Summer 2017

Towards the Substrate-bound Structure of Streptococcus pneumoniae Sortase A

Orion Banks
Western Washington University, bankso@wwu.edu

Towards the Substrate-bound Structure of
Streptococcus pneumoniae Sortase A

By

Orion Banks

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

Kathleen L. Kitto, Dean of the Graduate School

ADVISORY COMMITTEE

Chair, Dr. John Antos

Dr. P. Clint Spiegel

Dr. James Vyvyan

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.

Orion Banks

7/21/2017

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
Orion Banks
July 2017

Abstract

Bacterial sortases have been widely studied for their usefulness in protein modification, however, the variable substrate specificity and activity between homologs of these enzymes is not yet fully characterized. To attempt to further understand sorting signal recognition, we have made advances towards a substrate bound structure of *Streptococcus pneumoniae* sortase A ($SrtA_{pneu}$). This enzyme displays a wide tolerance for alternate amino acids within the canonical LPXTG sorting motif. Our strategy involves a non-cleavable peptide analog that can be docked into the active site, allowing for elucidation of a structure displaying the key contacts that allow the enzyme to recognize alternate sorting signals. To this end, ketomethylene-linked isosteres were designed and synthesized, one of which was incorporated into a peptide via solid phase synthesis to produce a non-cleavable sorting signal for $SrtA_{pneu}$. Preliminary analysis of the substrate analog LPAG(keto)G for inhibition of $SrtA_{pneu}$ activity in a model transpeptidation reaction suggested that this peptide was an effective inhibitor. Work towards understanding the activity of $SrtA_{pneu}$ in relation to its oligomeric state was also undertaken, revealing a strong relationship between the extent of oligomerization and relative activity of $SrtA_{pneu}$, where extensive oligomerization resulted in minimally active samples. Purification of $SrtA_{pneu}$ samples was optimized to produce pure monomeric samples of the enzyme, which showed improved transpeptidation activity. This work has helped lay the foundation for future efforts in producing a substrate-bound structure of $SrtA_{pneu}$.

Acknowledgements

I would like to thank Dr. John Antos for his consistent guidance and positive attitude throughout this project. I have learned and grown more than I could ever have anticipated in my time in his lab, and I am truly thankful for all the knowledge and experience I have gained from his tutelage. I also thank Dr. Serge Smirnov for his guidance during my undergraduate research career, without which I would not have been so well prepared for my work in Dr. Antos' lab.

Additionally, I would like to thank the current and former members of the Antos lab: Keyvan Nikghalb, Jesse Prelesnik, David Brzovic, Nicholas Horvath, Sierra Reed, Natasha Hessami, Ruth Valsquier, and Nathan Markus. I would especially like to thank a number of these people: Keyvan, whose work propelled my project into existence; Nick, who helped with some of the organic synthesis and will be taking over the project presented in this thesis; and Ruth, who was instrumental in developing our improved sortase preparation technique.

I would also like to thank the other members of my committee for their helpful guidance throughout the duration of my project, Dr. P. Clint Spiegel and Dr. James Vyvyan. I also thank Western Washington University and the Research Corporation for Science Advancement for funding this project.

Lastly, I would like to thank Ashley Banks, whose unwavering support has been the most important in my research career.

Table of Contents

List of Tables and Figures

List of Abbreviations and Acronyms

Abz	2-Aminobenzoyl
AcOH	Acetic acid
Boc	*tert*-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
FCC	Flash column chromatography
FGE	Formylglycine generating enzyme
Fmoc	Fluorenylmethyloxycarbonyl
FPLC	Fast protein liquid chromatography
HBTU	*O*-(Benzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High performance liquid chromatography
IMAC	Immobilized metal affinity chromatography

IPTG	Isopropyl-β-D-1-thiogalactopyranoside
LiHMDS	Lithium bis(trimethylsilyl)amide
MBHA	4-Methylbenzhydrylamine
MeCN	Acetonitrile
Ni-NTA	Nickel nitriloacetic acid
NMP	N-methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance
OtBu	*tert*-Butoxy
OSu	N-hydroxysuccinimide ester
PAGE	Polyacrylamide gel electrophoresis
RP	Reverse phase
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chomratography
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
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran

