffer via IMAC, followed by a rapid dilution of denaturing eluate in non-denaturing buffer to refold monomeric enzyme. The non-denatured diluted protein was reconcentrated by a non-denaturing IMAC purification, and eluted fractions were further purified by SE-FPLC to isolate residual dimers from reassembled monomeric enzyme. We also added tris(carboxyethyl)phosphine (TCEP), a non-sulfurous reducing agent, to our IMAC purification buffers to disfavor cysteine disulfide bridging between monomers and preserve the reduced form of the thiol. Overall, our enzyme refolding protocol in tandem with SE-FPLC enrichment significantly improved SrtA_{pneu} monomer recovery (**Figure 9**). Furthermore, the regenerated SrtA_{pneu} monomer performed identical to previous model SML reactions, where 95% conversion was observed for the refolded enzyme.

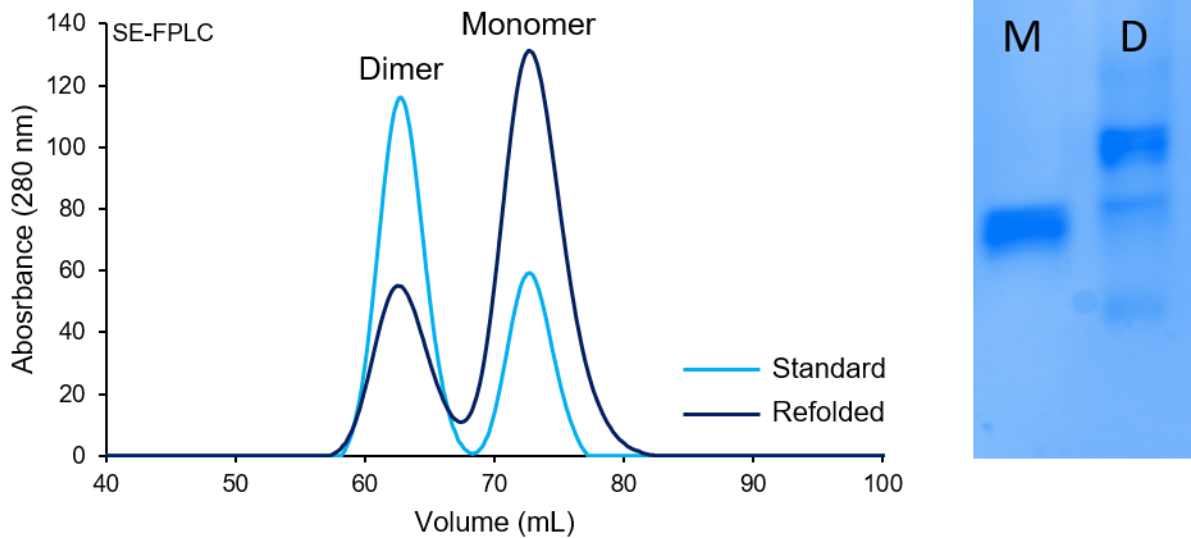


Figure 9. A comparison of standard vs. refolded SrtA_{pneu} preparations via SE-FPLC (left). Native-PAGE analysis of isolated monomeric (M) fractions and dimeric (D) fractions by SE-FPLC (right).

Supplementary investigations have revealed the presence of both monomeric and dimeric forms of SrtA_{staph} *in vivo* and *in vitro*.⁸⁹⁻⁹¹ The *in vitro* catalytic activity of monomeric SrtA_{staph} has been evaluated by Lu and coworkers, where ligation reactions with homodimeric preparations demonstrated superior product conversion compared to monomeric enzyme. In contrast, the insertion of a non-dimerizing SrtA_{staph} mutant in a knockout strain of *S. aureus* by Zhu et al. resulted in an increased presence of sortase-catalyzed surface anchored proteins *in vivo*, which provides evidence in support of a catalytically active SrtA_{staph} monomer. *In vivo* observations of SrtA_{staph} activity as a function of oligomerized state are in clear contrast to *in vitro* studies, but coincide our evidence indicating monomeric SrtA_{pneu} is the catalytically active form *in vitro*. The pervasive dimerization among SrtA homologs has raised questions regarding the biological significance of these dimers, as they are anticipated to serve as a means of regulating enzyme deactivation when extracellular protein appendage is unnecessary.^{89,91} This proposed mechanism of regulation is not unfounded, as there exists many enzymes (i.e. phospholipase) that are governed in this fashion.⁹² Although interplay between monomeric and dimeric forms of SrtA_{staph} and

SrtA_{pneu} has been evaluated, similar investigations have not been geared toward other SrtA homologs.

1.5 Overview of Project Goals

The modern utility of sortases as a tool for protein engineering has broadened the accessibility of site-directed ligation chemistry, which is substantiated by previous efforts toward understanding substrate tolerance among sortase homologs and the circumvention of limitations historically plaguing sortase-mediated ligation techniques. The continued development of this system hinges upon our ability to develop a structure-function relationship among sortases in an effort to improve our knowledge of substrate recognition among homologs with diverse substrate tolerances. To this end, the long-term goal of this project is to determine the structure of SrtA_{pneu} covalently docked with a peptide inhibitor as a means to identify novel interactions with active-site residues prompting a unique promiscuous substrate tolerance. As described in this thesis, preliminary investigations of monomeric SrtA_{pneu} structure involved protein crystallography followed by X-ray diffraction. Concurrently, we have attempted to construct a non-cleavable ketomethylene-based sorting motif analogs in an effort to prolong occupancy within the SrtA_{pneu} active site. Sorting motif substitution of 4th and 5th position residues with ketomethylene dipeptide was anticipated to mimic the performance of canonical SrtA_{pneu} substrates. Unfortunately, the challenging synthesis of a solid-phase ready ketomethylene dipeptide and the rapid degradation of ketomethylene-based substrates depreciated effectiveness of this approach. Correspondingly, we've redesigned our canonical SrtA_{pneu} substrate sorting motif with a cysteine residue in the 4th position. We believe this substrate analog will establish a disulfide linkage with the active site cysteine, allowing for elucidation of key interactions between the enzyme and residues positioned

along the sorting motif. Identification of substrate bound SrtA complex via LC-ESI-MS has propelled our efforts toward determining the structure using HSQC NMR.

Chapter 2 – Screening Studies for X-ray Crystallography of SrtA_{pneu}

2.1 Preparation of SrtA_{pneu} for Crystallization

Prior to screening crystallization conditions, it was first necessary to generate suitable preparations of the SrtA_{pneu} enzyme that were monomeric. To this end, an expression vector encoding a truncated version of SrtA_{pneu} lacking the first 80 residues (hereafter referred to as simply SrtA_{pneu}) and fused to an N-terminal His₆ tag was obtained. In this construct, the hydrophobic transmembrane domain was removed to increase the *in vitro* solubility of SrtA_{pneu} in aqueous buffers. A glycerol stock of transformed *E. coli* BL21 (DE3) was used to express SrtA_{pneu} using standard molecular biology techniques. After denaturing cell lysis using 8 M urea, the enzyme was separated from the cellular debris via centrifugation, then isolated from the supernatant using immobilized metal affinity chromatography (IMAC) under denaturing conditions (**Figure 10**, lanes 2-6). The denatured protein eluate from IMAC was then diluted ten-fold into a non-denaturing buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM TCEP) to refold SrtA_{pneu}, then repurified via IMAC under non-denaturing conditions to isolate soluble SrtA_{pneu} (**Figure 10**, lanes 7-10). The sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis of the refolded SrtA_{pneu} eluate revealed an intense band near 26 kDa, which was consistent with the calculated 20.1 kDa molecular weight of SrtA_{pneu}. This SDS-PAGE gel also showed the presence of a ~50 kDa SrtA_{pneu} dimer that persisted in the sample despite reducing and denaturing preparations.

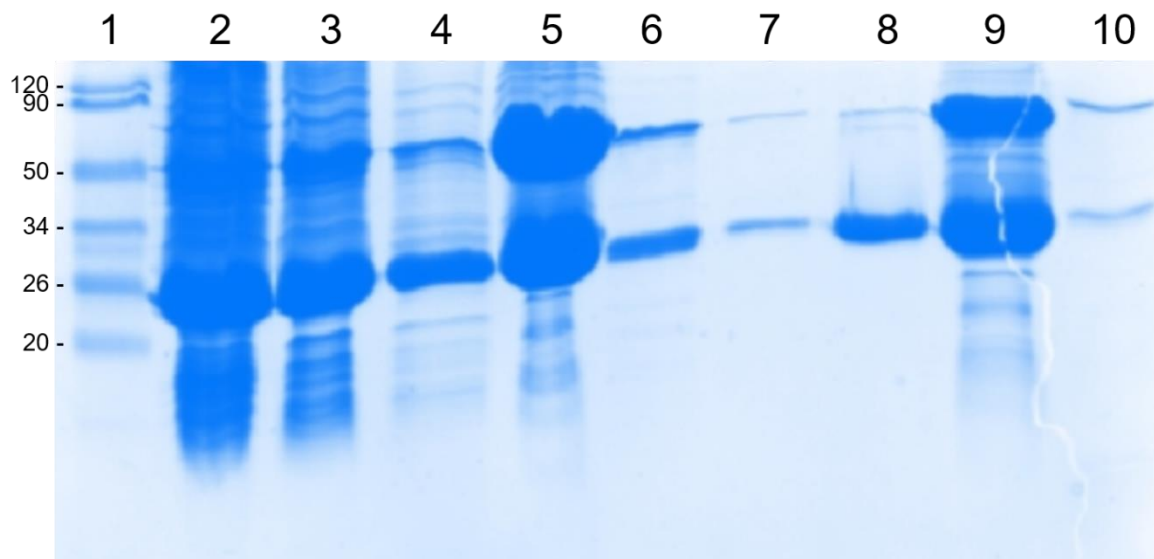


Figure 10. SDS-PAGE analysis of a SrtA_{pneu} IMAC/refolding purification scheme: (1) Protein molecular weight ladder, (2) cell lysate supernatant, (3) flow-through fraction of IMAC column under denaturing conditions, (4) denaturing wash, (5) 1st denaturing elution fraction, (6) 2nd denaturing elution fraction, (7) flow-through fraction of IMAC column following dilution of SrtA_{pneu} in non-denaturing buffer, (8) non-denaturing wash, (9) 1st non-denaturing elution fraction, (10) 2nd non-denaturing elution fraction.

In order to separate monomeric SrtA_{pneu} from higher order aggregates, the refolded protein solution was subjected to size exclusion chromatography (SEC). Monomeric fractions were collected and pooled, and analysis by analytical SEC revealed that final the protein preparation consisted of >85% monomer (**Figure 11**). Electrospray ionization mass spectrometry (ESI-MS) of the final monomer preparation reported a mass of 20,144 Da, in excellent agreement with the calculated molecular weight of 20,145 Da for SrtA_{pneu} (**Figure 11**). Taken together, SEC and ESI-MS analyses provided evidence for a predominantly monomeric batch of SrtA_{pneu}, which was subsequently concentrated to ~6 mg/mL for protein crystallization trials.

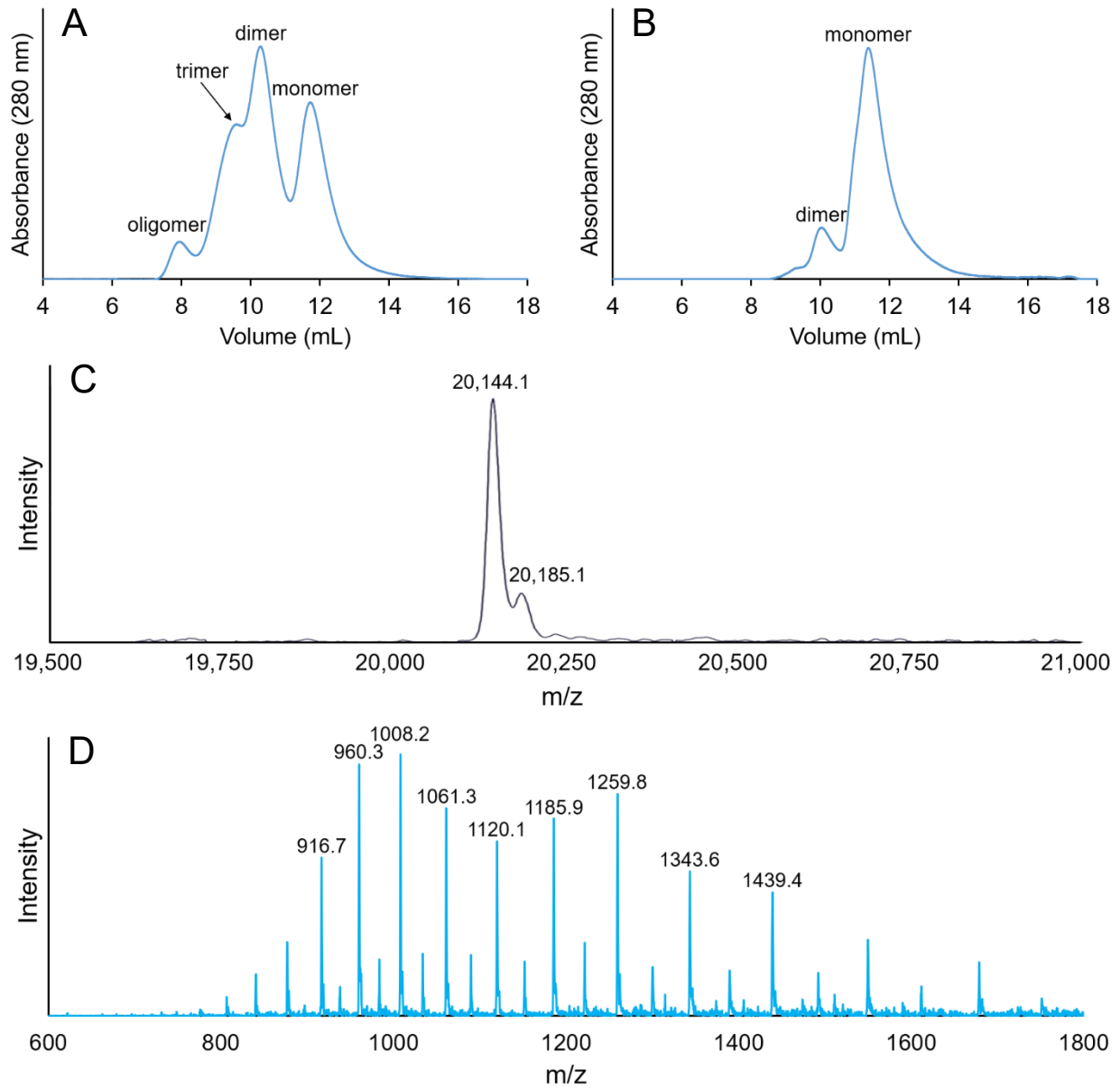


Figure 11. SEC traces of IMAC purified, refolded SrtA_{pneu} (A) and collected monomeric fractions (B). A deconvolved mass spectrum (C) of SrtA_{pneu} generated from the (D) raw ESI-MS spectrum of the purified enzyme.

2.2 Summary of Crystal Screening Efforts

With a monomeric batch of SrtA_{pneu} in hand, we turned our attention to screening conditions for crystallization of this enzyme. Initially, this involved utilizing vapor-drop diffusion methods for generating crystals. The concentrated stock of 6 mg/mL monomeric SrtA_{pneu} was screened against PEG/Ion2 and Index screening condition kits, each containing 48 and 96 different conditions, respectively. Four room temperature conditions from the Index screening kit produced a variety of crystal morphologies including wafer, rod, and asymmetric crystals (**Table 2, Figure 12**).

Table 2. Preliminary screening conditions resulting in crystal formation (pH 5.5, 25% PEG 3350, RT).

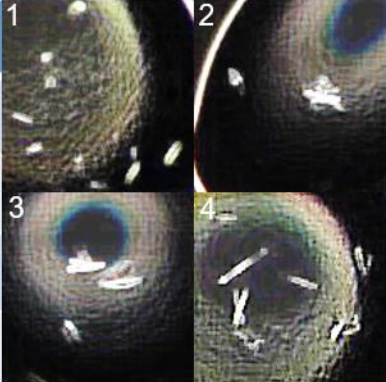
Condition #	Salt Composition	
1	0.1 M Bis-Tris (BT)	
2	0.1 M Bis-Tris 0.2 M Sodium Chloride (NaCl)	
3	0.1 M Bis-Tris 0.2 M Ammonium Acetate (NaOAc)	
4	0.1 M Bis-Tris 0.2 M Magnesium Chloride (MgCl ₂)	

Figure 12. Crystals formed by screening conditions in Table 2.

These crystallization conditions shared numerous similarities, which suggested that Bis-Tris, pH 5.5, and polyethylene glycol (PEG) 3350 were promising components for inducing SrtA_{pneu} crystallization at room temperature. Based on this analysis, we sought to optimize these conditions by determining SrtA_{pneu} nucleation dependence as a function of pH and PEG 3350 concentration (**Figure 13**).

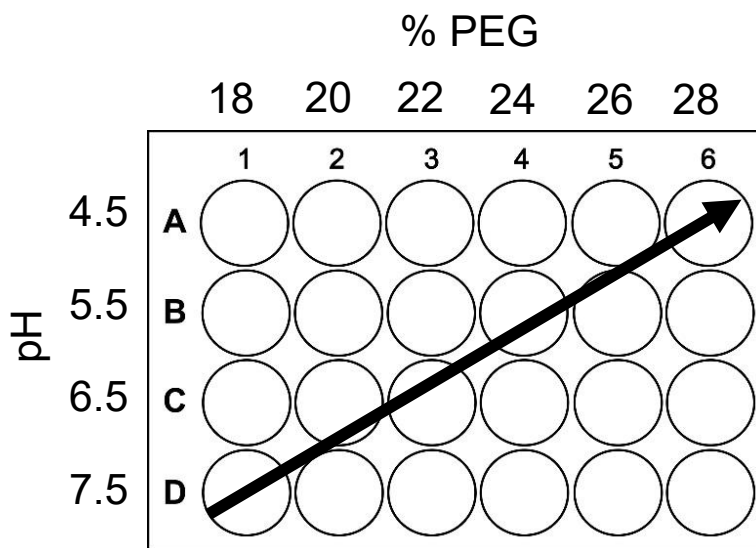


Figure 13. Crystal optimization as a function of pH and PEG 3350. The arrow represents the observed pattern of crystal formation, appearing more frequently as pH decreases and PEG 3350 increases.

After two weeks, we observed the presence of crystals in conditions with low pH (4.5-5.5) and relatively high PEG 3350 concentration (24-26%). Microscopic investigation of crystals formed in the presence of these conditions revealed relatively small crystals that did not appear to have the defined 3D structure (i.e. hexagonal prism) of diffractable protein crystals.

Crystal screening was continued by maintaining 25% PEG 3350 and pH 4.5 or 5.5, however this time adjusting the concentrations of Bis-Tris (0.05-0.30 M) and other tuning salts (0.05-0.40 M). Hits were detected that produced crystals after two weeks, which essentially replicated the appearance of crystals grown in the previous pH/PEG optimization trial. Unlike the pH/PEG optimization trial, there was not a noticeable pattern indicating which salt/buffer concentrations favored crystal formation. This suggested that crystal growth and morphology was neither dependent on Bis-Tris nor tuning salt concentrations. It should be noted that each optimization trial was designed to replicate the initial screening conditions (#1-4) as a control, however, the reproducibility of crystals generated by these conditions was inconsistent. Although

the crystals we had generated were far from ideal for structure determination, we attempted to diffract our top candidates to determine whether we were generating salt or protein crystals (**Table 3**). Every crystal subjected to X-ray diffraction analysis displayed a prominent “ice ring” pattern, which occurs when the protein crystal is thawed and refrozen during transfer to the goniometer head of the diffractometer. Beyond this artifact, we didn’t observe any indication of a diffraction pattern corresponding to a proteinaceous crystal.

Table 3. Diffracted crystals formed by conditions A-E (pH 5.5, 25% PEG 3350, RT).

Condition	Salt Composition
A	0.05 M BT
B	0.25 M NaOAc
C	0.10 M BT
D	0.10 M BT 0.15 M MgCl ₂
E	0.10 M BT 0.05 M NaCl

We next tried crystallization conditions that were not included in either the PEG/Ion2 or Index crystal screening kits. Our previous screening efforts suggested that chloride-containing salts may promote crystal formation. Therefore, additional trials with potassium chloride (KCl), ammonium chloride (NH₄Cl), or calcium chloride (CaCl₂) tuning salts (0.05-0.40 M) in addition to Bis-Tris (0.05-0.30 M) were prepared. However, these did not demonstrate any capacity to grow diffractable crystals. We also attempted to modulate the standard 1:1 (2 μ L drop) ratio of enzyme loading to mother liquor. Since we had observed some protein aggregation during pH/PEG optimization trials, we anticipated that a lower enzyme loading relative to mother liquor would slow nucleation and allow for improved crystal packing. Unfortunately, no crystals were observed using diluted SrtA_{pneu} preparations even months after plating.

Finally, several months after our initial pH/PEG optimization trials, we did observe the presence of well-defined opalescent crystals in a well containing a replicate of the #4 screening

condition (0.1 M BT, pH 5.5, 0.2 M MgCl₂, 25% PEG 3350, RT). We attempted to loop these crystals, however, they had formed contacts with neighboring crystals and we were unable to effectively loop a single crystal without fracturing their structural integrity. In consideration of this promising result, we focused on optimizing condition #4 to recreate, and possibly improve, the iridescent crystals observed previously. Through the advice of our collaborators, we decided to optimize condition #4 by modulating PEG 3350 concentration as well as enzyme to mother liquor loading ratio. Several weeks after, we observed protein aggregation across over half of the preparations, and no sign of crystal formation.

In summary, while we have been able to generate a monomeric preparation of SrtA_{pneu}, we have yet to identify conditions that produce crystals suitable for structure determination. Given the fact that multiple sortases have been successfully characterized using X-ray crystallography, we anticipate that monomeric SrtA_{pneu} will ultimately be amenable to X-ray characterization, however additional crystallization condition screening is required. In addition, it may be necessary to redesign the protein construct itself, as the 80 residue truncation or the presence of the N-terminal His₆ tag may not be optimal for crystal formation.

Chapter 3 – Preparation and NMR Characterization of ¹⁵N-labeled SrtA_{pneu}

3.1 Preparation of unlabeled TEV-SrtA_{pneu} and 1D ¹H-NMR Analysis

In parallel with our efforts to generate X-ray quality crystals of SrtA_{pneu}, we also began studies aimed on elucidating the enzyme's 3D structure using solution NMR. To this end, we first generated a new stock of SrtA_{pneu} in order to monitor its stability using one dimensional ¹H-NMR. Anticipating that we may need to remove the N-terminal His₆ tag, a new clone of SrtA_{pneu} (TEV-

SrtA_{pneu}) was obtained, which included a TEV cleavage site between the His₆ tag and the catalytic domain. Although numerous publications have demonstrated that His₆ tags typically do not perturb protein folding and function, we had contemplated that our inability to effectively crystallize SrtA_{pneu} may be a repercussion of the flexible His₆ tag and thus the TEV cleavage site was included as an option.

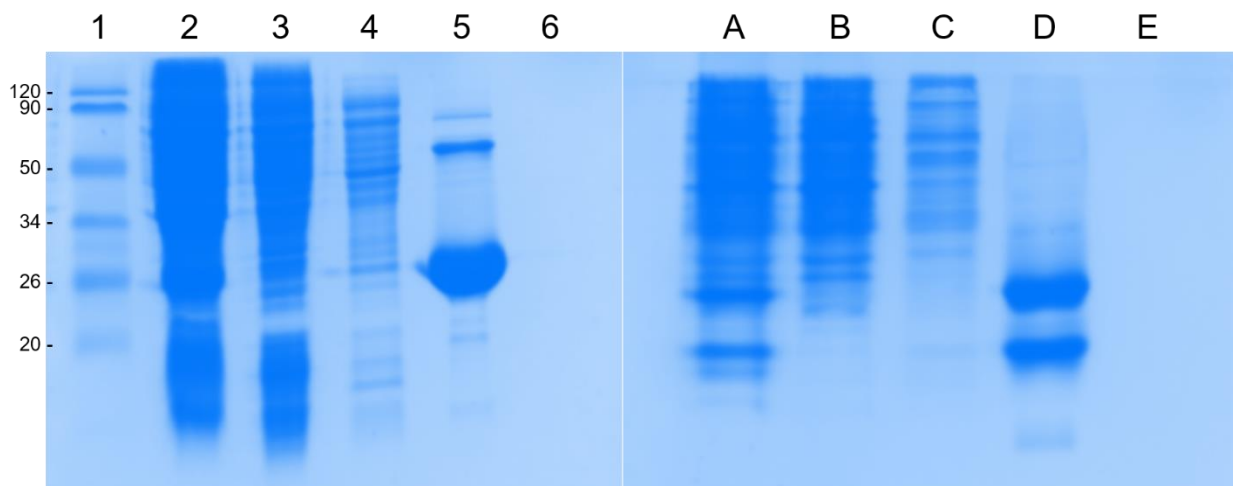


Figure 14. (Left) SDS-PAGE analysis of non-denaturing IMAC purification of TEV-SrtA_{pneu}: (1) Protein molecular weight ladder, (2) cell lysate supernatant, (3) flow-through fraction of IMAC column under non-denaturing conditions, (4) wash, (5) 1st non-denaturing elution fraction, (6) 2nd non-denaturing elution fraction. (Right) Native-PAGE analysis of non-denaturing IMAC purification of TEV-SrtA_{pneu}: (A) cell lysate supernatant, (B) flow-through fraction of IMAC column under non-denaturing conditions, (C) wash, (D) 1st non-denaturing elution fraction, (E) 2nd non-denaturing elution fraction.

Interestingly, and in contrast to the SrtA_{pneu} clone used for crystallization trials, we found that TEV-SrtA_{pneu} yielded sufficient quantities of active monomer without the need for refolding. As shown in **Figure 14**, a band with an appropriate molecular weight was observed following simple, non-denaturing IMAC purification, and native-PAGE analysis of the same purification revealed significant quantities of both a monomer and dimer. An SDS-PAGE analysis of the elution fraction revealed a high intensity band at ~26 kDa of SrtA_{pneu} as well as residual higher molecular weight polypeptides structures (**Figure 14**, lane 5). Additionally, we noticed a signature pair of bands representing monomeric (bottom) and dimeric (top) enzyme in a native-PAGE gel.

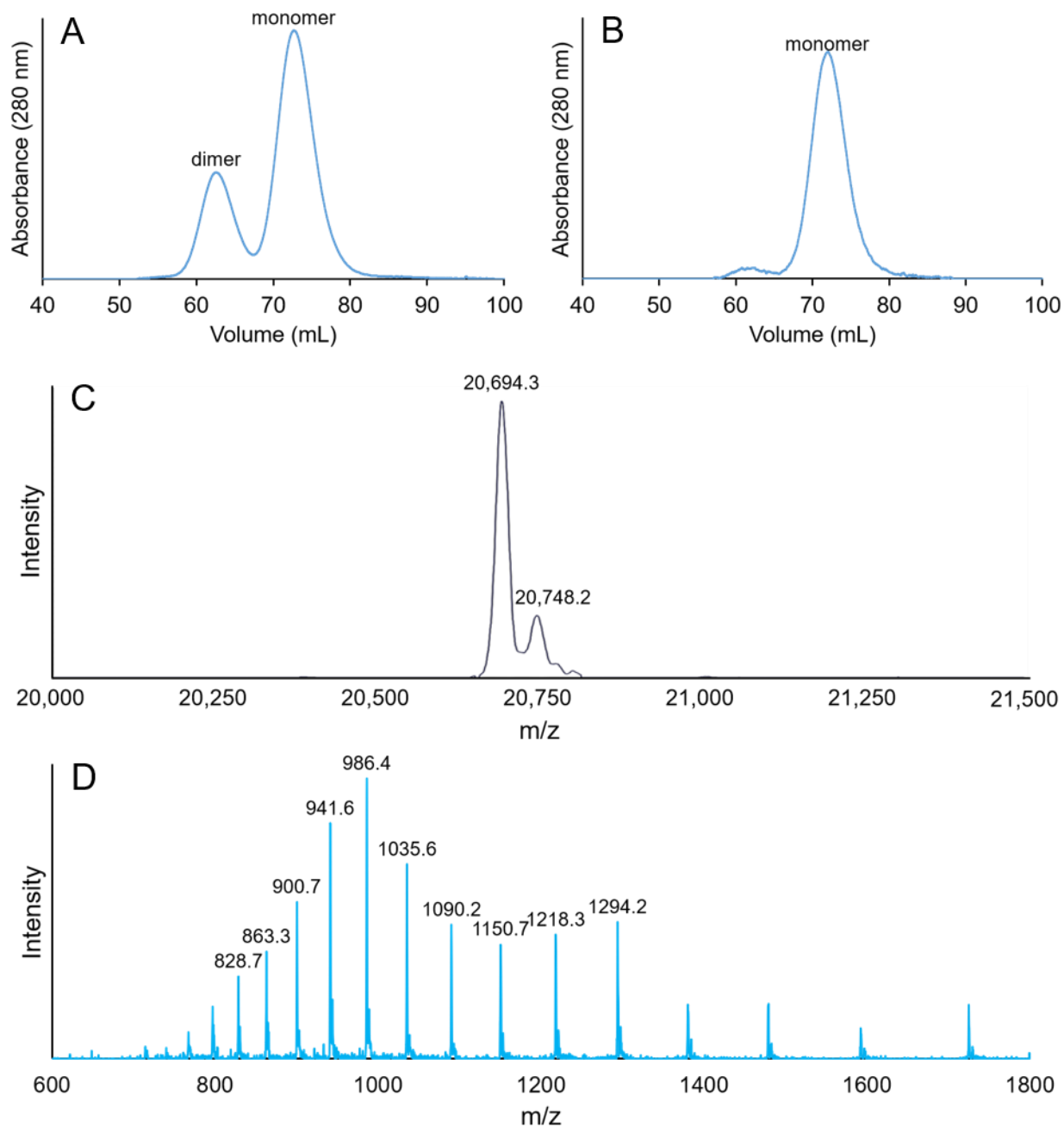


Figure 15. SEC traces of IMAC purified TEV-SrtA_{pneu} (A) and collected monomeric fractions (B). A deconvolved mass spectrum (C) of TEV-SrtA_{pneu} generated from the corresponding raw ESI-MS spectrum (D).

As noted above, a denaturing IMAC/refolding purification procedure was not applied to this protein expression, which provides a rationale for the relatively large band of dimer relative to monomer. In order to isolate the desired monomer of TEV-SrtA_{pneu}, we relied on a newly

acquired HiPrep 16/60 Sephacryl 200-HR size exclusion column, which provided excellent separation of the TEV-SrtA_{pneu} monomer from higher order assemblies (**Figure 15A,B**). Mass spectrometry also confirmed that the isolated protein had the expected molecular weight (**Figure 15C,D**).

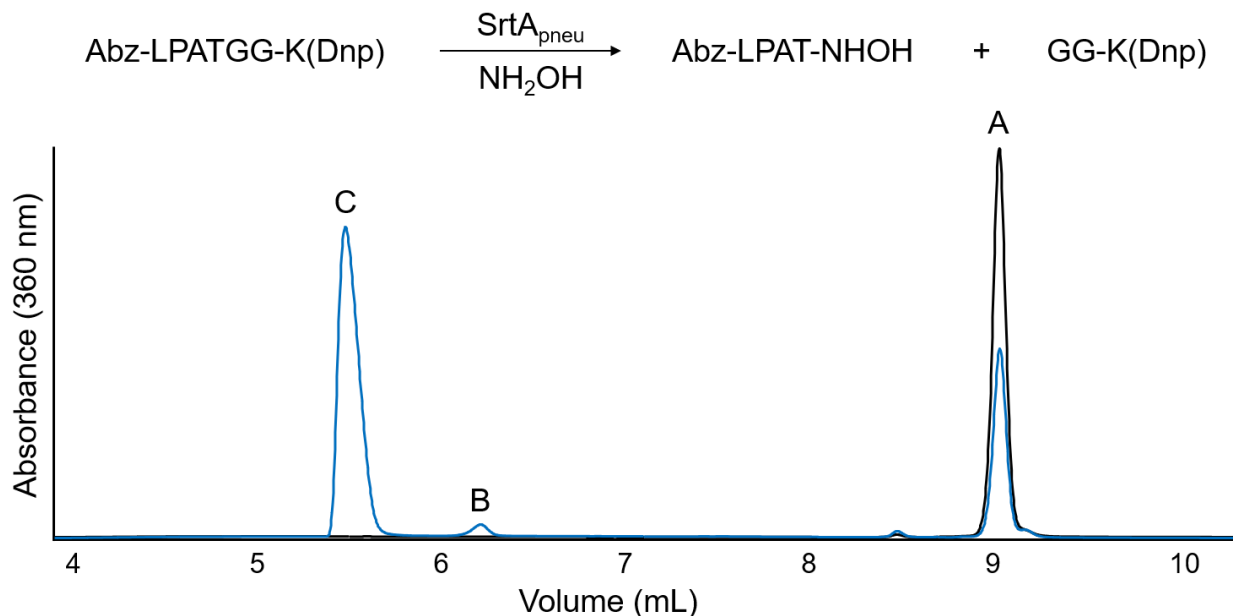


Figure 16. RP-HPLC analysis of model SML reaction using TEV-SrtA_{pneu} at 0 hr (Black) and after 150 minutes (Blue). (A) Abz-LPATGG-K(Dnp) substrate. (B) Abz-LPATG-NHOH product. (C) GG-K(Dnp) excised fragment.

To confirm that monomeric TEV-SrtA_{pneu} was active, our preparation was subjected to a model sortase-mediated ligation (SML) reaction using an Abz-LPATGG-K(Dnp) peptide substrate and a strong H₂NOH nucleophile to assess *in vitro* catalytic activity. Previous work from our lab revealed that monomeric SrtA_{pneu} was catalytically active while dimer was inactive, so we anticipated our enzyme stock to behave accordingly. The reactions were analyzed by reverse phase liquid chromatography (RP-HPLC) every 30 minutes during a 150-minute incubation period at room temperature. The UV/Vis chromatogram for the TEV-SrtA_{pneu} monomer revealed a ~65% conversion of substrate to modified product, which we deemed as sufficiently active compared to the minimal (<5%) product formation of inactive dimer even after a 24-hour incubation (**Figure**

16). Overall, these analyses provided evidence that the monomeric form of TEV-SrtA_{pneu} was indeed an active form of the enzyme, and the form of the enzyme that would be of interest for further structural characterization.

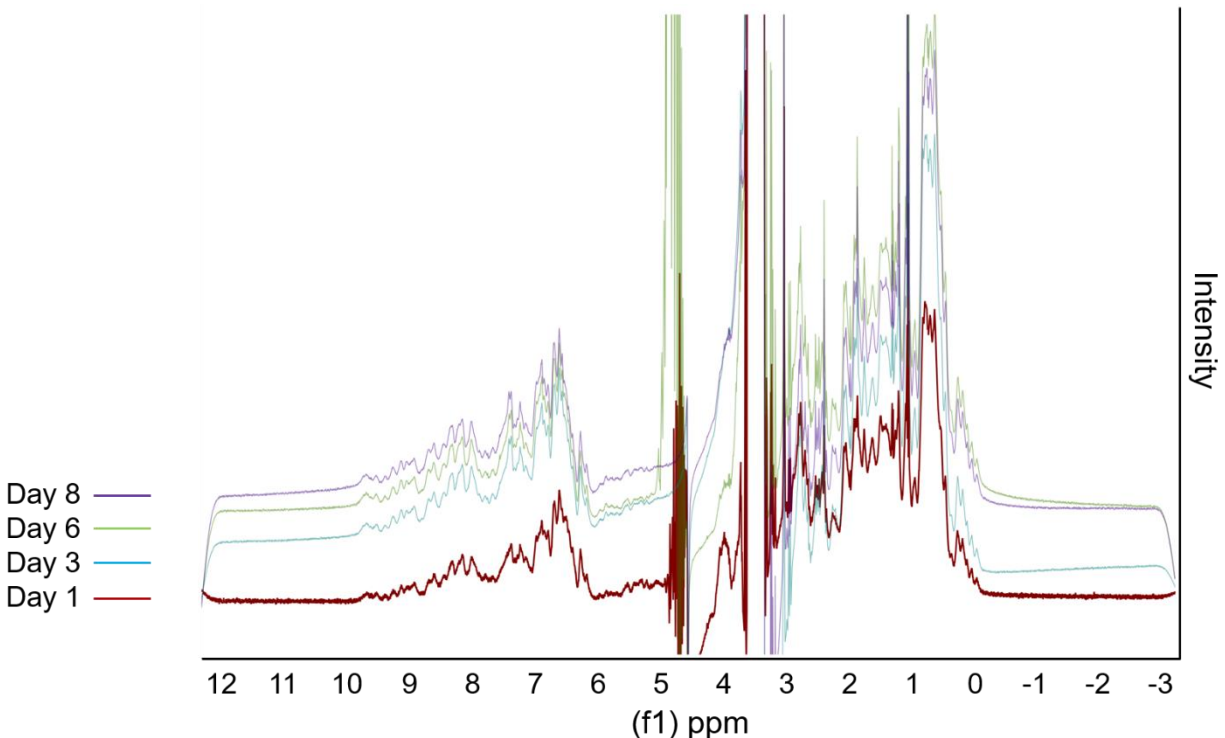


Figure 17. 1D ¹H-NMR analysis of monomeric SrtA_{pneu}-TEV over the course of several days.

A stock of SrtA_{pneu} was concentrated to 450 μM, then used to prepare a ¹H-NMR sample including deuterated water (D₂O, 10% v/v), ethylenediaminetetraacetic acid (EDTA) and sodium azide (NaN₃). This sample was subjected to several rounds of ¹H-NMR analysis (512 scans) over the course of eight days (4 °C) to monitor conformational stability and report signs of degradation (**Figure 17**). The spectral consistency observed among all acquisitions, in particular within the amide N-H region (6-10 ppm) and the aliphatic side chain region (1-3 ppm) suggested that the TEV-SrtA_{pneu} monomer was stable over lengthy periods of time, and presumably not degrading or aggregating into higher order structures.

3.2 Expression and Purification of ^{15}N -labeled TEV-SrtA_{pneu}

Having established that TEV-SrtA_{pneu} remained sufficiently stable in solution, we then began generating a stock of isotopically labelled (^{15}N) enzyme for two-dimensional NMR characterization. A glycerol stock of transformed *E. coli* BL21 (DE3) was used to express ^{15}N TEV-SrtA_{pneu} using a minimal media expression protocol in order to incorporate the ^{15}N label. Briefly, this procedure involves using an overnight seed culture grown in standard LB to initiate a large-scale growth, which should be gently centrifuged after an OD₆₀₀ of 0.5 is reached. The pelleted cells are then resuspended in a wash solution to remove residual nutrient rich media, which must be performed in a timely manner to avoid significant cellular arrest. The cells are then resuspended in minimal media containing 1.5 g of ^{15}N labeled ammonium chloride ($^{15}\text{NH}_4\text{Cl}$). In our hands, initial attempts at expressing ^{15}N -TEV-SrtA_{pneu} revealed poor protein yield, which we speculate was a consequence of inducing expression too late at an OD₆₀₀ of ~ 0.8 or above. By inducing expression in the OD₆₀₀ range of 0.4-0.6 and incubating for 5 hrs at 37 °C, we were able to significantly improve protein yields (**Figure 18**).

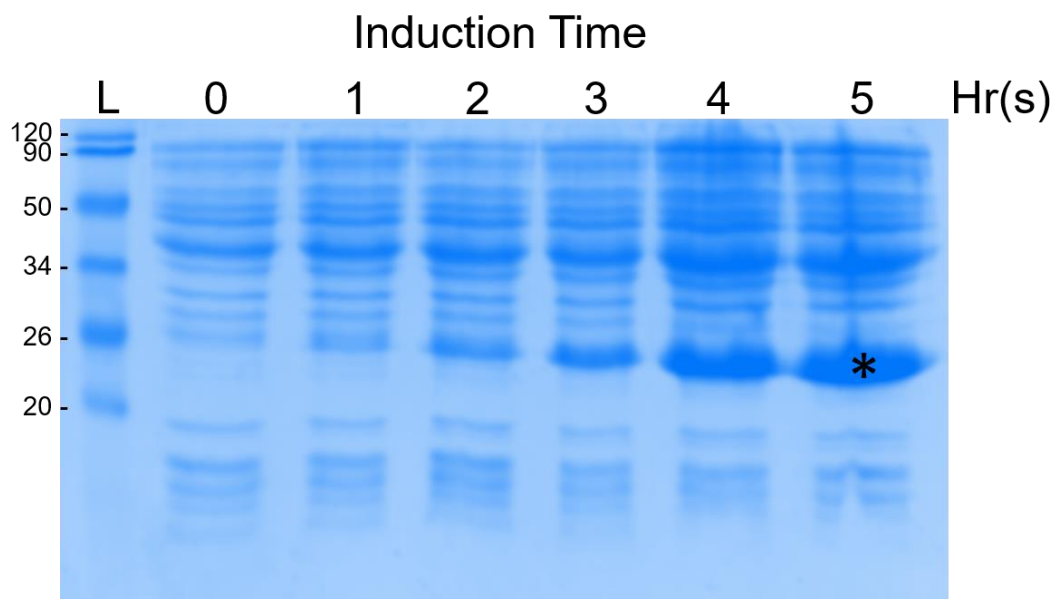


Figure 18. An SDS-PAGE analysis of ^{15}N TEV-SrtA_{pneu}: (L) Protein molecular weight ladder. *Band corresponding to ^{15}N TEV-SrtA_{pneu}.