Tris	Tris(hydroxymethyl)aminomethane
UV/Vis	Ultraviolet/visible spectroscopy

1. Introduction

1.1 Protein Engineering and Sortase A

Protein engineering chemistry is a rapidly growing field of research, proven to be broadly applicable to problems in chemistry, molecular biology, and medicine. In many cases, protein engineering seeks to endow proteins with expanded functionality through a variety of permanent modification methods.[1-5] Among the numerous methods available for protein modification, chemoenzymatic systems have seen increasing use, aiding in the production of fluorescent-labeled proteins,[6] antibody-drug conjugates,[7] and bioconjugated nanoparticles,[8] within a growing catalog of useful products. This method typically provides site-specific and rapid attachment of a desired modification, achieved through recognition and alteration of amino acids or sets of amino acids within a protein of interest.[2, 9] Frequently, genetic modification is used to produce a recognition site for an enzyme with the ability to catalyze addition of a desired modification. The range of modifications is vast, however; most enzymes for protein modification function in exactly the same manner: a specific sequence of amino acids is recognized and a modification is made to or within the recognized site (**Figure 1**).

In general, an individual enzyme catalyzes only a single reaction, meaning for each type of modification, a separate chemoenzymatic system exists. Formylglycine generating enzyme (FGE) recognizes a CXPXR amino acid sequence and modifies the cysteine residue to formylglycine co-translationally,[10,

[11] which can be used to generate oximes through reaction with hydroxyamines,[11, 12] a simple click reaction. Lipoic acid ligase attaches lipoic acid to lysine within its recognition sequence. This enzyme has been particularly useful in protein modification, as alteration of its substrates to include the azide and alkyne click handles has allowed for site specific incorporation of these selective and rapidly derivatized groups.[13] This enzyme has also been employed in studies of protein-protein interactions.[14] Biotin ligase also modifies the lysine sidechain within its recognition sequence, attaching a biotin molecule to the ε-amine.[15] Biotin often functions as a site specific tag for binding biomolecules to materials, examples of which include nanoparticles[16] and quantum dots,[17] as this small molecule selectively and tightly binds to proteins avidin and streptavidin.

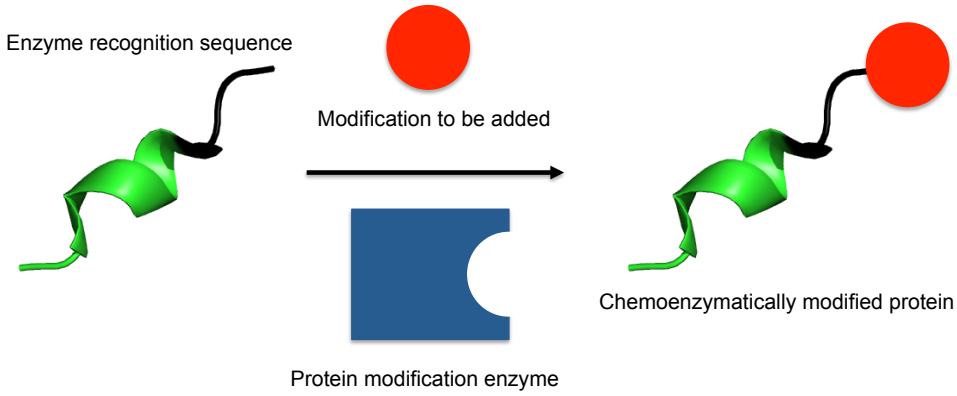


Figure 1. Generalized schematic of chemoenzymatic modification of proteins. A protein with an endogenous or inserted peptide recognition sequence (shown in black) is recognized by the protein modification enzyme. The modification is attached covalently within the recognition sequence.

The sortase enzyme family has been well studied for use in protein modification. Sortases are transpeptidases found in Gram-positive bacteria[18, 19],

all of which maintain a nucleophilic cysteine within their active site that serves as the primary residue for catalysis.[18, 20, 21] Sortases can be divided into multiple classes (A-F)[21, 22], each exhibiting different structural and biochemical traits, some of which are shown in **Table 1**. Of most relevance to this study is the class A sortase (SrtA), which resides on the extracellular membrane of bacteria, embedded via a transmembrane domain[18] that can be removed to produce soluble, recombinant SrtA for use *in vitro. In vivo*, SrtA typically performs a "housekeeping" function by anchoring multiple protein types to the extracellular wall.[18, 19, 23-26] It has been determined that many proteins appended to the cell wall by SrtA are key virulence factors,[27] including collagen adhesion proteins,[28] fibronectin binding proteins,[29-31] and immunoglobulin binders[32] that aide bacterial cell colonization and inhibit the host immune response. This *in vivo* function of SrtA makes it a viable drug target in Gram-positive bacteria, as several studies have shown that Gram-positive bacterial strains without SrtA exhibit reduced virulence.[33-37] This makes it prudent to develop an in-depth understanding of sortase enzymology, as it may benefit the development of novel therapeutics to replace the rapidly failing catalog of current antibiotic drugs.[38] While many advances have been made towards understanding specific homologs of SrtA, other less studied homologs may serve as excellent targets for further analysis as tools and drug targets.[9, 29, 31, 39]

Table 1. Sorting signals, substrates, and species specificity of all known sortase classes. [from Bradshaw *et al*.[21]]

Sortase Class	Motif	Example Substrates	Species
A	LPXTG	Surface proteins[40]	All low GC content Gram-positive bacteria
B	NP(Q/K)TN	Haem acquisition proteins[41]	Low GC content Gram-positive bacilli and cocci
C	(I/L)(P/A)XTG	Pillin subunits[42, 43]	Both low and high GC content Gram-positive bacteria
D	LPNTA	Endospore envelope proteins[44, 45]	*Bacillus* species
E	LAXTG	Pili[46]	High GC content Gram-positive bacteria
F	Unknown	Unknown	*Actinobacteria*

As each class of sortase performs a separate set of functions, each also recognizes its own sorting motif; typically a five amino acid sequence that is accepted into the enzyme active site for modification.[21, 22] Specifically, SrtA homologs typically recognize an LPXTG motif,[18, 19, 23, 24] where X is any amino acid. In proteins containing this motif, it is followed by a string of hydrophobic residues and a highly polar tail, dominated by arginine and lysine residues. This full segment (the sorting motif, hydrophobic, and hydrophilic residues) is referred to as a cell wall sorting signal (CWSS), as it indicates which proteins should be embedded in the cell membrane and targeted for binding, cleavage, and ligation

to the peptidoglycan of the extracellular wall.[23, 47, 48] While sortases vary in terms of their sorting motif preferences, they are all believed to catalyze transpeptidation reactions similiar to that depicted in **Figure 2** for sortase A from *Staphylococcus aureus* ($SrtA_{staph}$). In the case of $SrtA_{staph}$, the enzyme cleaves between the threonine and glycine residues in the sorting motif, forming an acyl-enzyme intermediate that is resolved by an N-terminal amine nucleophile (in *S. aureus*, the lipid II pentaglycine moiety within the peptidoglycan).[26] After transferring the substrate, the enzyme releases the newly fused protein, now appended to the cell wall.