For optimal production of monomeric ^{15}N -TEV-SrtA_{pneu}, we once again found that a denaturing/refolding protocol was preferred. Thus, as described in section 2.1, cells were first lysed under denaturing conditions (8 M urea). This was followed by denaturing IMAC purification, refolding, and non-denaturing IMAC purification. As shown in below, an SDS-PAGE analysis of the refolded ^{15}N TEV-SrtA_{pneu} following non-denaturing IMAC purification revealed an intense band between the 20 and 26 kDa molecular weight markers, consistent with the formation of the desired ^{15}N -labeled enzyme (**Figure 19**). This SDS gel displayed a significant monomer concentration relative to the residual dimer above. The signature presence of bands representing monomer (bottom) and dimer (top) were also observed on a native-PAGE gel, where the monomer band intensity was significantly more prominent than all other polypeptides present.

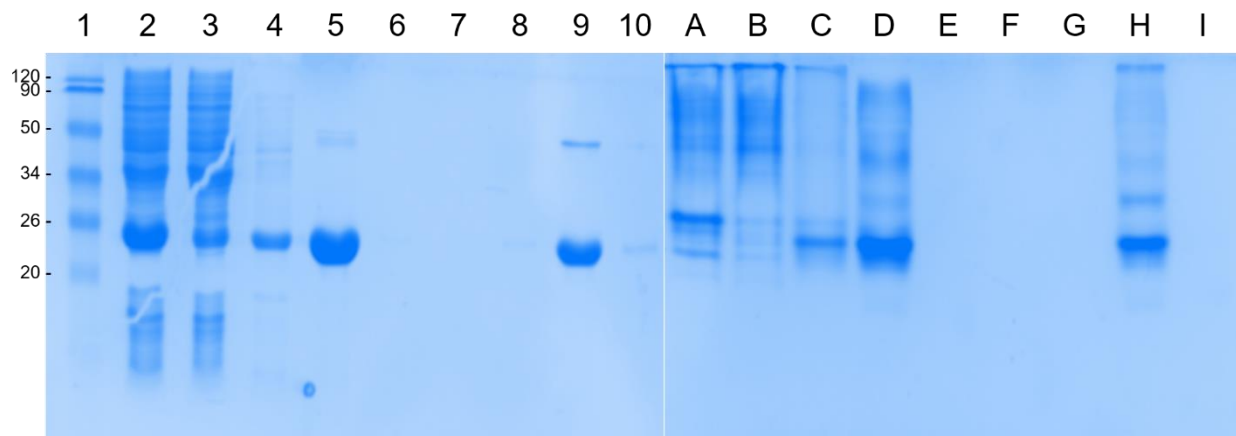


Figure 19. An SDS-PAGE analysis of ^{15}N SrtA_{pneu}-TEV refolding IMAC purification scheme (Left): (1) Protein ladder. (2) Denatured supernatant. (3) Denatured supernatant flow-through. (4) Denaturing wash flow-through. (5) 1st denaturing elution fraction. (6) 2nd denaturing elution fraction. (7) Rapid dilution flow-through. (8) Native wash flow-through. (9) 1st native elution fraction. (10) 2nd native elution fraction. A native-PAGE analysis of ^{15}N SrtA_{pneu}-TEV refolding IMAC purification scheme (Right): (A) Denatured supernatant. (B) Denatured supernatant flow-through. (C) Denaturing wash flow-through. (D) 1st denaturing elution fraction. (E) 2nd denaturing elution fraction. (F) Rapid dilution flow-through. (G) Native wash flow-through. (H) 1st native elution fraction. (I) 2nd native elution fraction.

As a final means of purification, the refolded ^{15}N TEV-SrtA_{pneu} was subjected to SEC to isolate the monomer. The chromatograms from these SEC separations displayed a monomer peak

significantly larger than dimer (left) and higher molecular weight oligomer (leftmost) peaks. Fractions encompassing the right-half of the monomer peak were collected to avoid dimer contamination, then concentrated. These fractions were characterized by ESI-MS to confirm the identity of the protein, as well as to assess the level of ^{15}N incorporation (**Figure 20**).

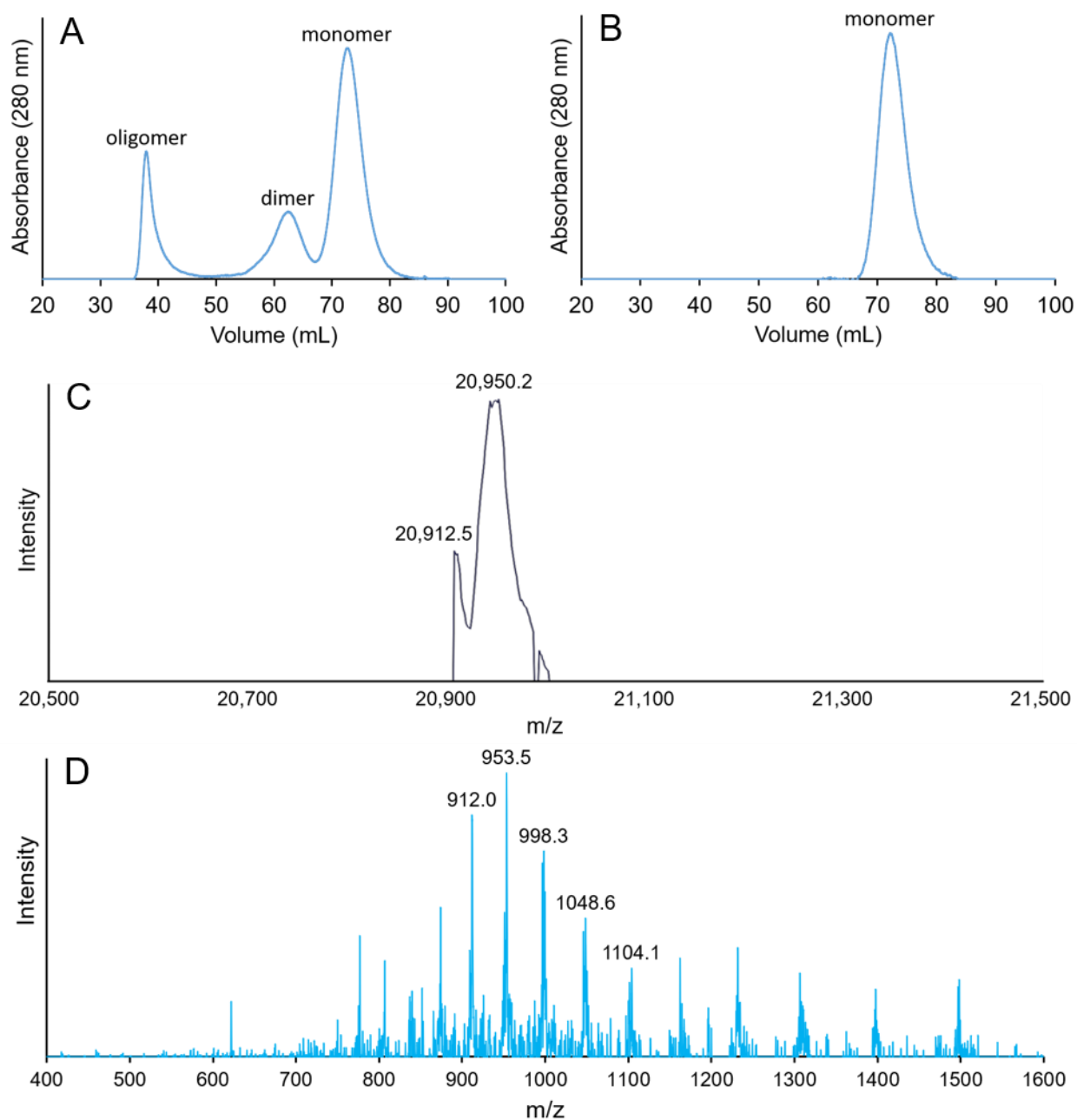


Figure 20. SEC-FPLC traces of IMAC purified ^{15}N SrtA_{pneu}-TEV (A) and collected monomeric fractions (B). A deconvolved mass spectrum (C) of ^{15}N SrtA_{pneu}-TEV generated from the corresponding raw ESI-MS spectrum (D).

This result indicated an ~84% isotope incorporation for the lighter mass peak, but the heavier mass peak is implying an impossible percent incorporation. Mass readouts by our ESI-MS instrument have historically been prone to error when processing biological samples larger than peptides, so the observed discrepancy in mass was disregarded and the 20,950.5 Da mass was assumed to resemble ~100% isotope incorporation of SrtA_{pneu}-TEV. Although the ESI-MS readouts didn't reliably report the extent of SrtA_{pneu}-TEV isotopic labelling, we anticipated this enzyme stock to be sufficient for HSQC NMR.

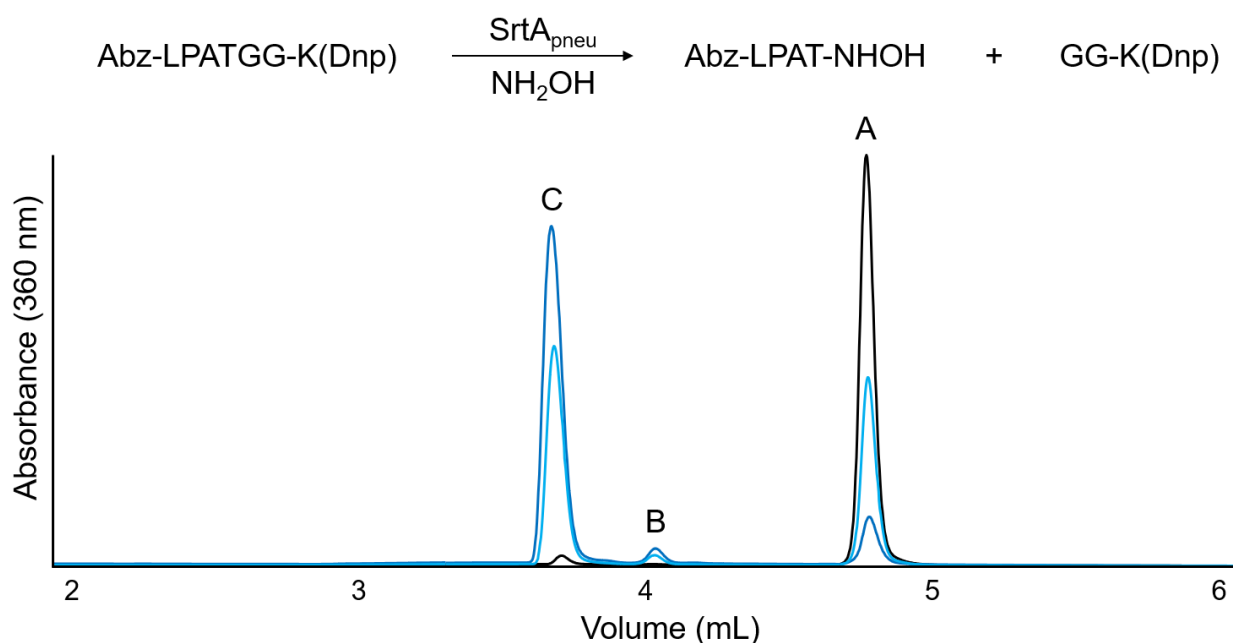


Figure 21. RP-HPLC analysis of model SML reaction using ¹⁵N SrtA_{pneu}-TEV at 0-hr (Black), at 1-hr (Cyan), and after 24-hrs (Marine). (A) Abz-LPATGG-K(Dnp) substrate. (B) Abz-LPATG-NH₂OH product. (C) GG-K(Dnp) excised fragment.

To ensure that the monomeric ¹⁵N TEV-SrtA_{pneu} preparation was active, it was subjected to a model sortase-mediated ligation reaction using an Abz-LPATG-GK(Dnp) peptide substrate and a strong H₂NOH nucleophile. The reaction was analyzed via RP-HPLC after an overnight incubation at room temperature and compared to a control immediately acquired after additional of the enzyme (i.e. time = 0 h). The UV/Vis chromatogram reported an ~82% conversion of substrate to modified product, which is in excellent agreement with previous model reactions using

unlabeled SrtA_{pneu} (**Figure 21**). Based on these results, we were able to determine that incorporating ¹⁵N isotopes into TEV-SrtA_{pneu} did not adversely affect its catalytic activity, which also suggested that its structure was relatively unperturbed.

3.3 2D-HSQC NMR analysis of ¹⁵N-labeled TEV-SrtA_{pneu}

With a purified sample of monomeric ¹⁵N TEV-SrtA_{pneu} in hand, we turned our attention to the acquisition ¹⁵N-HSQC spectra. For our initial sample, ¹⁵N TEV-SrtA_{pneu} was concentrated to 98 μM, and then combined with D₂O (10% v/v), EDTA (0.5 mM) and NaN₃ (0.02% w/v) prior to NMR analysis. The acquired spectrum displayed a number of resonances within the expected chemical shift range along the ¹⁵N axis (100-130 ppm), however the resolution of many signals, particularly within a central cluster of peaks was poor (**Figure 22**). While the lack of resolution made it difficult to discern every signal, we detected ~125 unique resonances, which unfortunately was well below the more than 200 unique cross peaks we had anticipated for the full-length protein. Based on this result, we hypothesized that the resolution of ¹H-¹⁵N couplings could be improved by modulating the acquisition temperature. Therefore, ¹⁵N-HSQC spectra were acquired at 15 and 45 °C. Unfortunately, neither temperature improved the quality of the spectra. Lowering the temperature reduced the resolution of peaks centered within the cluster, and raising the temperature caused the enzyme to precipitate.

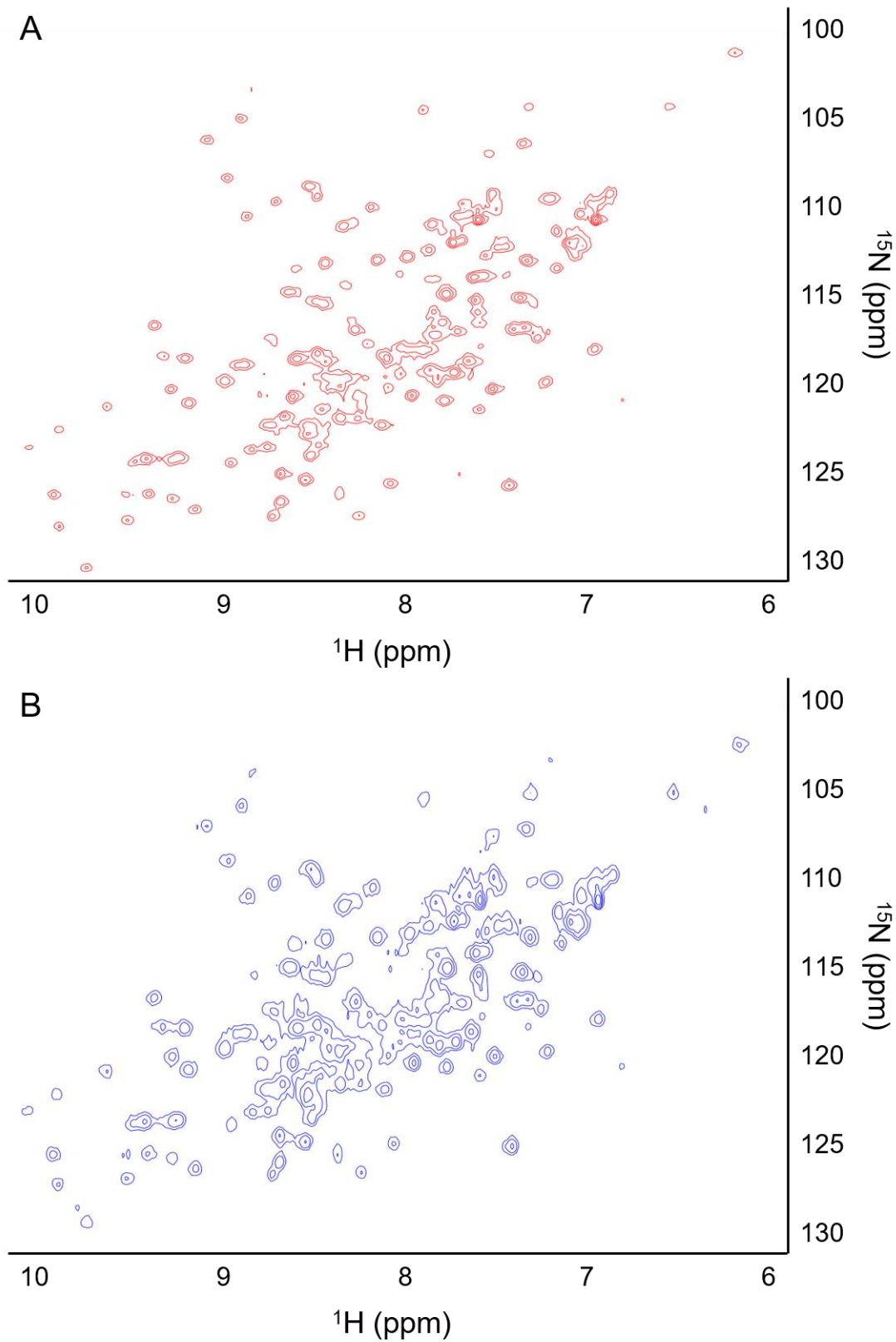


Figure 22. 2D HSQC NMR acquisitions of ^{15}N SrtA_{pneu}-TEV (98 μM) at RT (Red) and 15 $^{\circ}\text{C}$ (Blue).

Next, we attempted to improve the resolution of the ^{15}N -HSQC spectrum by adjusting the enzyme concentration. This idea was based on the possibility that the enzyme may exist in solution as an equilibrium between monomeric and higher order aggregates, and therefore adjusting the concentration may alter the ratio between those species. First, we prepared a 5x diluted sample derived from our original $98\ \mu\text{M}$ enzyme stock. In this case, the NMR was unable to detect any ^1H - ^{15}N signatures, likely due to the lower overall concentration and corresponding reduction in signal-to-noise. We then proceeded to generate a more concentrated ($225\ \mu\text{M}$) stock of ^{15}N TEV-SrtA_{pneu} and were encouraged to see a ^{15}N -HSQC spectrum with smooth and well-defined contours outlining the perimeter of peaks (**Figure 23**). However, peaks in the center of the cluster remained largely undefined, and therefore unsuitable for resonance assignment and structure determination.

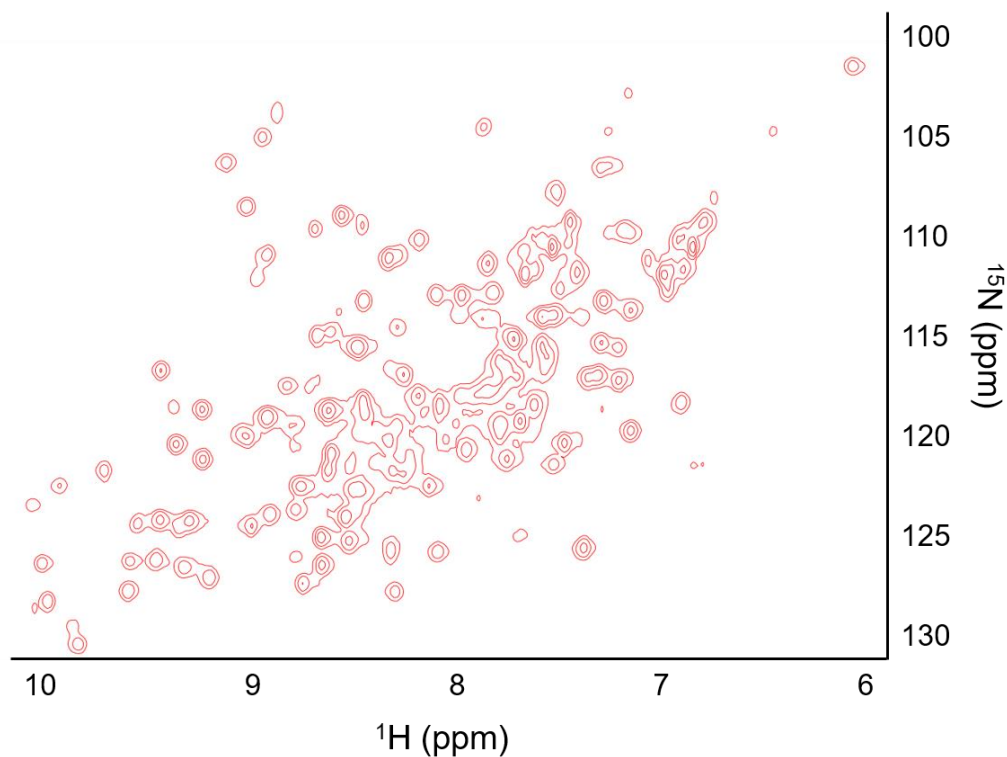


Figure 23. 2D HSQC NMR acquisition of ^{15}N SrtA_{pneu}-TEV ($225\ \mu\text{M}$) at RT.