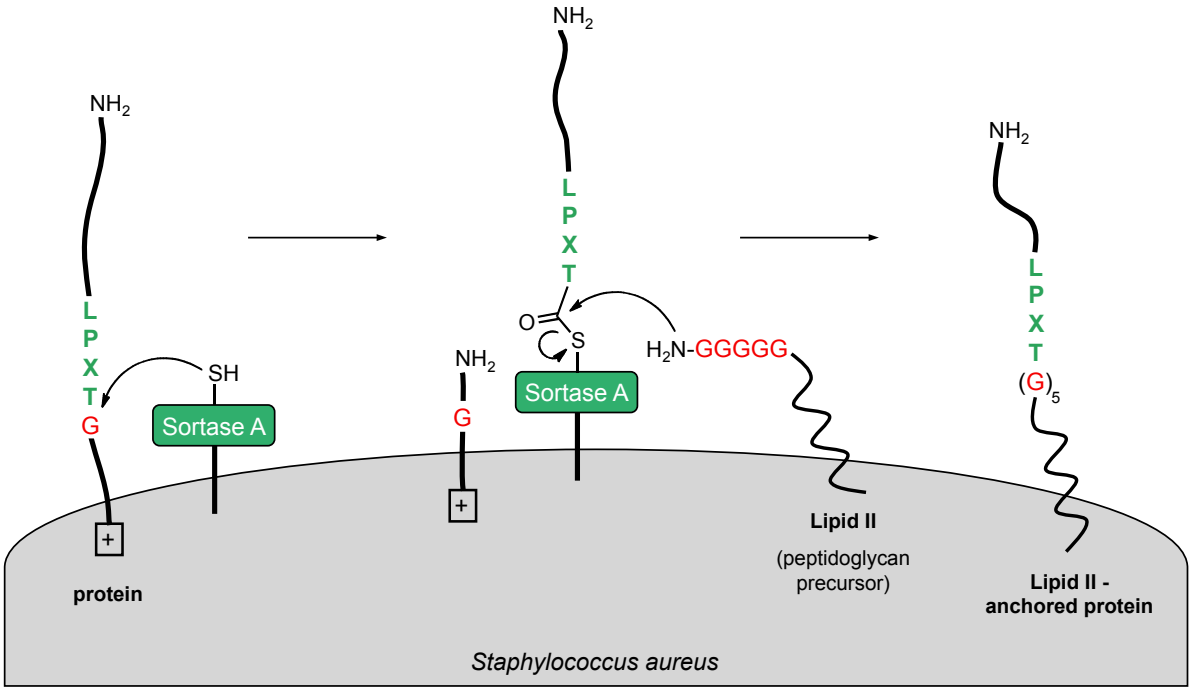


Figure 2. *In vivo* mechanism of transpeptidation by $SrtA_{staph}$.

The innate ability of sortases to catalyze site-specific ligations has been used extensively in protein engineering chemistry to produce a large catalog of proteins and peptides conjugated to fluorophores,[49] nanoparticles,[8] lipid nanodiscs,[50] other proteins,[51] and even live cells[52, 53] (**Figure 3**). The protein or peptide ligation partners can function as either the amine nucleophile or substrate (sorting signal) in the transpeptidation reaction, which further enhances the versatility of this approach.[51, 54-56] This has been used to modify the C- and N-termini of proteins, in addition to surface loops.[51, 56] Several homologs of sortase have been used in protein modification efforts. While most studies are carried out with $SrtA_{staph}$, several modifications have been demonstrated with the SrtA homolog from *Streptococcus pyogenes*[9, 49, 57] ($SrtA_{pyogenes}$) and, more recently, with the evolved mutants of $SrtA_{staph}$, which demonstrate higher reaction rates,[58] altered substrate profiles,[58-60] and Ca^{2+} independence.[61, 62]

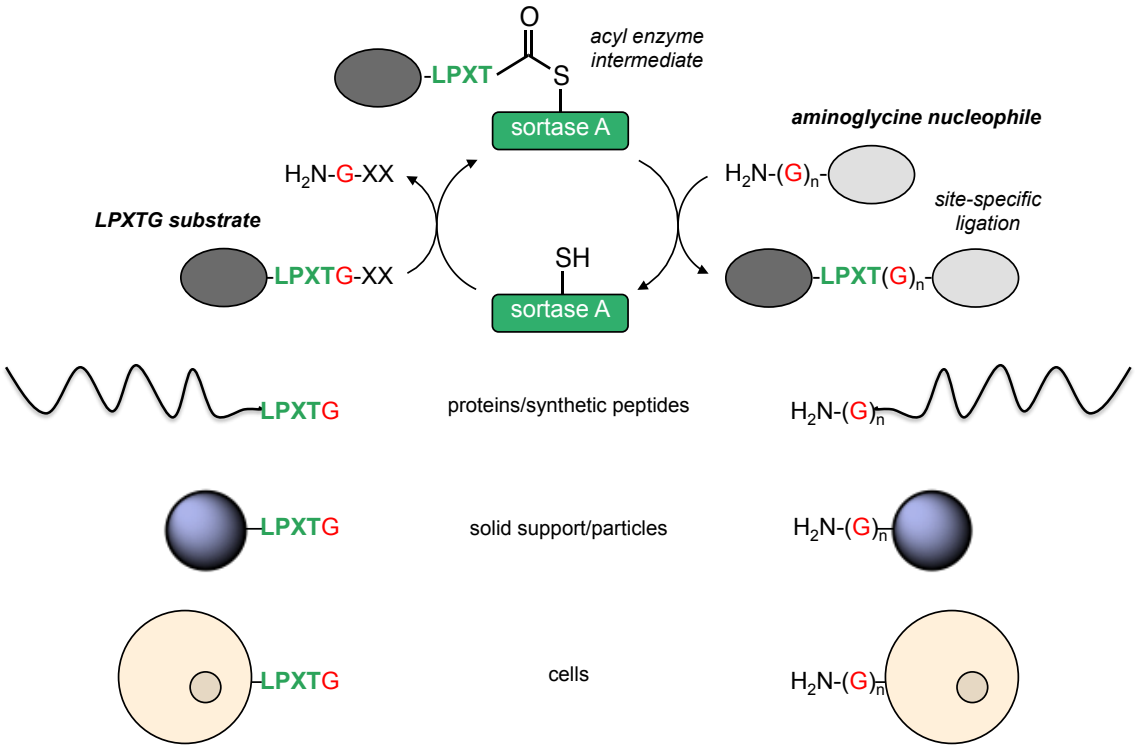


Figure 3. Protein modification using *in vitro* transpeptidation catalyzed by SrtA. This strategy can be applied to a variety of ligation partners, shown in the bottom half of the figure, facilitating the construction of diverse protein fusions.

1.2 Expanded Substrate Tolerance of Sortase Homologs

In addition to increases in the diversity of protein engineering applications that employ sortases, there have also been a number of studies focused on circumventing critical limitations of the method itself. With respect to $SrtA_{staph}$, slow reaction rates, reversibility, strict substrate specificity,[63] and limited acceptance of structurally diverse amine nucleophiles[63-65] reduce the usefulness of the enzyme for certain applications. Slow reaction rates can often be overcome by altering reaction conditions and reversibility can be diminished by

the use of specially designed substrates.[66, 67] Our lab has previously developed substrates containing the Ni^{2+} binding motif GGHG on the C-terminus, which, upon cleavage, demonstrated a reduced capacity to function as a nucleophile in the reverse reaction to reform the substrate.[66] Depsipeptides have also been used effectively modify the N-terminus of proteins, as processing substrates with an ester linkeage produces C-terminal fragments with no available amine, eliminating the possibility the reverse reaction to regenerate the starting material.[67] In contrast to the issue of reversibility, substrate specificity and limits of accepted nucleophiles cannot be overcome by modifying the materials or conditions used in the reaction, as they are founded in enzyme structure and the mechanism of catalysis. In response, multiple groups have now generated evolved SrtA mutants with improved catalytic activity, modified substrate tolerance, and independence from certain reaction cofactors.[48-52] While useful, these strategies often rely on laborious directed evolution strategies in order to identify useful mutations. As a complement to these approaches, our group has been exploring the reactivity of *naturally occurring sortase homologs* as a way to further expand the scope and versatility of sortase-mediated protein engineering.