Given that increasing the enzyme concentration did not entirely alleviate the resolution issues with our spectrum, we speculated that the poorly defined central cluster of peaks may be the result of conformationally labile portions of the enzyme, for example at the N-terminus where a TEV cleavage site and His₆ tag were present. To begin to probe this, we first generated a ¹⁵N-labeled version of SrtA_{pneu} lacking the TEV cleavage site, but retaining the His₆ tag. This protein was prepared and characterized following the same protocol for ¹⁵N TEV-SrtA_{pneu}. A concentrated stock (171 μM) of SrtA_{pneu} without a TEV cleavage site was generated, and an NMR sample was prepared with standard conditions.

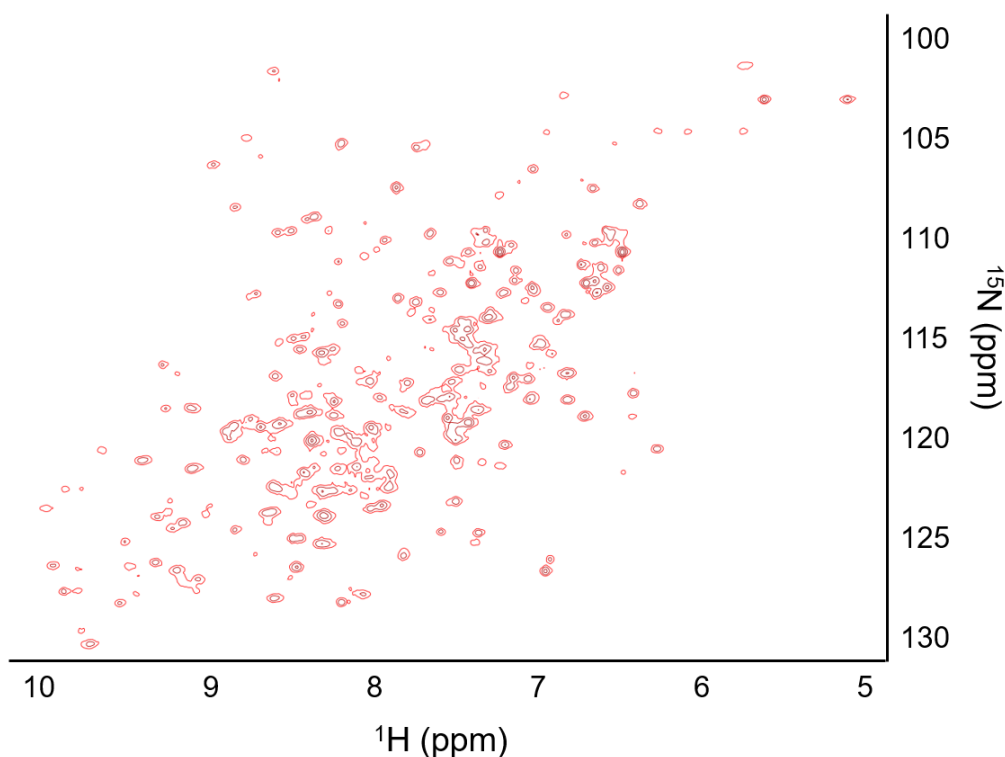


Figure 24. 2D HSQC NMR acquisition of ¹⁵N SrtA_{pneu} (171 μM) at RT.

The acquired ¹⁵N-HSQC spectrum did yield some improvement and unique ~175 peaks were observed, including the appearance of new peaks around the exterior of the peak cluster a more resolved interior peak cluster as compared to ¹⁵N TEV-SrtA_{pneu} (**Figure 24**). While this slight improvement in resolution was encouraging, this sample SrtA_{pneu} lacking a TEV cleavage site still

did not provide sufficient resolution for determining the structure of SrtA_{pneu} as there remained a significant margin of error for discerning individual peaks within the central cluster.

As a final means of probing the impact of the enzyme N-terminus, we also used TEV protease to generate a sample of ¹⁵N SrtA_{pneu} (80.5 μM) lacking both the TEV cleavage site and the His₆ tag. While TEV cleavage was successful, as indicated by ESI-MS, the resulting the ¹⁵N-HSQC spectrum did not provide significant improvements in signal resolution (**Figure 25**).

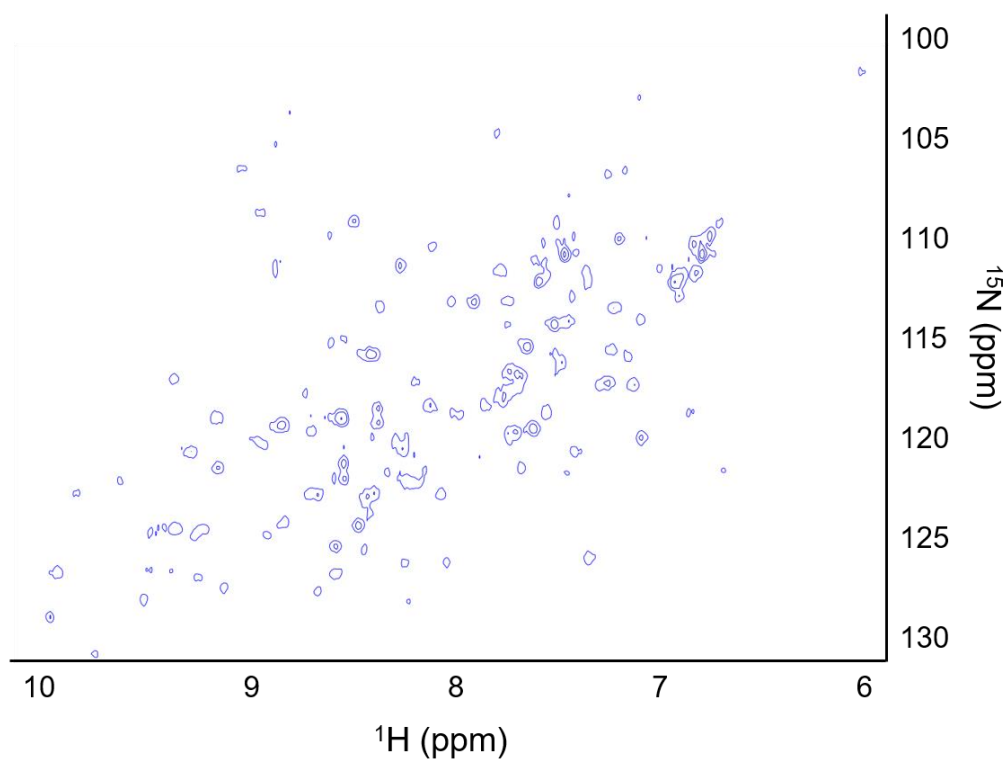


Figure 25. 2D HSQC NMR acquisition of TEV-cleaved ¹⁵N SrtA_{pneu} (80.5 μM) at RT.

Overall, while we have successfully generated a sample of ¹⁵N-labeled enzyme that is active and appears to be monomeric, we have yet to acquire two-dimensional data that is suitable for full resonance assignment and subsequent structure determination. The reasons for this are not entirely clear, however we speculate that portions of SrtA_{pneu} may be conformationally labile in solution, leading to poorly defined regions in the ¹⁵N-HSQC spectrum. Consistent with this

interpretation, we note that the central cluster of poorly defined peaks falls within the range of 110-125 ppm on the ^{15}N axis, which is where disordered regions of proteins are often observed. We also note that regions of disorder have been observed in sortase A homologs in other organisms, for example in the case of sortase A from *S. aureus* where binding of the LPXTG substrate appears to induce a disorder-to-order transition with loops flanking the enzyme active site.

Chapter 4 – Progress Toward Preparation of Substrate-docked derivatives of SrtA_{pneu}

4.1 – Design and Synthesis of ketomethylene isosteres

In addition to determining the structure of SrtA_{pneu} in the absence of substrate, we have also been pursuing strategies for generating enzyme bound to substrate mimetics in order to clearly delineate the interactions between the sorting motif and the enzyme active site. As described in Chapter 1, structures of sortase A from *S. aureus* and *B. anthracis* have been reported in which the enzymes are bound to a substrate analog that mimics the acyl enzyme intermediate. This approach has provided excellent insight into the recognition of the first four residues of the LPXTG sorting motif, however it fails to clearly identify contacts that dictate substrate selectivity involving the 5th position (often G) of the sorting signal. To address this issue, previous work in the Antos lab sought to replace the scissile amide linkage between the 4th and 5th position residues with a non-cleavable carbon-carbon bond. Specifically, a ketomethylene dipeptide isostere (5-aminolevulinic) was incorporated into a peptide substrate in place of the native threonine and glycine residues (**Figure 26**).

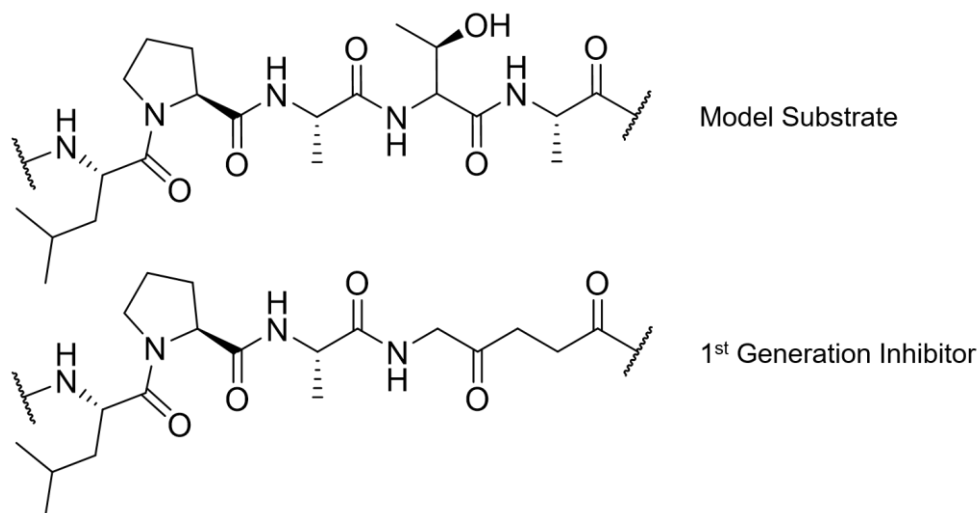


Figure 26. Structural comparison of a model LPATG substrate to a 1st generation inhibitor substrate. The 4th and 5th position residues of the model substrate have been replaced with a ketomethylene linked diglycine.

Preliminary evaluation showed that this substrate was not cleaved by SrtA_{pneu}, and additionally it was able to inhibit enzyme activity was added to a model *in vitro* reaction.⁷⁰ While encouraging, it was recognized that 5-aminolevulinic was not the optimal diketomethylene building block for this approach as it mimicked a diglycine dipeptide structure without any of the relevant amino acid side chains.

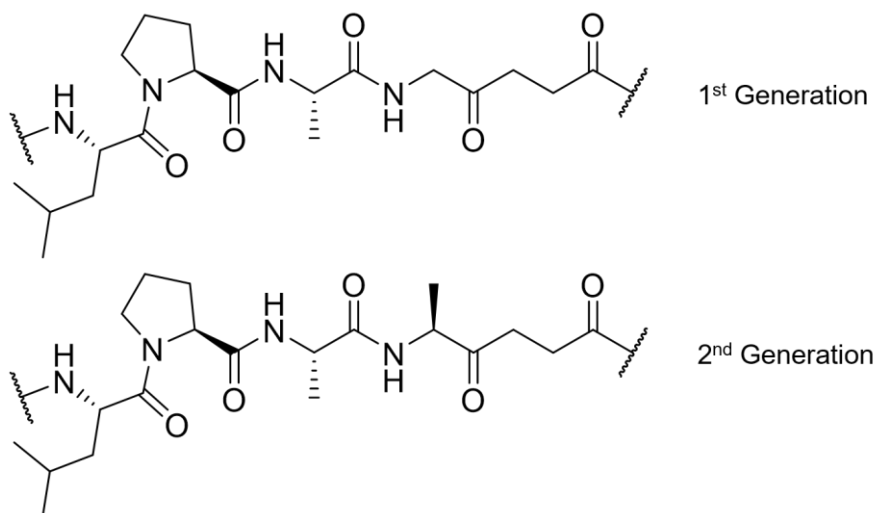


Figure 27. Structural comparison of 1st generation and 2nd generation inhibitor substrates. The 4th position residue of the 1st generation substrate has been replaced with an alanine to serve as an improved mimic of the preferential 4th position threonine.

To address this limitation, we sought to develop a synthetic approach for preparing diketomethylene analogs that included substituents in positions that mirrored natural amino acids. Moreover, we wanted to prepare building that would be compatible with standard solid-phase synthesis techniques. To that end, we designed ketomethylene analog A[keto]G as an initial synthetic target, which would mimic an alanine-glycine dipeptide rather than a glycine-glycine dipeptide (**Figure 27**).

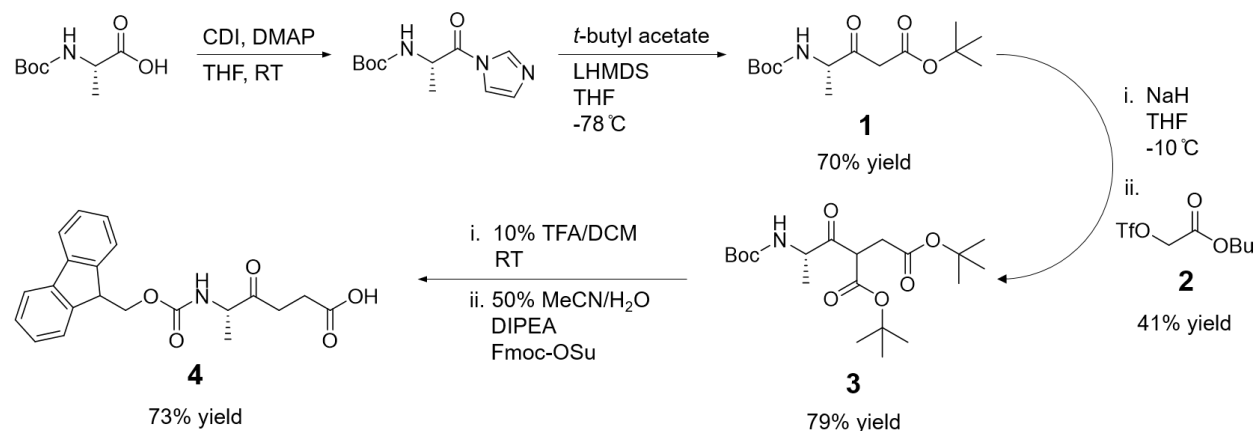


Figure 28. Overview of the proposed synthetic scheme based on procedures from Budnjo et al. and Mathieu et al. using a Boc-protected amino acid starting material.

We adopted a synthetic strategy by Budnjo et al.¹ and Mathieu et al.² in order to produce analog **4** (**Figure 28**). The synthesis began by combining excess lithium enolate of *t*-butyl acetate and carbonyldiimidazole (CDI)-activated Boc-Ala-OH in the presence of a 4-dimethylamino-pyridine (DMAP) catalyst. The resulting Boc-ketoester (**1**) was isolated in 70% yield, and then used as a nucleophile in a stereospecific substitution of a triflate (**2**) derived from *t*-butyl-2-hydroxyacetate. Triflate (**2**) was prepared separately using *t*-butyl-2-hydroxyacetate, triflic anhydride and 2,6-lutidine. Deprotonation of the Boc-ketoester by excess NaH, followed by dropwise addition of **2** resulted in the production of Boc-ketomethylene (**3**) (41% yield). The identity of Boc-ketomethylene (**3**) was confirmed by proton nuclear magnetic resonance spectroscopy (¹H-

NMR), as well as low resolution LC-ESI-MS which reported a molecular weight consistent with the Na²⁺-adduct of the Boc-ketomethylene. Subsequent exposure to trifluoroacetic acid (TFA) provided a means to remove *t*-butyl ester and Boc protecting groups, which was followed by an in-situ decarboxylation to generate an unprotected precursor. This intermediate was not isolated and was immediately reprotected using Fmoc-OSu in the presence of excess diisopropylethylamine (DIPEA). While the formation of the target compound (**4**) was confirmed via low resolution LC-ESI-MS and ¹H-NMR, overall yields were very poor (<7%) and indicated the need for further synthesis optimization.

Table 4. Summary of Boc-Ketomethylene reaction optimization results.

Trial	Base	Solvent	Electrophile	Yield (3)
1	0.140 g NaH	25 mL THF	0.46 g (2)	0.100 g (14% yield)
2	0.655 g KOtBu	25 mL THF	0.46 g (2)	None
3	1.95 g LHMDS	25 mL THF	0.46 g (2)	0.200 g (29% yield)
4	0.140 g NaH	25 mL DMF	0.46 g (2)	None
5	0.126 g NaH	25 mL THF	390 μ L <i>t</i> -Bu-bromoacetate	0.550 g (79% yield)

Optimization of the reaction cascade began by focusing on the synthesis of Boc-ketomethylene (**Table 4**). Alternate bases (KOtBu or LiHMDS) were used in place of NaH, and the replacement of THF with DMF was also attempted. Unfortunately, all variations produced lower yields than the original reaction conditions. We then tried to replace the sensitive *t*-butyl-2-hydroxyacetate triflate (**2**) with the less reactive and commercially available *t*-butyl bromoacetate, and were encouraged to observe significant improvements in Boc-ketomethylene yield.

Next, we sought to investigate the Fmoc reprotection step used to generate the final ketomethylene product (**Figure 29**). Using the Fmoc protection of 5-aminolevulinic acid as a model, we found that replacement of DIPEA with sodium carbonate (Na_2CO_3) and the use of 0.96 eq of Fmoc-OSu improved gave acceptable reaction yields.

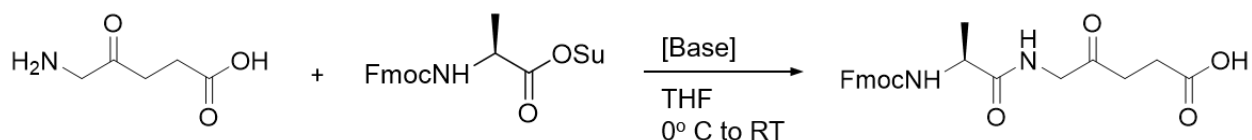


Figure 29. Reaction scheme for synthesizing a G[keto]G building block for SPPS.

Unfortunately, use of these conditions to prepare the final ketomethylene target (A[keto]G) did not significantly increase the amounts we were able to recover, as purification via column chromatography consistently failed to produce entirely pure product, despite variations in eluent, the inclusion of 0.1% acetic acid in the mobile phase, or the use of dry loading techniques.

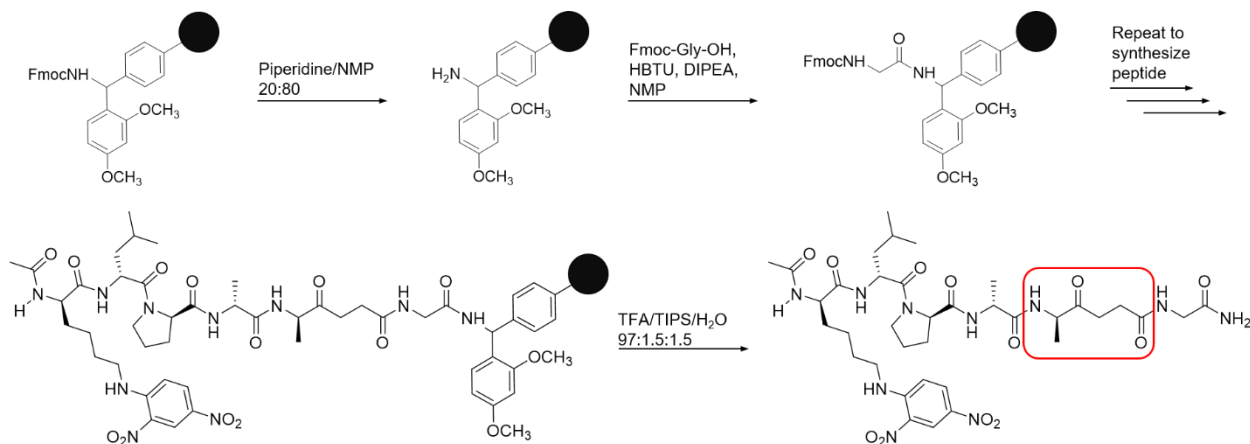


Figure 30. Synthesis of 2nd generation ketomethylene substrate analog using the A(keto)G building block (highlighted) to install a non-cleavable linkage between 4th and 5th residues.

While the synthesis of **4** remained problematic, we felt that some initial proof-of-concept work on its incorporation into a substrate analog were warranted to establish whether any additional synthesis optimization would be worthwhile. Thus, a large-scale SPPS preparation of **4**

beginning with 300 mg of **3** was performed, resulting in the recovery of 200 mg of the final Fmoc-protected product (<73% yield), which was confirmed by both ¹H-NMR and low-resolution LC-ESI-MS. This provided enough material for use in solid-phase synthesis, which was initiated using Rink amide resin and a standard Fmoc-based approach (**Figure 30**). Unfortunately, following cleavage of the crude peptide from the resin, no evidence for the formation of the desired substrate analog would be detected by LC-ESI-MS or RP-HPLC. Overall, this prompted us to abandon this approach and redirect efforts to the alternate substrate analog design described below.

4.2 – Third generation design using disulfide linked analog

Having struggled with synthesis of a ketomethylene-containing substrate analog, we developed an alternate strategy for generating a substrate bound analog of SrtA_{pneu} that relied entirely on standard amino acid residues and standard solid-phase synthetic techniques. This strategy involves replacing the standard LPATG motif with a derivative containing cysteine (C) in place of the native threonine (T) (**Figure 31**).

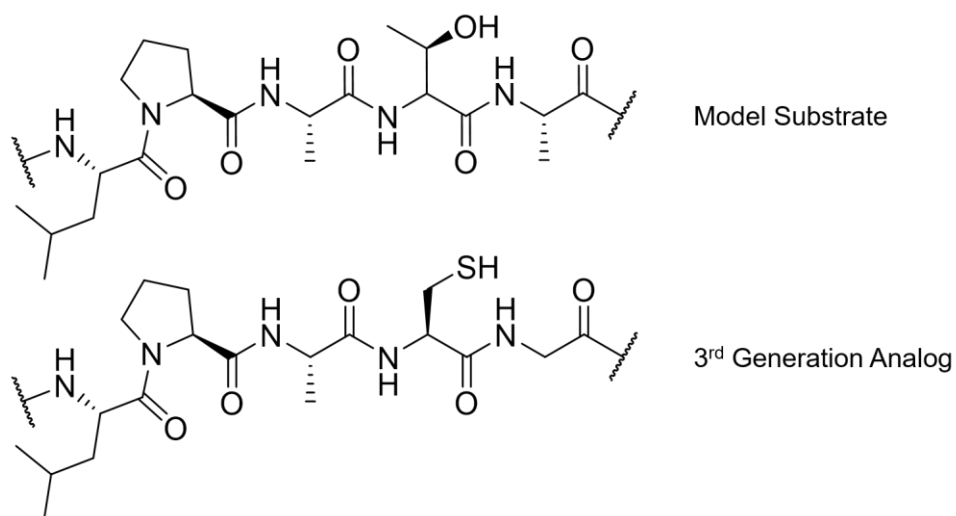


Figure 31. Structural comparison of a model LPATG substrate to a 3rd generation inhibitor substrate. The 4th and 5th position residues of the model substrate have been replaced with cysteine and glycine, respectively.

Subsequently, disulfide bond formation with the enzyme active site would be used to anchor the peptide. While this approach would yield a somewhat artificial enzyme analog that does not exactly reproduce all the contacts between enzyme and substrate, this approach has been reported in the context sortase A from *S. aureus* and shown to provide a means for identifying enzyme residues by solution NMR that are likely to be involved in substrate recognition. Additionally, we hypothesized that the presence of a substrate analog could serve to stabilize the enzyme structure, and alleviate the disorder observed in ^{15}N -HSQC spectra of the free enzyme described in Chapter 3.

Work on this approach began by synthesizing a Bz-GLPACGG peptide using standard solid phase synthesis. The peptide was purified by RP-HPLC, and its identity was confirmed by LC-ESI-MS (**Figure 32**). With the peptide in hand, we then activated the cysteine by conversion to a mixed disulfide using 2,2'-dithiopyridine. Quantitative formation of the mixed disulfide was clearly evident after 30 minutes as determined by LC-ESI-MS, and the product was subsequently isolated by RP-HPLC. Test reactions were then performed in which the thiopyridyl modified peptide was combined TEV-SrtA_{pneu} in various ratios (2-50 equivalents) at room temperature. Using LC-ESI-MS, it was observed that an excess of peptide (50 equiv.) was ideal for rapid and quantitative formation of the desired enzyme adduct (**Figure 33**).

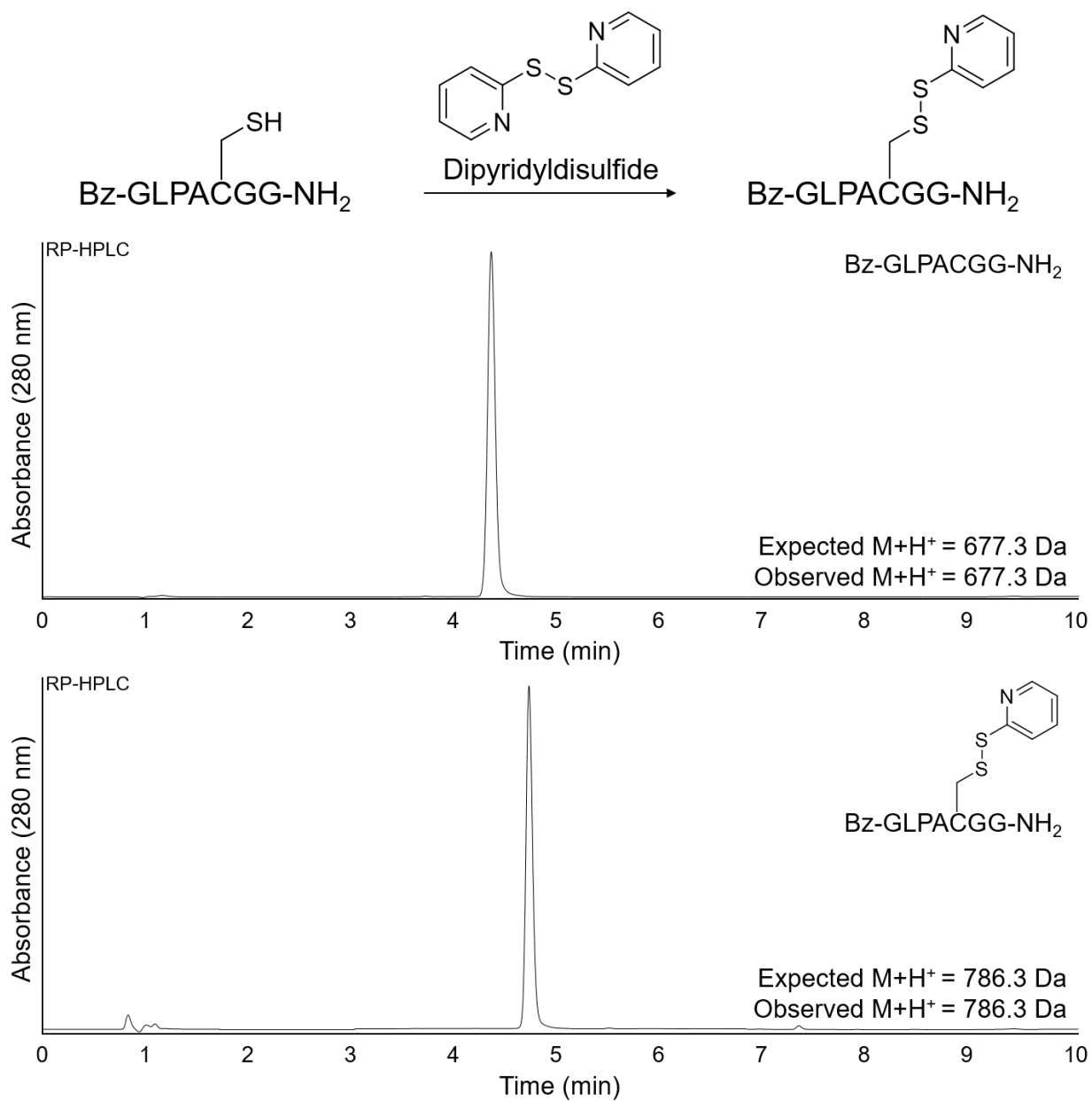


Figure 32. Shown above is a reaction for generating a Bz-GLPACGG peptide appended to a thiopyridine group. RP-HPLC/LC-ESI-MS analysis for determining the purity and molecular weight of Bz-GLPACGG starting material (top) and modified peptide product (bottom).

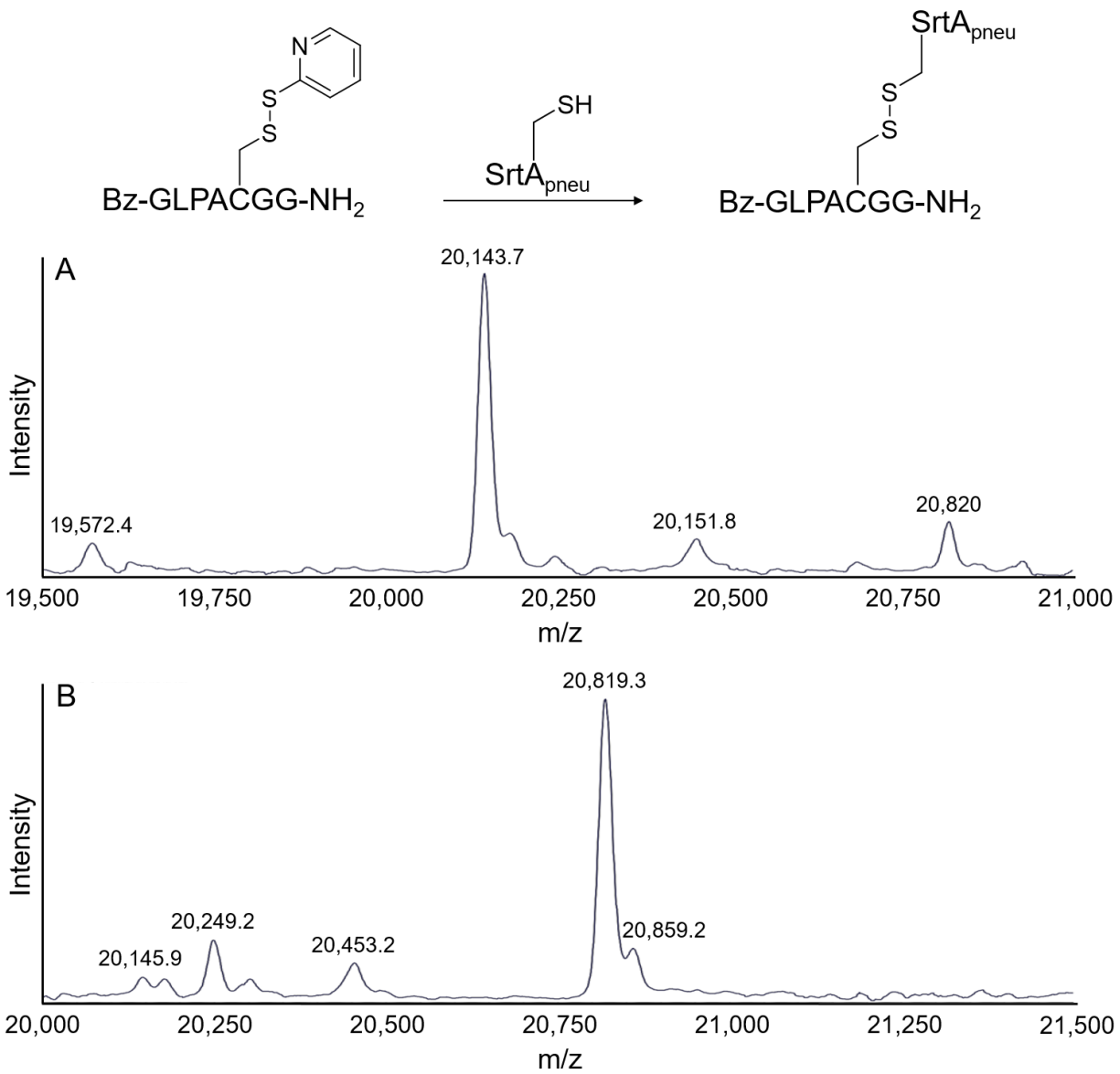


Figure 33. Shown above is a reaction for appending Bz-GLPACGG to the active-site cysteine of SrtA_{pneu}. Deconvolved mass spectra of unmodified SrtA_{pneu} (A) and substrate bound enzyme (B).

Chapter 5 – Conclusion and Future Directions

The results reported in this thesis describe our progress toward characterizing the structure of SrtA_{pneu} bound to a substrate. Chapter 2 detailed the utilization of x-ray crystallography to determine the unbound structure of SrtA_{pneu}, which began by employing methods of reducing the assembled forms of the enzyme by denaturing IMAC purification and isolating the monomer by SE-FPLC. Spin concentrated monomeric preparations were subjected to INDEX and PEG/Ion2 screening kits, which revealed several crystal hits featuring Bis-Tris as a component of each condition. Therefore, we anticipated that the presence of Bis-Tris favored crystal formation and endeavored toward optimizing the screening conditions in an effort to grow diffractable crystals.

Among all of the optimization trials we had attempted, crystal growth appeared to be significantly dependent on pH and PEG 3350 concentration. In particular, our top candidates were grown in conditions with a pH ranging from 4.5-5.5, and concentrations of PEG 3350 between 24-26%. Other optimization trials we had conducted deviated Bis-Tris, or tuning salt, concentrations as well as alternative enzyme to mother liquor hanging drop compositions. We had even tried exploring chloride-containing conditions not included within the preliminary screening kits, but we were still unable to find an optimal crystal condition. Regardless, we sought to diffract the suboptimal crystals generated thus far, however, we were unable to glean any structural information.

In consideration of the theory guiding crystal packing, even the slightest presence of dimeric SrtA_{pneu} may have significantly perturbed unit cell assembly of monomers. This may provide a rationale for our challenges with protein crystallization, where the enzyme stock may be more at fault than the conditions we had employed. To our credit, we prepared the sample to the

best of our ability considering our experiences with dimeric SrtA persisting regardless of denaturing and reducing preparations as well as the poor resolution afforded by our size exclusion column. Taken together, we were unable to solve the structure of unbound SrtA_{pneu} via x-ray crystallography in a timely manner, which propelled our efforts toward utilizing solution NMR as described in Chapter 3.

Chapter 3 detailed our efforts toward characterizing the structure of substrate-bound SrtA_{pneu} by solution NMR, which began by utilizing standard IMAC purification and SEC isolation of monomeric SrtA_{pneu}-TEV. We opted to express a SrtA_{pneu} construct with a TEV cleavage site prior to the N-terminal His₆ tag in the anticipation that its presence may perturb native folding. Initially, we sought to determine if SrtA_{pneu}-TEV is conformationally stable over the course of several days, as we were concerned for the integrity of enzyme stocks moving forward to numerous lengthy HSQC acquisitions. We were confident that SrtA_{pneu}-TEV remained stable based on the spectra consistency observed in the amide region reported by several 1D ¹H-NMR acquisitions.

Next, we redirected our efforts toward expressing our construct in minimal media to isotopically label SrtA_{pneu}-TEV with ¹⁵N for prospective HSQC NMR. The overall success of the isotopically labelled protein expression was largely dependent on the OD₆₀₀ (0.4-0.6) of the minimal media prior to inducing expression with IPTG, and the period of expression (5 hrs). These efforts generated a 98 μM stock of ¹⁵N SrtA_{pneu}-TEV, where HSQC acquisition revealed a cluster of poorly resolved peaks containing approximately 125 peaks. LC-ESI-MS of ¹⁵N SrtA_{pneu}-TEV demonstrated a high degree of isotope incorporation, which was inconsistent with the number of couplings reported by our HSQC spectrum. Based on the advice of our NMR collaborator, we

subjected the 98 μM enzyme stock to 15 $^{\circ}\text{C}$ and 45 $^{\circ}\text{C}$ during acquisition to observe improvements in the resolution. Unfortunately, lowering the temperature did not enhance the spectral resolution, and increasing the temperature caused the enzyme to precipitate out of solution.

Sequential HSQC acquisitions of a stock with a significantly higher concentration resulted in refinement of peaks confined in the perimeter of the cluster, however, the cluster interior remained largely undefined and the number of observable couplings was unchanged. We suspected that insertion of the TEV recognition sequence may have compromised the folding dynamics of the enzyme. Accordingly, we expressed isotopically labeled a SrtA_{pneu} construct lacking a TEV-site. HSQC acquisition of SrtA_{pneu} displayed the appearance of 50 new peaks in addition to those observed previously. Although promising, the removal of the TEV-site did not facilitate detection of all ^1H - ^{15}N couplings implied by preliminary MS data. Furthermore, HSQC acquisition of a TEV cleaved ^{15}N SrtA_{pneu}-TEV expression construct provided no indication of improvement.

Taken together, the evidence suggested that neither His₆ tag nor TEV-site was the direct cause of our resolution dilemma. Rather, we hypothesized the possibility that active unbound SrtA_{pneu} could exist as multiple transient conformations compared to the stable substrate-bound acyl-enzyme intermediate. We began to justify our complications with enzyme crystallization and poor HSQC resolution based on the premise that the *in vitro* behavior of the active enzyme is more dynamic than we had initially anticipated. As discussed in Chapter 4, we sought to design a non-cleavable peptide analog to dock in the active site in order to stabilize the acyl-enzyme conformation and constrain the movement of nuclei during HSQC acquisition for improved resolution.

Chapter 4 encompassed the synthesis of 2nd and 3rd generation peptide inhibitors for their use towards determining the structure of substrate-bound SrtA_{pneu}-TEV via solution NMR. Previous efforts had demonstrated a reasonable capacity for 1st generation LPA[G(keto)G]G inhibitors to compromise the activity of SrtA_{pneu} in the presence of a model substrate and nucleophile. However, the suboptimal efficacy of this proof of concept inhibitor was reflective of missing active site contacts with canonical Thr and Ala residues in the 4th and 5th positions of the substrate, respectively. Attempts to synthesize the 2nd generation LPA[A(keto)G]G peptide analog were founded on prior efforts toward generating a ketomethylene-linked dipeptide isostere mimicking Ala and Gly, which more closely resembled a model LPATAG substrate upon SPPS incorporation.

The original synthesis cascade of an SPPS amenable Fmoc protected [A(keto)G] construct was burdened by a very poor overall yield. The most notable alteration in the procedure that led to improved yields involved the substitution of *t*-Bu-2-hydroxyacetate triflate for *t*-Bu-bromoacetate, which also abrogated the necessity for triflate preparation prior to Boc-ketoester synthesis. Numerous rounds of optimization afforded a sizeable stockpile of Fmoc-[A(keto)G], but we were unable to successfully incorporate it into a peptide via SPPS. Prior to SPPS, we had confirmed the identity of Fmoc-[A(keto)G] by ¹H-NMR and LC-ESI-MS. In theory, Fmoc-[A(keto)G] should behave similarly to our commercially acquired SPPS coupling agents and we've yet to develop an explanation for our observations.

Having struggled with the synthesis of the 2nd generation peptide analog, we developed an alternative 3rd generation substrate design based on standard amino acid residues and standard solid-phase synthetic techniques. This strategy involved replacing the 4th position threonine of a model LPATG substrate with cysteine, where the enzyme active site cysteine was anticipated to form a disulfide bridge upon substrate docking. Therefore, characterization of substrate-bound SrtA_{pneu}-TEV would demonstrate key binding interactions with the 5th position residue, beyond the contacts established by the first four residues reported by Suree et al. (PDB ID: 2KID). A Bz-GLPACGG-NH₂ peptide prepared with a thiopyridine leaving group in the 4th position cysteine was successfully synthesized and demonstrated the capacity to covalently bind to SrtA_{pneu} using excess substrate leaving miniscule traces of unbound enzyme.

Considering that the 3rd generation peptide was capable binding to SrtA_{pneu}, future efforts should be directed toward observing if the peptide is accepted by ¹⁵N SrtA_{pneu}-TEV. Upon confirming the identity of substrate-bound enzyme via LC-ESI-MS, subsequent steps include isolating the substrate-enzyme complex, determining the lifetime, followed by HSQC acquisition. If the appearance of highly resolved (>200) peaks is observed, then our hypothesis pointing blame on enzyme mobility for compromising peak detection and resolution would likely be correct. Lastly, a 3D NMR would be acquired to begin assigning peaks to individual amino acids composing the enzyme primary sequence. Based on this analysis, we may finally assess active site residues prompting key interactions with each position along the substrate.

Regardless of the shortcomings discussed in this thesis, we've been able to establish a foundation for unveiling aspects of SrtA_{pneu} specificity and structural characterization that are not

currently reported in literature. Pursuance of determining the structure of substrate-bound SrtA_{pneu} with our 3rd generation peptide analog may reveal unpublished active site interactions with the 5th position residue. In light of this, we may begin to assemble a complete understanding of SrtA substrate specificity based on published interactions with the first four substrate residues and our newly reported 5th position contacts. As we continue to develop a structure-function relationship for SrtA homologs, we can begin to utilize their unique properties to expand the scope of sortase mediated ligation.

Chapter 6 – Experimental

6.1 Expression of SrtA_{pneu}

The following construct was obtained via commercial gene synthesis from DNA 2.0.

Full sequence of $\Delta 80$ SrtA_{pneu}:

MESSHHHHHHA VLTSQWDAQKLPVIGGIAIPELEMNLP IFKGLDNVNL FYGAGTMKRE
QVMGEGNYSLASHHIFGVDNANKMLFSPLDNAKNGMKIYLTDKNKVYTYEIREVKRVT
PDRVDEVDDR DGVNEITLVTCEDLAATERIIVKGD LKETKDYSQTSDEILTAFNQPYKQF
Y

The following construct was obtained via commercial gene synthesis from ATUM.

Full sequence of $\Delta 80$ SrtA_{pneu}-TEV:

MHHHHHHENLYFQGA VLTSQWDAQKLPVIGGIAIPELEMNLP IFKGLDNVNL FYGAGT
MKREQVMGEGNYSLASHHIFGVDNANKMLFSPLDNAKNGMKIYLTDKNKVYTYEIRE
VKRVTPDRVDEVDDR DGVNEITLVTCEDLAATERIIVKGD LKETKDYSQTSDEILTAFNQ
PYKQFY

Non-Isotopically Labelled SrtA_{pneu} Expression: A 50 uL aliquot of BL21(DE3) cells in 50% glycerol containing the plasmid for SrtA_{pneu} was added to 50 mL of LB broth containing 100 µg/mL ampicillin and incubated with shaking at 37 °C overnight. Roughly 25 mL of culture was then added per 1 L of LB broth containing 100 µg/mL ampicillin to initiate largescale growth. This culture was allowed to grow to an OD₆₀₀ reading of 0.7-0.8 at 37 °C in a shaking incubator (210 RPM) before induction with 1 mL of 1 M IPTG. Cells remained at 37 °C with shaking for at least three hours to express SrtA_{pneu}, and were then isolated by centrifugation at 6000 RPM. Pelleted cells were subsequently stored at -80 °C.

Minimal Media Isotopically Labeled SrtA_{pneu} Expression: A 50 uL aliquot of BL21(DE3) cells in 50% glycerol containing the plasmid for SrtA_{pneu}-TEV-His6 was added per 50 mL of LB broth containing 100 µg/mL kanamycin to initiate a 100 mL seed culture growth, which was incubated with shaking (210 RPM) at 37 °C overnight. Roughly 25 mL of culture was added per 1 L of LB broth containing 100 µg/mL kanamycin to initiate a large scale 4 L growth. This culture was allowed to grow to an OD₆₀₀ reading of 0.4-0.5 at 37 °C in a shaking incubator (210 RPM), and cells were isolated by centrifugation at 4000xg for 20 minutes at 4 °C. Pelleted cells were resuspended in 500 mL of a 1 L salt wash (22 mM Na₂PO₄, 22 mM KH₂PO₄, 8.6 mM NaCl, 1 mM MgCl₂, 2.5 µM B1 Vitamin, 100 µM CaCl₂, 100 ug/mL kanamycin) solution, and cells were isolated by centrifugation at 4000xg for 20 minutes at 4 °C. Pelleted cells were resuspended in 500 mL of minimal growth media (22 mM Na₂PO₄, 22 mM KH₂PO₄, 8.6 mM NaCl, 1 mM MgCl₂, 2.5 µM B1 Vitamin, 100 µM CaCl₂, 100 ug/mL kanamycin, 25 mM D-glucose, 27.5 mM ¹⁵N-NH₄Cl), and the culture was allowed to grow to an OD₆₀₀ reading of 0.6 at 37 °C in a shaking incubator (210 RPM) before induction with 1 mL of 1 M IPTG. Cells remained at 37 °C with shaking for

five hours to express ^{15}N SrtA_{pneu}, and were then isolated by centrifugation at 5000 RPM for 20 minutes at 4 °C. Pelleted cells were subsequently stored at -80 °C.

Native purification: Frozen cell pellets were resuspended in 30 mL denaturing lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA). The resuspended cells were sonicated for two 30 second intervals at 50% power output and the lysate was clarified by centrifugation at 17,500 RPM (Thermo Scientific Fiberlite F20-12x50 LEX rotor). This clarified lysate was added to 5 mL of His-Bind resin (Thermo-Fisher) column pre-equilibrated in denaturing wash buffer (50 mM Tris pH 7.5, 150 mM NaCl, 20 mM imidazole). Bound protein was washed with 10 column volumes of wash buffer and then eluted in two 1 column volume portions of denaturing elution buffer (50 mM Tris pH 7.5, 150 mM NaCl, 300 mM imidazole). Collected fractions were analyzed by native and SDS-PAGE. SrtA_{pneu} monomer was further purified on an NGC FPLC system (Bio-Rad) by size-exclusion chromatography using an Enrich SEC 70 column (Bio-Rad) with running buffer (50 mM Tris pH 7.5, 150 mM NaCl) or a HiPrep 16/60 Sephacryl 200-HR column as the eluent at either 0.2 mL/min or 0.5 mL/min. Monomeric protein fractions were pooled, and if necessary, concentrated using centrifugal concentrators (10 KDa MW cutoff). Samples were stored at 4 °C for temporary storage or -20 °C for long term storage.

Refolding Purification: Frozen cell pellets were resuspended in 30 mL denaturing lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM TCEP, 0.5 mM EDTA, 8 M urea). The resuspended cells were sonicated for two 30 second intervals at 50% power output and the lysate was clarified by centrifugation at 17,500 RPM. This clarified lysate was added to 5 mL of His-Bind resin (Thermo-Fisher) column pre-equilibrated in denaturing wash buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1

mM TCEP, 20 mM imidazole, 8 M urea). Bound protein was washed with 10 column volumes of wash buffer and then eluted in two 1 column volume portions of denaturing elution buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM TCEP, 300 mM imidazole, 8 M urea). The first eluted fraction was then rapidly diluted (100x) by addition to dilution buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM TCEP). This material was then recirculated through a 5 mL Ni-NTA column equilibrated in native wash buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM TCEP, 20 mM imidazole). Bound protein was further washed with 10 column volumes of native wash buffer, then eluted in two 1 column volume aliquots of native elution buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM TCEP, 300 mM imidazole). Collected fractions were analyzed by native and SDS-PAGE. SrtA_{pneu} monomer was further purified on an NGC FPLC system (Bio-Rad) by size-exclusion chromatography using an Enrich SEC 70 column (Bio-Rad) or a HiPrep 16/60 Sephacryl 200-HR column with running buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM TCEP) as the eluent at either 0.2 mL/min or 0.5 mL/min. Monomeric protein fractions were pooled, and if necessary, concentrated using centrifugal concentrators (10 KDa MW cutoff). Samples were stored at 4 °C for temporary storage or -20 °C for long term storage.

Evaluation of protein concentration. UV/Vis spectroscopy for determining concentrations of the prepared samples was performed on a NanodropTM ND1000 spectrophotometer (Thermo Scientific) at 280 nm using 17,420 M⁻¹ cm⁻¹ (SrtA_{pneu}) or 18,910 M⁻¹ cm⁻¹ (SrtA_{pneu}-TEV) as the estimated molar extinction coefficient from analysis of the protein sequence by ExPASy ProtParam.

Protein LC-ESI-MS Analysis. Liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS) was performed using a Dionex Ultimate 3000 HPLC system (Thermo Scientific) connected to an expression^L high performance compact mass spectrometer (Advion, Inc.) through analytical scale separations using a Phenomenex Kinetex 2.6 μm , 100 \AA C4 column (2.0 x 100 mm) with Method B. Data analysis was conducted by Advion Data Express software version 3.0. Mass spectrum deconvolution was achieved through a max entropy algorithm to determine uncharged masses of samples.

6.2 Protein Crystal Preparations & Diffraction

Crystallization of SrtA_{pneu}. Efforts to produce crystals suitable for x-ray diffraction were grown via hanging drop vapor diffusion. All hanging drop loadings were composed of 1:1 mother liquor to enzyme (2 μL drop), unless stated otherwise, using a 6 mg/mL SrtA_{pneu} stock. PEG/ION2 and INDEX screening kits were utilized, and crystal formation was observed under the following conditions: 0.2 M sodium acetate pH 7.0 20% w/v PEG 3350 4 °C (C1), 0.2 M sodium formate pH 7.0 20% w/v PEG 3350 4 °C (C2), 0.2 M sodium malonate pH 6.0 20% w/v PEG 3350 4°C (C3), 0.1 M Bis-Tris pH 5.5 25% w/v PEG 3350 21 °C (C4), 0.1 M Bis-Tris pH 5.5 0.2 M Sodium Chloride 25% PEG 3350 21 °C (C5), 0.1 M Bis-Tris pH 5.5 0.2 M Ammonium Acetate 25% w/v PEG 3350 21 °C(C6), 0.1 M Bis-Tris pH 5.5 0.2 M Magnesium Chloride hexahydrate 25% w/v PEG 3350 21 °C (C7). Room temperature screening conditions (C4-7) were modified by optimizing pH, PEG 3350 and salt concentrations, as well as mother liquor:enzyme drop loading ratio. Optimization efforts around the C4-7 conditions varied pH (pH 4.5-7.5 in steps of 1 pH, pH 4.5-5.5 in steps of 0.2 pH) and PEG concentration (18-28% w/v in steps of 2% w/v). Optimization of C4 (pH 4.5/5.5) varied Bis-Tris concentrations (0.05-0.30 M in steps of 0.05 M), and

optimization of C5-7 (pH 4.5/5.5) varied concentrations of the salts in addition to 0.1 M Bis-Tris (0.05-0.2 M in steps of 0.05 M, and 0.2-0.4 M in steps of 0.1 M), where 1:1 (2 uL drop) and 2:1 (3 uL drop) mother liquor to enzyme ratios were used for C4-7. Furthermore, conditions C4-7 were optimized by varying PEG 3350 concentration (20-30% w/v in steps of 2% w/v) and mother liquor to enzyme ratio (1:1, 1:2, 2:1, 2:2, 1:3, and 3:1). Observed crystal conditions not evaluated in the preliminary screening, yet were anticipated to induce crystal growth, included: 0.1 M Bis-Tris pH 5.5 0.2 M Potassium Chloride 25% w/v PEG 3350 21 °C (C8), 0.1 M Bis-Tris pH 5.5 0.2 M Ammonium Chloride 25% w/v PEG 3350 21 °C (C9), 0.1 M Bis-Tris pH 5.5 0.2 M Calcium Chloride 25% w/v PEG 3350 21 °C (C10). Crystal screening of C8-10 (pH 4.5/5.5) involved varying concentrations of the salts in addition to 0.1 M Bis-Tris (0.05-0.2 M in steps of 0.05 M, and 0.2-0.4 M in steps of 0.1 M).

X-ray Diffraction: Crystals from the following conditions were analyzed via x-ray diffraction: 0.05 M Bis-Tris pH 5.5 25% w/v PEG 3350 21 °C; 0.25 M Bis-Tris pH 5.5 25% w/v PEG 3350 21 °C, 0.05 M Bis-Tris pH 5.5 0.05 M Ammonium Acetate 25% w/v PEG 3350 21 °C; 0.1 M Bis-Tris pH 5.5 0.15 M Magnesium Chloride hexahydrate 25% w/v PEG 3350 21 °C; 0.1 M Bis-Tris pH 5.5 0.05 M Sodium Chloride 25% PEG 3350 21 °C. Crystals looped from these conditions were cryoprotected by washing each crystal with respective crystal inducing conditions with 30% v/v glycerol, followed by immediate flash freezing in liquid nitrogen. All x-ray diffraction data was collected on a Rigaku XtaLAB PRO diffractometer.

6.3 NMR sample Preparation & Acquisition

NMR samples contained 50–300 μM of SrtA_{pneu}-TEV-His6 or SrtA_{pneu}-His6 (unlabeled as well as labeled with ^{15}N), which were all prepared under NMR conditions (10% v/v D₂O, 0.5 mM EDTA, 0.02% w/v NaN₃). NMR spectra were collected with a Bruker Avance spectrometer at 500 MHz for both 1D ^1H and 2D HSQC FID processing, and figure generation was done using Mestrelab MestReNova software version 10.0.2-15465.

6.4 Synthesis of ketomethylene isosteres

All chemicals were obtained from commercial sources and were used without further purification. NMR spectra were collected with a Bruker Avance spectrometer at 500 MHz for ^1H . FID processing and figure generation was done using Mestrelab MestReNova software version 10.0.2-15465. All reactions were performed in flame-dried glassware under argon atmosphere. HPLC purification and analysis was performed using a Dionex Ultimate 3000 HPLC system. LC-ESI-MS was performed with a Dionex Ultimate 3000 HPLC system connected in line to an expression^L high performance compact mass spectrometer (Advion, Inc.). Analytical separations for MS analysis of synthetic products were achieved with a Phenomenex Kinetex 2.6 μm , 100 \AA C18 column (2.1 x 100 mm) with the following method: MeCN (0.1% formic acid) / 95% H₂O, 5% MeCN (0.1% formic acid) mobile phase. Flow rate = 0.3 mL/min. Gradient = 5% MeCN (0.0-0.5 min), 5% MeCN to 90% MeCN (0.5-5.0 min), hold 90% MeCN (5.0-7.0 min), 90% MeCN to 10% MeCN (7.0-7.1 min), re-equilibrate to 10% MeCN (7.1-10.0 min).

tert-butyl 2-(((Trifluoromethyl)sulfonyl)oxy)acetate (**1**). A solution of *t*-butyl 2-hydroxyacetate (0.66 g, 5.0 mmol) in dry DCM (20 mL) was combined with 2,6-lutidine (0.87 mL, 5.0 mmol).

The mixture was cooled to 0 °C and triflic anhydride (1.18 mL, 5.0 mmol) was added dropwise over 70 minutes, during which time the color changed to light red then orange. After stirring for 1 hour at 0 °C, the reaction mixture was diluted with n-hexane (100 mL), washed with 1:3 1 M HCl/sat. NaCl (3x, 50 mL), and dried over MgSO₄. The extract was concentrated by rotary evaporation and dried under vacuum to afford the product as a red/orange oil which was used without further purification (0.71 g, 41% yield). ¹H NMR (500 MHz, CDCl₃): δ 4.80 (s, 2H), 1.54 (s, 9H).

tert-butyl (S)-4-((*tert*-butoxycarbonyl)amino)-3-oxopentanoate (**2**). Boc-Ala-OH (1.32 g, 7.0 mmol) was dissolved in dry THF (20 mL) and then treated with CDI (1.08 g, 7.7 mmol), which was added in three portions while stirring, resulting in bubble formation. Within five minutes of CDI addition, DMAP (26 mg, 0.21 mmol) was added to the reaction mixture. This was left to stir for one hour. In a separate flask, *t*-butyl acetate (4.1 mL, 28.7 mmol) was added dropwise to 1 M LiHMDS (28 mL, 28 mmol) in THF (28 mL) at -78 °C under stirring over the course of ~10 minutes. This reaction was left to stir for 20 min at -78 °C, and then removed from cooling and stirred at room temperature for an additional 10 minutes. The enolate solution was then again cooled to -78 °C and stirred for 20 additional minutes, followed by the dropwise addition of the CDI-activated Boc-Ala-OH over 10 minutes. The combined reaction was allowed to stir for 1.5 hrs at -78 °C before being quenched with 10% w/v citric acid (50 mL). The mixture was extracted with ethyl acetate (2x, 30 mL), washed with sat. NaHCO₃ (30 mL) and sat. NaCl (3x, 30 mL), and then dried over MgSO₄. After concentration by rotary evaporation, the crude product was purified by flash column chromatography (1:3 EtOAc/n-hexane) yielding the product as a white solid (1.38

g, 70% yield). ¹H NMR (500 MHz, CDCl₃): δ 5.15 (m, 1H), 4.38 (m, 1H), 3.46 (q, *J* = 15.6 Hz, 2H), 1.46 (s, 9H), 1.44 (s, 9H), 1.36 (d, *J* = 7.2, 3H).

General procedure for synthesis of di-tert-butyl 2-((tert-butoxycarbonyl)-L-alanyl)succinate (3).

Boc-Ala ketoester (**2**) (0.50 g, 1.74 mmol) was dissolved in dry THF (10 mL) and added dropwise to a stirred suspension of NaH (60% in mineral oil, 0.126 g, 3.2 mmol) in dry THF (10 mL) at -5 °C. This mixture was allowed to stir for 20 min, after which *t*-Bu-bromoacetate (390 μL, 2.64 mmol) was added at -5 °C. The reaction was allowed to stir overnight at room temperature before being quenched with 10% w/v citric acid (15 mL). The quenched reaction was extracted with EtOAc (3x, 30 mL) washed with sat. NaCl (90 mL) and dried over MgSO₄ before being concentrated via rotary evaporation to yield a yellow oil. This residue was purified by flash column chromatography with 1:5 EtOAc/hexane and the desired product fractions identified by TLC were pooled, and concentrated by rotary evaporation (0.55 g, 79% yield). ¹H NMR (500 MHz, CDCl₃) δ 5.24 (dd, *J* = 13.6, 5.6 Hz, 1H), 4.68-4.44 (m, 1H), 4.11 (ddd, *J* = 16.9, 8.2, 6.5 Hz, 1H), 2.88-2.68 (m, 2H), 1.50-1.37 (m, 27H), 1.37-1.33 (m, 3H).

(*S*)-5-((((9H-fluoren-9-yl)methoxy)carbonyl)-amino)-4-oxohexanoic acid (**4**). Compound **3** (0.3 g, 0.747 mmol) was solvated in 10% TFA/DCM (25 mL) and allowed to stir overnight at room temperature. After concentrating the resulting mixture by rotary evaporation, the residue was dissolved in DCM (10 mL) and reconcentrated by rotary evaporation (3x), after which the remaining residue was dried under high vacuum. The vacuum dried residue was then dissolved in 1:1 water/MeCN (15 mL) and DIPEA (0.375 mL, 2.15 mmol). Fmoc-OSu (0.252 g, 0.747 mmol) was then added and allowed to react for 24 hours before the addition of 10 mL of 1 M HCl, which

formed a precipitate. The reaction was extracted into DCM (3x, 30 mL), washed with sat. NaCl (1x, 30 mL) and dried over MgSO₄ before being concentrated under rotary evaporation. The residue was solubilized in 3:1 EtOAc/n-hexane and subjected to flash chromatography using 3:1 EtOAc/n-hexane to purify Fmoc-ketomethylene (**4**) (0.2 g, 73% yield). ¹H NMR (500 MHz, CDCl₃) δ 7.78 (d, *J* = 7.5 Hz, 2H), 7.61 (dd, *J* = 7.6, 3.6 Hz, 2H), 7.45-7.40 (m, 2H), 7.34 (td, *J* = 7.5, 1.2 Hz, 2H), 5.53 (d, *J* = 7.1 Hz, 1H), 4.43 (m, 3H), 4.24 (t, *J* = 6.8 Hz, 1H), 2.86 (m, 2H), 2.72 (m, 2H), 1.41 (d, *J* = 7.2 Hz, 3H). LC-ESI-MS: calculated exact mass 368.14 Da, observed 368.13 Da.

6.5 Peptide Synthesis & Analysis

General procedure for solid-phase peptide synthesis. All chemicals were obtained from commercial sources and were used without further purification. All peptides were synthesized in glass or plastic synthesis vessels. Peptides were synthesized on a 0.1 mmol scale using Rink amide MBHA resin. Deprotection was achieved by washing with 20% piperidine/NMP (10 mL, 2x, 20 min) and was followed by washing with NMP (10 mL, 3x, 10 min). To the deprotected resin, a mixture containing an Fmoc protected amino acid (0.3 mmol), HBTU (0.3 mmol) and DIPEA solvated in NMP was added, which was left to incubate for 1-24 hrs at room temperature with shaking. Unreacted coupling components were removed, and the resin washed with NMP (10 mL, 3x, 10 min) before repetition of this process to couple all amino acids. Where appropriate, acetyl capping of the N-terminus was achieved by combining acetic anhydride (0.3 mmol), DIPEA (0.5 mmol), and NMP (10 mL), which was added to the resin to couple for 2 hrs. Each peptide generated as a substrate for SML reactions contained the 2-aminobenzoyl (Abz) and 2,4-dinitrophenyl (Dnp) fluorophore-quencher pair to simplify analysis by UVVis spectroscopy, where Dnp was conjugated

to the ϵ -amine of a lysine side chain [Fmoc-K(Dnp)-OH]. After completion of the peptide, the resin was washed with DCM (10 mL, 3x, 10 min) and incubated with cleavage solution (9.5 mL TFA, 0.25 mL H₂O, 0.25 mL TIPS) for 30 min (5 mL, 2x). The cleaved peptide was collected and concentrated via rotary evaporation before being precipitated into dry ice-cooled diethyl ether. The precipitate was centrifuged at 4000xg for 5 min and the ether discarded to afford a peptide pellet, which was dried under vacuum for 24 hrs. Peptides were solubilized using a mixture of water and acetonitrile that was variable based on the amino acid composition. Purification from this state was achieved by RP-HPLC with Method A and the molecular weight of the peptides verified via LC-ESI-MS with Method B. Peptides were lyophilized and resolubilized in 1:1 water/DMSO or DMSO to produce stock solutions for use in reactions, which were further analyzed for purity by RP-HPLC analysis using Method B. For peptides containing the Dnp chromophore, concentrations were estimated by UV/Vis spectroscopy on a NanodropTM ND-1000 spectrophotometer (Thermo 80 Scientific) at 365 nm using the molar extinction coefficient 17,300 M⁻¹ cm⁻¹ for the Dnp chromophore.

Activation of Bz-GLPACGG-NH₂ (5). After RP-HPLC purification with Method A and subsequent lyophilization, **5** was combined with 2,2'-dipyridyldisulfide (2x) and solvated in NMP. Purification of the activated product (**6**) from reaction mixture was achieved by RP-HPLC with Method A, and the molecular weight was verified using LC-ESI-MS with Method B. Purified **6** was lyophilized and subsequently resolubilized in 1:10 water/DMSO. An aliquot (1 μ L) of **6** was diluted (100x) in 100 mM DTT and incubated at room temperature for 15-30 minutes. The concentration was estimated by UV/Vis spectroscopy on a NanodropTM ND-1000

spectrophotometer (Thermo 80 Scientific) at 343 nm using the molar extinction coefficient 8,080 M⁻¹ cm⁻¹ for the excised 2-mercaptopyridine chromophore.

HPLC purification and analysis was performed using a Dionex Ultimate 3000 HPLC system. Semi-preparative separations for the purification of peptides were performed with a Phenomenex Luna 5 μm 100 Å C18 column (10 x 250 mm) fitted with a Phenomenex SecurityGuard SemiPrep Guard cartridge (10 mm ID). Purification separations were carried out with the following method: (Method A): MeCN (0.1% formic acid) / 95% H₂O, 5% MeCN (0.1% formic acid) mobile phase. Flow rate = 4.0 mL/min. Gradient = 20% MeCN (0.0-2.0 min), 20% MeCN to 90% MeCN (2.0-15.0 min), hold 90% MeCN (15.0-17.0 min), 90% MeCN to 10% MeCN (17.0-17.01 min), re-equilibrate to 10% MeCN (17.01-19.0 min).

Analytical assessments of peptide purity by UV/Vis, following purification with Method A, were performed with a Phenomenex Kinetex 2.6 μm, 100 Å C18 column (2.1 x 100 mm) with the method (Method B): MeCN (0.1% formic acid) / 95% H₂O, 5% MeCN (0.1% formic acid) mobile phase. Flow rate = 0.3 mL/min. Gradient = 10% MeCN (0.0-0.5 min), 10% MeCN to 90% MeCN (0.5-5.0 min), hold 90% MeCN (5.0-7.0 min), 90% MeCN to 10% MeCN (7.0-7.1 min), reequilibrate to 10% MeCN (7.1-10.0 min).

LC-ESI-MS was performed with a Dionex Ultimate 3000 HPLC system connected inline to an expression^L high performance compact mass spectrometer (Advion, Inc.). Analytical separations for UV/Vis and mass spectrometry analysis were performed with a Phenomenex Kinetex 2.6 μm, 100 Å C18 column (2.1 x 100 mm) with Method B.

6.6 Analysis of Enzyme Transpeptidation Activity

Reactions were prepared by combining all components shown in Table 5. except enzyme, which was added to initiate the reaction. Conversion was analyzed by UV/Vis of analytical RP-HPLC using a Dionex Ultimate 3000 HPLC system (Thermo Scientific) with a Phenomenex Kinetex 2.6 μm , 100 Å C18 column (2.1 x 100 mm) with Method B.

Table 5. Reaction conditions for SML. Water was added to 50 μL total reaction volume unless otherwise stated.

Stock Solution	Reaction Concentration
Buffer (500 mM Tris pH 7.5, 1500 mM NaCl)	50 mM Tris pH 7.5, 150 mM NaCl
Substrate (1:1 water/DMSO) (0.5-100 mM)	200 μM
Nucleophile (water) (100 mM)	10 mM
SrtApneu (50-500 μM)	25 μM

Chapter 7 – Literature Cited

- (1) Rashidian, M.; Dozier, J. K.; Distefano, M. D. Enzymatic Labeling of Proteins: Techniques and Approaches. *Bioconjug. Chem.* **2013**, *24* (8), 1277–1294. <https://doi.org/10.1021/bc400102w>.
- (2) Koniev, O.; Wagner, A. Developments and Recent Advancements in the Field of Endogenous Amino Acid Selective Bond Forming Reactions for Bioconjugation. *Chem Soc Rev* **2015**, *44* (15), 5495–5551. <https://doi.org/10.1039/C5CS00048C>.
- (3) Brannigan, J. A.; Wilkinson, A. J. Protein Engineering 20 Years On. *Nat. Rev. Mol. Cell Biol.* **2002**, *3* (12), 964–970. <https://doi.org/10.1038/nrm975>.
- (4) Lang, K.; Chin, J. W. Bioorthogonal Reactions for Labeling Proteins. *ACS Chem. Biol.* **2014**, *9* (1), 16–20. <https://doi.org/10.1021/cb4009292>.
- (5) Carter, P. Site-Directed Mutagenesis. *Biochem. J.* **1986**, *237* (1), 1–7. <https://doi.org/10.1042/bj2370001>.
- (6) Könnig, D.; Kolmar, H. Beyond Antibody Engineering: Directed Evolution of Alternative Binding Scaffolds and Enzymes Using Yeast Surface Display. *Microb. Cell Factories* **2018**, *17* (1), 32. <https://doi.org/10.1186/s12934-018-0881-3>.

- (7) Caucheteur, D.; Robin, G.; Parez, V.; Martineau, P. Construction of a Synthetic Antibody Gene Library for the Selection of Intrabodies and Antibodies. In *Phage Display*; Hust, M., Lim, T. S., Eds.; Springer New York: New York, NY, 2018; Vol. 1701, pp 239–253. https://doi.org/10.1007/978-1-4939-7447-4_12.
- (8) Lee, T. C.; Kang, M.; Kim, C. H.; Schultz, P. G.; Chapman, E.; Deniz, A. A. Dual Unnatural Amino Acid Incorporation and Click-Chemistry Labeling to Enable Single-Molecule FRET Studies of P97 Folding. *ChemBioChem* **2016**, *17* (11), 981–984. <https://doi.org/10.1002/cbic.201500695>.
- (9) Lang, K.; Chin, J. W. Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem. Rev.* **2014**, *114* (9), 4764–4806. <https://doi.org/10.1021/cr400355w>.
- (10) Wals, K.; Ovaa, H. Unnatural Amino Acid Incorporation in E. Coli: Current and Future Applications in the Design of Therapeutic Proteins. *Front. Chem.* **2014**, *2*. <https://doi.org/10.3389/fchem.2014.00015>.
- (11) Metildi, C. A.; Kaushal, S.; Luiken, G. A.; Talamini, M. A.; Hoffman, R. M.; Bouvet, M. Fluorescently Labeled Chimeric Anti-CEA Antibody Improves Detection and Resection of Human Colon Cancer in a Patient-Derived Orthotopic Xenograft (PDOX) Nude Mouse Model: In Vivo Colon Cancer Labeling With CEA. *J. Surg. Oncol.* **2014**, *109* (5), 451–458. <https://doi.org/10.1002/jso.23507>.
- (12) de Bruin, B.; Kuhnast, B.; Hinnen, F.; Yaouancq, L.; Amessou, M.; Johannes, L.; Samson, A.; Boisgard, R.; Tavitian, B.; Dollé, F. 1-[3-(2-[¹⁸F]Fluoropyridin-3-Yloxy)Propyl]Pyrrole-2,5-Dione: Design, Synthesis, and Radiosynthesis of a New [¹⁸F]Fluoropyridine-Based Maleimide Reagent for the Labeling of Peptides and Proteins. *Bioconjug. Chem.* **2005**, *16* (2), 406–420. <https://doi.org/10.1021/bc0497463>.
- (13) Ducry, L.; Stump, B. Antibody–Drug Conjugates: Linking Cytotoxic Payloads to Monoclonal Antibodies. *Bioconjug. Chem.* **2010**, *21* (1), 5–13. <https://doi.org/10.1021/bc9002019>.
- (14) Jevš̃evar, S.; Kunstelj, M.; Porekar, V. G. PEGylation of Therapeutic Proteins. *Biotechnol. J.* **2010**, *5* (1), 113–128. <https://doi.org/10.1002/biot.200900218>.
- (15) Dozier, J.; Distefano, M. Site-Specific PEGylation of Therapeutic Proteins. *Int. J. Mol. Sci.* **2015**, *16* (10), 25831–25864. <https://doi.org/10.3390/ijms161025831>.
- (16) Mao, H.; Hart, S. A.; Schink, A.; Pollok, B. A. Sortase-Mediated Protein Ligation: A New Method for Protein Engineering. *J. Am. Chem. Soc.* **2004**, *126* (9), 2670–2671. <https://doi.org/10.1021/ja039915e>.
- (17) Beerli, R. R.; Hell, T.; Merkel, A. S.; Grawunder, U. Sortase Enzyme-Mediated Generation of Site-Specifically Conjugated Antibody Drug Conjugates with High In Vitro and In Vivo Potency. *PLOS ONE* **2015**, *10* (7), e0131177. <https://doi.org/10.1371/journal.pone.0131177>.
- (18) Chen, Q.; Sun, Q.; Molino, N. M.; Wang, S.-W.; Boder, E. T.; Chen, W. Sortase A-Mediated Multi-Functionalization of Protein Nanoparticles. *Chem. Commun.* **2015**, *51* (60), 12107–12110. <https://doi.org/10.1039/C5CC03769G>.
- (19) Antos, J. M.; Chew, G.-L.; Guimaraes, C. P.; Yoder, N. C.; Grotenbreg, G. M.; Popp, M. W.-L.; Ploegh, H. L. Site-Specific N- and C-Terminal Labeling of a Single Polypeptide Using Sortases of Different Specificity. *J. Am. Chem. Soc.* **2009**, *131* (31), 10800–10801. <https://doi.org/10.1021/ja902681k>.

- (20) Roeser, D.; Preusser-Kunze, A.; Schmidt, B.; Gasow, K.; Wittmann, J. G.; Dierks, T.; von Figura, K.; Rudolph, M. G. A General Binding Mechanism for All Human Sulfatases by the Formylglycine-Generating Enzyme. *Proc. Natl. Acad. Sci.* **2006**, *103* (1), 81–86. <https://doi.org/10.1073/pnas.0507592102>.
- (21) Rush, J. S.; Bertozzi, C. R. New Aldehyde Tag Sequences Identified by Screening Formylglycine Generating Enzymes *in Vitro* and *in Vivo*. *J. Am. Chem. Soc.* **2008**, *130* (37), 12240–12241. <https://doi.org/10.1021/ja804530w>.
- (22) Wu, P.; Shui, W.; Carlson, B. L.; Hu, N.; Rabuka, D.; Lee, J.; Bertozzi, C. R. Site-Specific Chemical Modification of Recombinant Proteins Produced in Mammalian Cells by Using the Genetically Encoded Aldehyde Tag. *Proc. Natl. Acad. Sci.* **2009**, *106* (9), 3000–3005. <https://doi.org/10.1073/pnas.0807820106>.
- (23) Fernández-Suárez, M.; Baruah, H.; Martínez-Hernández, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y. Redirecting Lipoic Acid Ligase for Cell Surface Protein Labeling with Small-Molecule Probes. *Nat. Biotechnol.* **2007**, *25* (12), 1483–1487. <https://doi.org/10.1038/nbt1355>.
- (24) Slavoff, S. A.; Liu, D. S.; Cohen, J. D.; Ting, A. Y. Imaging Protein–Protein Interactions inside Living Cells via Interaction-Dependent Fluorophore Ligation. *J. Am. Chem. Soc.* **2011**, *133* (49), 19769–19776. <https://doi.org/10.1021/ja206435e>.
- (25) Beckett, D.; Kovaleva, E.; Schatz, P. J. A Minimal Peptide Substrate in Biotin Holoenzyme Synthetase-Catalyzed Biotinylation. *Protein Sci.* **2008**, *8* (4), 921–929. <https://doi.org/10.1110/ps.8.4.921>.
- (26) Algar, W. R.; Prasuhn, D. E.; Stewart, M. H.; Jennings, T. L.; Blanco-Canosa, J. B.; Dawson, P. E.; Medintz, I. L. The Controlled Display of Biomolecules on Nanoparticles: A Challenge Suited to Bioorthogonal Chemistry. *Bioconjug. Chem.* **2011**, *22* (5), 825–858. <https://doi.org/10.1021/bc200065z>.
- (27) Howarth, M.; Takao, K.; Hayashi, Y.; Ting, A. Y. Targeting Quantum Dots to Surface Proteins in Living Cells with Biotin Ligase. *Proc. Natl. Acad. Sci.* **2005**, *102* (21), 7583–7588. <https://doi.org/10.1073/pnas.0503125102>.
- (28) Ton-That, H.; Liu, G.; Mazmanian, S. K.; Faull, K. F.; Schneewind, O. Purification and Characterization of Sortase, the Transpeptidase That Cleaves Surface Proteins of *Staphylococcus Aureus* at the LPXTG Motif. *Proc. Natl. Acad. Sci.* **1999**, *96* (22), 12424–12429. <https://doi.org/10.1073/pnas.96.22.12424>.
- (29) Novick, R. P. Sortase: The Surface Protein Anchoring Transpeptidase and the LPXTG Motif. *Trends Microbiol.* **2000**, *8* (4), 148–151. [https://doi.org/10.1016/S0966-842X\(00\)01741-8](https://doi.org/10.1016/S0966-842X(00)01741-8).
- (30) Ilangovan, U.; Ton-That, H.; Iwahara, J.; Schneewind, O.; Clubb, R. T. Structure of Sortase, the Transpeptidase That Anchors Proteins to the Cell Wall of *Staphylococcus Aureus*. *Proc. Natl. Acad. Sci.* **2001**, *98* (11), 6056–6061. <https://doi.org/10.1073/pnas.101064198>.
- (31) Bradshaw, W. J.; Davies, A. H.; Chambers, C. J.; Roberts, A. K.; Shone, C. C.; Acharya, K. R. Molecular Features of the Sortase Enzyme Family. *FEBS J.* **2015**, *282* (11), 2097–2114. <https://doi.org/10.1111/febs.13288>.

- (32) Dramsi, S.; Trieu-Cuot, P.; Bierne, H. Sorting Sortases: A Nomenclature Proposal for the Various Sortases of Gram-Positive Bacteria. *Res. Microbiol.* **2005**, *156* (3), 289–297. <https://doi.org/10.1016/j.resmic.2004.10.011>.
- (33) Ton-That, H.; Schneewind, O. Anchor Structure of Staphylococcal Surface Proteins: IV. INHIBITORS OF THE CELL WALL SORTING REACTION. *J. Biol. Chem.* **1999**, *274* (34), 24316–24320. <https://doi.org/10.1074/jbc.274.34.24316>.
- (34) Mazmanian, S. K. Staphylococcus Aureus Sortase, an Enzyme That Anchors Surface Proteins to the Cell Wall. *Science* **1999**, *285* (5428), 760–763. <https://doi.org/10.1126/science.285.5428.760>.
- (35) Mazmanian, S. K.; Liu, G.; Jensen, E. R.; Lenoy, E.; Schneewind, O. Staphylococcus Aureus Sortase Mutants Defective in the Display of Surface Proteins and in the Pathogenesis of Animal Infections. *Proc. Natl. Acad. Sci.* **2000**, *97* (10), 5510–5515. <https://doi.org/10.1073/pnas.080520697>.
- (36) Mazmanian, S. K.; Ton-That, H.; Schneewind, O. Sortase-Catalysed Anchoring of Surface Proteins to the Cell Wall of Staphylococcus Aureus. *Mol. Microbiol.* **2001**, *40* (5), 1049–1057. <https://doi.org/10.1046/j.1365-2958.2001.02411.x>.
- (37) Marraffini, L. A.; DeDent, A. C.; Schneewind, O. Sortases and the Art of Anchoring Proteins to the Envelopes of Gram-Positive Bacteria. *Microbiol. Mol. Biol. Rev.* **2006**, *70* (1), 192–221. <https://doi.org/10.1128/MMBR.70.1.192-221.2006>.
- (38) Patti, J. M.; Jonsson, H.; Guss, B.; Switalski, L. M.; Wiberg, K.; Lindberg, M.; Höök, M. Molecular Characterization and Expression of a Gene Encoding a Staphylococcus Aureus Collagen Adhesin. *J. Biol. Chem.* **1992**, *267* (7), 4766–4772.
- (39) Lalioui, L.; Pellegrini, E.; Dramsi, S.; Baptista, M.; Bourgeois, N.; Doucet-Populaire, F.; Rusniok, C.; Zouine, M.; Glaser, P.; Kunst, F.; et al. The SrtA Sortase of Streptococcus Agalactiae Is Required for Cell Wall Anchoring of Proteins Containing the LPXTG Motif, for Adhesion to Epithelial Cells, and for Colonization of the Mouse Intestine. *Infect. Immun.* **2005**, *73* (6), 3342–3350. <https://doi.org/10.1128/IAI.73.6.3342-3350.2005>.
- (40) Signas, C.; Raucci, G.; Jonsson, K.; Lindgren, P. E.; Anantharamaiah, G. M.; Hook, M.; Lindberg, M. Nucleotide Sequence of the Gene for a Fibronectin-Binding Protein from Staphylococcus Aureus: Use of This Peptide Sequence in the Synthesis of Biologically Active Peptides. *Proc. Natl. Acad. Sci.* **1989**, *86* (2), 699–703. <https://doi.org/10.1073/pnas.86.2.699>.
- (41) Vanier, G.; Sekizaki, T.; Domínguez-Punaro, M. C.; Esgleas, M.; Osaki, M.; Takamatsu, D.; Segura, M.; Gottschalk, M. Disruption of SrtA Gene in Streptococcus Suis Results in Decreased Interactions with Endothelial Cells and Extracellular Matrix Proteins. *Vet. Microbiol.* **2008**, *127* (3–4), 417–424. <https://doi.org/10.1016/j.vetmic.2007.08.032>.
- (42) Gómez, M. I.; Lee, A.; Reddy, B.; Muir, A.; Soong, G.; Pitt, A.; Cheung, A.; Prince, A. Staphylococcus Aureus Protein A Induces Airway Epithelial Inflammatory Responses by Activating TNFR1. *Nat. Med.* **2004**, *10* (8), 842–848. <https://doi.org/10.1038/nm1079>.

- (43) Cossart, P.; Jonquieres, R. Sortase, a Universal Target for Therapeutic Agents against Gram-Positive Bacteria? *Proc. Natl. Acad. Sci.* **2000**, *97* (10), 5013–5015. <https://doi.org/10.1073/pnas.97.10.5013>.
- (44) Bierne, H.; Mazmanian, S. K.; Trost, M.; Pucciarelli, M. G.; Liu, G.; Dehoux, P.; the European Listeria Genome Consortium; Jansch, L.; Portillo, F. G.; Schneewind, O.; et al. Inactivation of the SrtA Gene in *Listeria Monocytogenes* Inhibits Anchoring of Surface Proteins and Affects Virulence. *Mol. Microbiol.* **2002**, *43* (4), 869–881. <https://doi.org/10.1046/j.1365-2958.2002.02798.x>.
- (45) Garandeau, C. The Sortase SrtA of *Listeria Monocytogenes* Is Involved in Processing of Internalin and in Virulence. *Infect. Immun.* **2002**, *70* (3), 1382–1390. <https://doi.org/10.1128/IAI.70.3.1382-1390.2002>.
- (46) Chen, S.; Paterson, G. K.; Tong, H. H.; Mitchell, T. J.; DeMaria, T. F. Sortase A Contributes to Pneumococcal Nasopharyngeal Colonization in the Chinchilla Model. *FEMS Microbiol. Lett.* **2005**, *253* (1), 151–154. <https://doi.org/10.1016/j.femsle.2005.09.052>.
- (47) Paterson, G. K.; Mitchell, T. J. The Role of *Streptococcus Pneumoniae* Sortase A in Colonisation and Pathogenesis. *Microbes Infect.* **2006**, *8* (1), 145–153. <https://doi.org/10.1016/j.micinf.2005.06.009>.
- (48) Maresso, A. W.; Schneewind, O. Sortase as a Target of Anti-Infective Therapy. *Pharmacol. Rev.* **2008**, *60* (1), 128–141. <https://doi.org/10.1124/pr.107.07110>.
- (49) Theile, C. S.; Witte, M. D.; Blom, A. E. M.; Kundrat, L.; Ploegh, H. L.; Guimaraes, C. P. Site-Specific N-Terminal Labeling of Proteins Using Sortase-Mediated Reactions. *Nat. Protoc.* **2013**, *8* (9), 1800–1807. <https://doi.org/10.1038/nprot.2013.102>.
- (50) Petrache, A. I.; Machin, D. C.; Williamson, D. J.; Webb, M. E.; Beales, P. A. Sortase-Mediated Labelling of Lipid Nanodiscs for Cellular Tracing. *Mol. Biosyst.* **2016**, *12* (6), 1760–1763. <https://doi.org/10.1039/C6MB00126B>.
- (51) Witte, M. D.; Cragolini, J. J.; Dougan, S. K.; Yoder, N. C.; Popp, M. W.; Ploegh, H. L. Preparation of Unnatural N-to-N and C-to-C Protein Fusions. *Proc. Natl. Acad. Sci.* **2012**, *109* (30), 11993–11998. <https://doi.org/10.1073/pnas.1205427109>.
- (52) Tanaka, T.; Yamamoto, T.; Tsukiji, S.; Nagamune, T. Site-Specific Protein Modification on Living Cells Catalyzed by Sortase. *ChemBioChem* **2008**, *9* (5), 802–807. <https://doi.org/10.1002/cbic.200700614>.
- (53) Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. Sortagging: A Versatile Method for Protein Labeling. *Nat. Chem. Biol.* **2007**, *3* (11), 707–708. <https://doi.org/10.1038/nchembio.2007.31>.
- (54) van 't Hof, W.; Maňásková, S. H.; Veerman, E. C. I.; Bolscher, J. G. M. Sortase-Mediated Backbone Cyclization of Proteins and Peptides. *Biol. Chem.* **2015**, *396* (4). <https://doi.org/10.1515/hsz-2014-0260>.
- (55) Haridas, V.; Sadanandan, S.; Dheepthi, N. U. Sortase-Based Bio-Organic Strategies for Macromolecular Synthesis. *ChemBioChem* **2014**, *15* (13), 1857–1867. <https://doi.org/10.1002/cbic.201402013>.

- (56) Ritzefeld, M. Sortagging: A Robust and Efficient Chemoenzymatic Ligation Strategy. *Chem. - Eur. J.* **2014**, *20* (28), 8516–8529. <https://doi.org/10.1002/chem.201402072>.
- (57) Schmohl, L.; Schwarzer, D. Sortase-Mediated Ligations for the Site-Specific Modification of Proteins. *Curr. Opin. Chem. Biol.* **2014**, *22*, 122–128. <https://doi.org/10.1016/j.cbpa.2014.09.020>.
- (58) Voloshchuk, N.; Liang, D.; Liang, J. Sortase A Mediated Protein Modifications and Peptide Conjugations. *Curr. Drug Discov. Technol.* **2016**, *12* (4), 205–213. <https://doi.org/10.2174/1570163812666150903115601>.
- (59) Proft, T. Sortase-Mediated Protein Ligation: An Emerging Biotechnology Tool for Protein Modification and Immobilisation. *Biotechnol. Lett.* **2010**, *32* (1), 1–10. <https://doi.org/10.1007/s10529-009-0116-0>.
- (60) Matsumoto, T.; Takase, R.; Tanaka, T.; Fukuda, H.; Kondo, A. Site-Specific Protein Labeling with Amine-Containing Molecules Using *Lactobacillus Plantarum* Sortase. *Biotechnol. J.* **2012**, *7* (5), 642–648. <https://doi.org/10.1002/biot.201100213>.
- (61) Guimaraes, C. P.; Witte, M. D.; Theile, C. S.; Bozkurt, G.; Kundrat, L.; Blom, A. E. M.; Ploegh, H. L. Site-Specific C-Terminal and Internal Loop Labeling of Proteins Using Sortase-Mediated Reactions. *Nat. Protoc.* **2013**, *8* (9), 1787–1799. <https://doi.org/10.1038/nprot.2013.101>.
- (62) Chen, I.; Dorr, B. M.; Liu, D. R. A General Strategy for the Evolution of Bond-Forming Enzymes Using Yeast Display. *Proc. Natl. Acad. Sci.* **2011**, *108* (28), 11399–11404. <https://doi.org/10.1073/pnas.1101046108>.
- (63) Piotukh, K.; Geltinger, B.; Heinrich, N.; Gerth, F.; Beyermann, M.; Freund, C.; Schwarzer, D. Directed Evolution of Sortase A Mutants with Altered Substrate Selectivity Profiles. *J. Am. Chem. Soc.* **2011**, *133* (44), 17536–17539. <https://doi.org/10.1021/ja205630g>.
- (64) Schmohl, L.; Bierlmeier, J.; Gerth, F.; Freund, C.; Schwarzer, D. Engineering Sortase A by Screening a Second-Generation Library Using Phage Display: Directed Evolution of Sortase A. *J. Pept. Sci.* **2017**, *23* (7–8), 631–635. <https://doi.org/10.1002/psc.2980>.
- (65) Hirakawa, H.; Ishikawa, S.; Nagamune, T. Design of Ca²⁺-Independent *Staphylococcus Aureus* Sortase A Mutants. *Biotechnol. Bioeng.* **2012**, *109* (12), 2955–2961. <https://doi.org/10.1002/bit.24585>.
- (66) Kruger, R. G.; Otvos, B.; Frankel, B. A.; Bentley, M.; Dostal, P.; McCafferty, D. G. Analysis of the Substrate Specificity of the *Staphylococcus Aureus* Sortase Transpeptidase SrtA[†]. *Biochemistry* **2004**, *43* (6), 1541–1551. <https://doi.org/10.1021/bi035920j>.
- (67) Perry, A. M.; Ton-That, H.; Mazmanian, S. K.; Schneewind, O. Anchoring of Surface Proteins to the Cell Wall of *Staphylococcus Aureus*: III. LIPID II IS AN *IN VIVO* PEPTIDOGLYCAN SUBSTRATE FOR SORTASE-CATALYZED SURFACE PROTEIN ANCHORING. *J. Biol. Chem.* **2002**, *277* (18), 16241–16248. <https://doi.org/10.1074/jbc.M109194200>.
- (68) Huang, X.; Aulabaugh, A.; Ding, W.; Kapoor, B.; Alksne, L.; Tabei, K.; Ellestad, G. Kinetic Mechanism of *Staphylococcus Aureus* Sortase SrtA. *Biochemistry* **2003**, *42* (38), 11307–11315. <https://doi.org/10.1021/bi034391g>.

- (69) Dorr, B. M.; Ham, H. O.; An, C.; Chaikof, E. L.; Liu, D. R. Reprogramming the Specificity of Sortase Enzymes. *Proc. Natl. Acad. Sci.* **2014**, *111* (37), 13343–13348. <https://doi.org/10.1073/pnas.1411179111>.
- (70) Nikghalb, K. D.; Horvath, N. M.; Prelesnik, J. L.; Banks, O. G. B.; Filipov, P. A.; Row, R. D.; Roark, T. J.; Antos, J. M. Expanding the Scope of Sortase-Mediated Ligations by Using Sortase Homologues. *ChemBioChem* **2018**, *19* (2), 185–195. <https://doi.org/10.1002/cbic.201700517>.
- (71) Schmohl, L.; Bierlmeier, J.; von Kügelgen, N.; Kurz, L.; Reis, P.; Barthels, F.; Mach, P.; Schutkowski, M.; Freund, C.; Schwarzer, D. Identification of Sortase Substrates by Specificity Profiling. *Bioorg. Med. Chem.* **2017**, *25* (18), 5002–5007. <https://doi.org/10.1016/j.bmc.2017.06.033>.
- (72) Boekhorst, J.; de Been, M. W. H. J.; Kleerebezem, M.; Siezen, R. J. Genome-Wide Detection and Analysis of Cell Wall-Bound Proteins with LPxTG-Like Sorting Motifs. *J. Bacteriol.* **2005**, *187* (14), 4928–4934. <https://doi.org/10.1128/JB.187.14.4928-4934.2005>.
- (73) Kruger, R. G.; Dostal, P.; McCafferty, D. G. Development of a High-Performance Liquid Chromatography Assay and Revision of Kinetic Parameters for the *Staphylococcus Aureus* Sortase Transpeptidase SrtA. *Anal. Biochem.* **2004**, *326* (1), 42–48. <https://doi.org/10.1016/j.ab.2003.10.023>.
- (74) Schleifer, K. H.; Kandler, O. Peptidoglycan Types of Bacterial Cell Walls and Their Taxonomic Implications. *Bacteriol. Rev.* **1972**, *36* (4), 407–477.
- (75) Jacobitz, A. W.; Kattke, M. D.; Wereszczynski, J.; Clubb, R. T. Sortase Transpeptidases: Structural Biology and Catalytic Mechanism. In *Advances in Protein Chemistry and Structural Biology*; Elsevier, 2017; Vol. 109, pp 223–264. <https://doi.org/10.1016/bs.apcsb.2017.04.008>.
- (76) Jacobitz, A. W.; Wereszczynski, J.; Yi, S. W.; Amer, B. R.; Huang, G. L.; Nguyen, A. V.; Sawaya, M. R.; Jung, M. E.; McCammon, J. A.; Clubb, R. T. Structural and Computational Studies of the *Staphylococcus Aureus* Sortase B-Substrate Complex Reveal a Substrate-Stabilized Oxyanion Hole. *J. Biol. Chem.* **2014**, *289* (13), 8891–8902. <https://doi.org/10.1074/jbc.M113.509273>.
- (77) Frankel, B. A.; Kruger, R. G.; Robinson, D. E.; Kelleher, N. L.; McCafferty, D. G. *Staphylococcus Aureus* Sortase Transpeptidase SrtA: Insight into the Kinetic Mechanism and Evidence for a Reverse Protonation Catalytic Mechanism[†]. *Biochemistry* **2005**, *44* (33), 11188–11200. <https://doi.org/10.1021/bi050141j>.
- (78) Suree, N.; Liew, C. K.; Villareal, V. A.; Thieu, W.; Fadeev, E. A.; Clemens, J. J.; Jung, M. E.; Clubb, R. T. The Structure of the *Staphylococcus Aureus* Sortase-Substrate Complex Reveals How the Universally Conserved LP X TG Sorting Signal Is Recognized. *J. Biol. Chem.* **2009**, *284* (36), 24465–24477. <https://doi.org/10.1074/jbc.M109.022624>.
- (79) Chan, A. H.; Yi, S. W.; Terwilliger, A. L.; Maresso, A. W.; Jung, M. E.; Clubb, R. T. Structure of the *Bacillus Anthracis* Sortase A Enzyme Bound to Its Sorting Signal: A FLEXIBLE AMINO-TERMINAL APPENDAGE MODULATES SUBSTRATE ACCESS. *J. Biol. Chem.* **2015**, *290* (42), 25461–25474. <https://doi.org/10.1074/jbc.M115.670984>.
- (80) Race, P. R.; Bentley, M. L.; Melvin, J. A.; Crow, A.; Hughes, R. K.; Smith, W. D.; Sessions, R. B.; Kehoe, M. A.; McCafferty, D. G.; Banfield, M. J. Crystal Structure of *Streptococcus Pyogenes* Sortase A:

IMPLICATIONS FOR SORTASE MECHANISM. *J. Biol. Chem.* **2009**, *284* (11), 6924–6933.
<https://doi.org/10.1074/jbc.M805406200>.

(81) Khare, B.; Krishnan, V.; Rajashankar, K. R.; I-Hsiu, H.; Xin, M.; Ton-That, H.; Narayana, S. V. Structural Differences between the Streptococcus Agalactiae Housekeeping and Pilus-Specific Sortases: SrtA and SrtC1. *PLoS ONE* **2011**, *6* (8), e22995. <https://doi.org/10.1371/journal.pone.0022995>.

(82) Chambers, C. J.; Roberts, A. K.; Shone, C. C.; Acharya, K. R. Structure and Function of a Clostridium Difficile Sortase Enzyme. *Sci. Rep.* **2015**, *5* (1), 9449. <https://doi.org/10.1038/srep09449>.

(83) Marraffini, L. A.; Ton-That, H.; Zong, Y.; Narayana, S. V. L.; Schneewind, O. Anchoring of Surface Proteins to the Cell Wall of *Staphylococcus Aureus*: A CONSERVED ARGININE RESIDUE IS REQUIRED FOR EFFICIENT CATALYSIS OF SORTASE A. *J. Biol. Chem.* **2004**, *279* (36), 37763–37770.
<https://doi.org/10.1074/jbc.M405282200>.

(84) Kappel, K.; Wereszczynski, J.; Clubb, R. T.; McCammon, J. A. The Binding Mechanism, Multiple Binding Modes, and Allosteric Regulation of *Staphylococcus Aureus* Sortase A Probed by Molecular Dynamics Simulations: Staphylococcus Aureus SrtA Probed by MD Simulations. *Protein Sci.* **2012**, *21* (12), 1858–1871. <https://doi.org/10.1002/pro.2168>.

(85) Moritsugu, K.; Terada, T.; Kidera, A. Disorder-to-Order Transition of an Intrinsically Disordered Region of Sortase Revealed by Multiscale Enhanced Sampling. *J. Am. Chem. Soc.* **2012**, *134* (16), 7094–7101. <https://doi.org/10.1021/ja3008402>.

(86) Pang, X.; Zhou, H.-X. Disorder-to-Order Transition of an Active-Site Loop Mediates the Allosteric Activation of Sortase A. *Biophys. J.* **2015**, *109* (8), 1706–1715. <https://doi.org/10.1016/j.bpj.2015.08.039>.

(87) Suree, N.; Liew, C. K.; Villareal, V. A.; Thieu, W.; Fadeev, E. A.; Clemens, J. J.; Jung, M. E.; Clubb, R. T. The Structure of the *Staphylococcus Aureus* Sortase-Substrate Complex Reveals How the Universally Conserved LP X TG Sorting Signal Is Recognized. *J. Biol. Chem.* **2009**, *284* (36), 24465–24477.
<https://doi.org/10.1074/jbc.M109.022624>.

(88) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* **2015**, *10* (6), 845–858.
<https://doi.org/10.1038/nprot.2015.053>.

(89) Lu, C.; Zhu, J.; Wang, Y.; Umeda, A.; Cowmeadow, R. B.; Lai, E.; Moreno, G. N.; Person, M. D.; Zhang, Z. *Staphylococcus Aureus* Sortase A Exists as a Dimeric Protein In Vitro[†]. *Biochemistry* **2007**, *46* (32), 9346–9354. <https://doi.org/10.1021/bi700519w>.

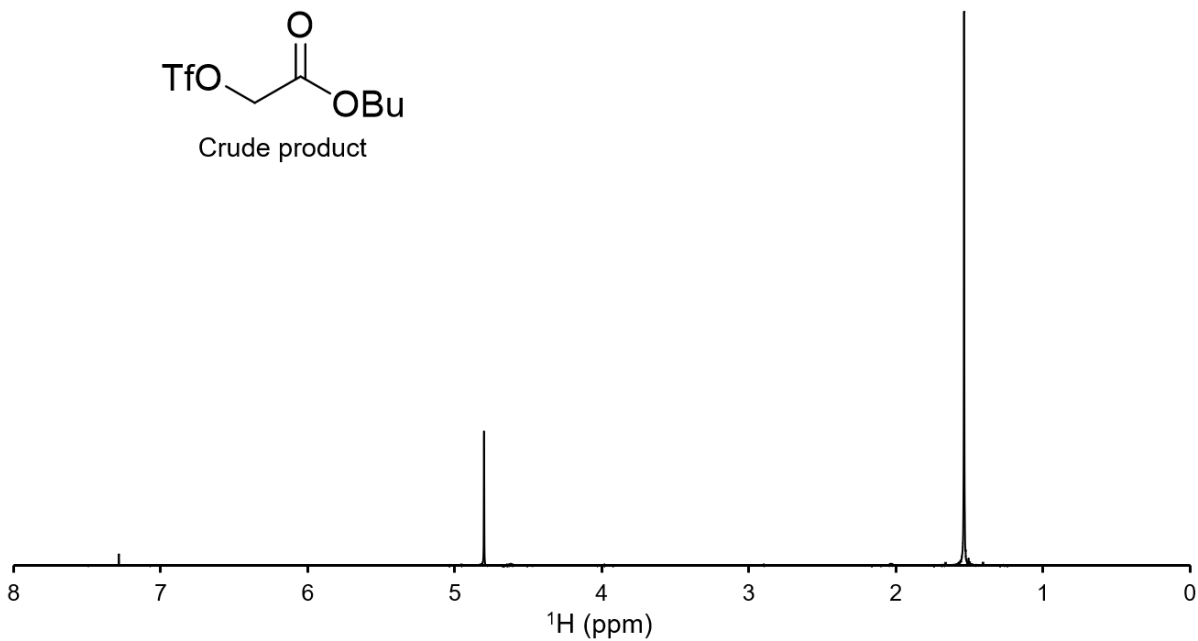
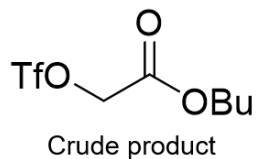
(90) Zhu, J.; Lu, C.; Standland, M.; Lai, E.; Moreno, G. N.; Umeda, A.; Jia, X.; Zhang, Z. Single Mutation on the Surface of *Staphylococcus Aureus* Sortase A Can Disrupt Its Dimerization[†]. *Biochemistry* **2008**, *47* (6), 1667–1674. <https://doi.org/10.1021/bi7014597>.

(91) Zhu, J.; Xiang, L.; Jiang, F.; Zhang, Z. J. Equilibrium of Sortase A Dimerization on *Staphylococcus Aureus* Cell Surface Mediates Its Cell Wall Sorting Activity. *Exp. Biol. Med.* **2016**, *241* (1), 90–100.
<https://doi.org/10.1177/1535370215592122>.

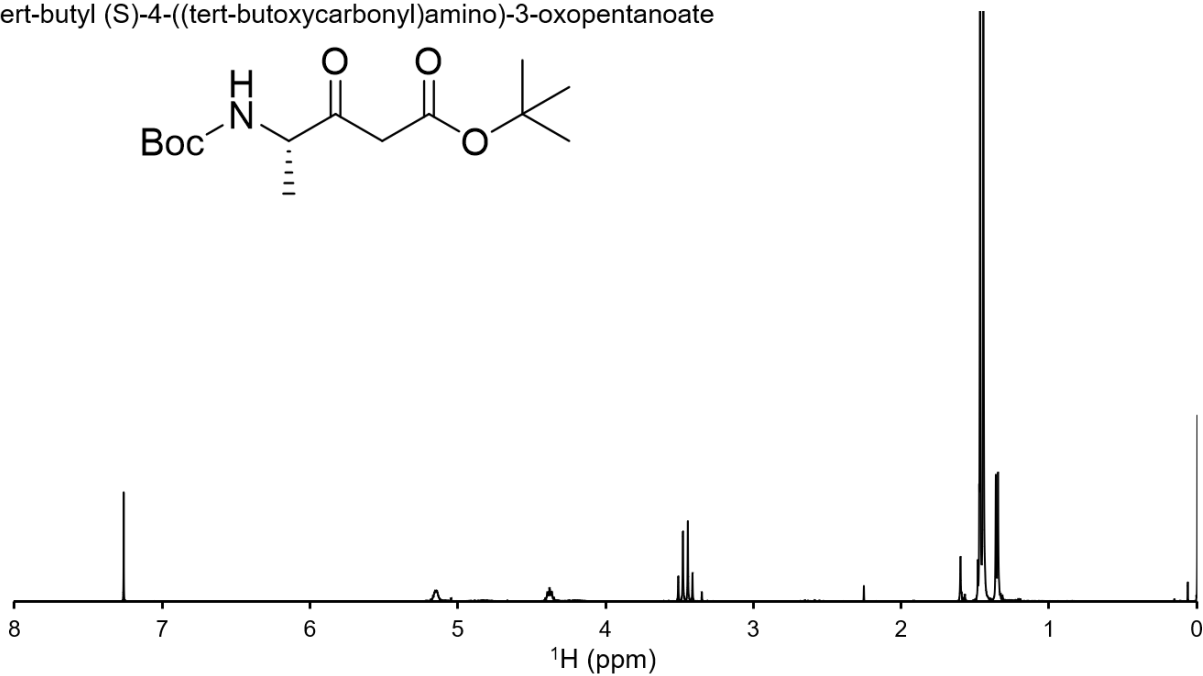
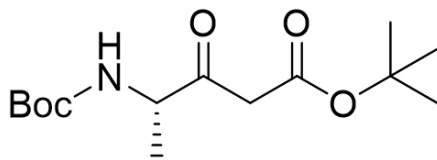
(92) Dekker, N.; Tommassen, J.; Lustig, A.; Rosenbusch, J. P.; Verheij, H. M. Dimerization Regulates the Enzymatic Activity of *Escherichia Coli* Outer Membrane Phospholipase A. *J. Biol. Chem.* **1997**, *272* (6), 3179–3184. <https://doi.org/10.1074/jbc.272.6.3179>.

Chapter 8 – Appendix

tert-butyl 2-(((Trifluoromethyl)sulfonyl)oxy)acetate



tert-butyl (S)-4-((tert-butoxycarbonyl)amino)-3-oxopentanoate



di-tert-butyl 2-((tert-butoxycarbonyl)-L-alanyl)succinate