With respect to substrate specificity, bioinformatics studies suggest that the vast majority of SrtA homologs are specific for LPXTG motifs[68]. Using the CW-PRED2 algorithm, sortase substrates have been predicted from more than 177 bacterial genomes.[61] $SrtA_{staph}$ was predicted to have between 15 and 21 substrates across 13 strains, all of which demonstrated a high preference for

leucine in position 1, proline in position 2, threonine in position 4, and glycine in position 5. Position 6 was also analyzed and demonstrated no preference for this position in sortase substrates. Position 3, the variable position in the sorting motif, was frequently occupied by a charged residue; however, no strong preference for specific amino acids was determined for this position. Interestingly, substrate prediction demonstrated that in about 10% of substrates, the 4th position was occupied by alanine. *In vitro* analysis of these preferences with discrete peptide substrates shows there are some discrepancies between the actual and predicted sorting sequences of some sortase homologs. Specifically, it has been demonstrated that $SrtA_{staph}$ has a preference for glycine in the 6th position, outside of the canonical LPXTG sorting motif.[69] Further, studies by Kruger et al. have demonstrated that $SrtA_{staph}$ accepts several amino acids in the 4th position of the sorting motif, a trait not predicted by bioinformatics algorithms.[69] Building from these initial studies, our group has now shown that SrtA homologs can exhibit substantial differences in tolerance for amino acids in multiple positions of the sorting sequence. The most substantial deviations from the canonical LPXTG sorting motif were found in the 4th and 5th positions of the amino acid sorting sequence. An example for the 5th position tolerance is shown in **Table 2**. These data revealed that many SrtA homologs accept several amino acids in the 5th position of the sorting sequence. A particularly wide tolerance to substrates with non-canonical amino acids in the 5th position can be seen in *Streptococcus pneumoniae* SrtA ($SrtA_{pneu}$).

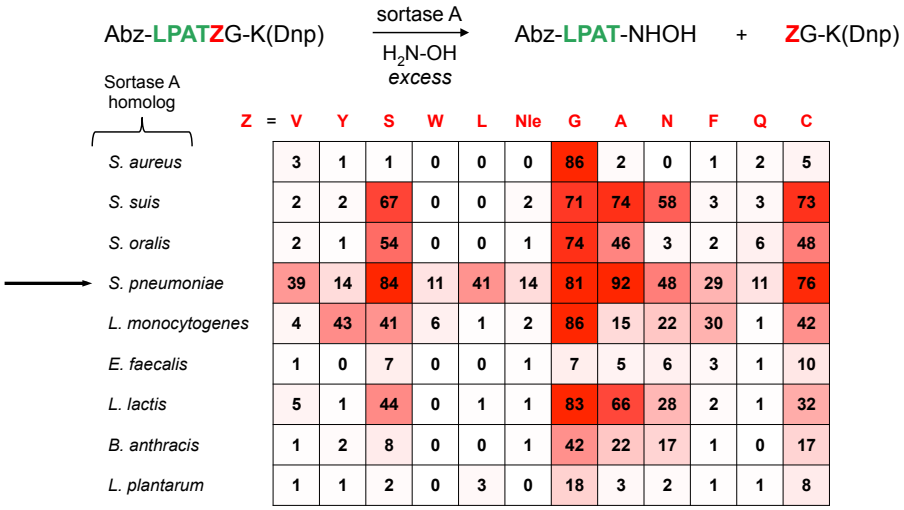


Figure 4. Amino acid preferences of selected SrtA homologs for the 5th position of the SrtA recognition sequence LPXTG [Nikghalb, Antos unpublished data].

In addition to the expanded substrate tolerance, the reaction completion percentages determined by HPLC and MS analysis of these model sortase-catalyzed ligation reactions demonstrate that $SrtA_{pneu}$ prefers the substrate LPATA *in vitro*, instead of the canonical LPXTG substrate. This is in clear contrast to $SrtA_{staph}$, which prefers only LPATG containing substrates. While some conversions are sub-optimal for use in protein engineering applications, the tolerance data shown in **Figure 4** above is the result of unoptimized reactions, and even low percent conversions may be improved by modifying the reaction conditions. Importantly, this would provide a route for expanding the scope of the sortase-mediated protein engineering, potentially increasing the range of endogenous protein targets that are compatible with sortase without the need to insertion of a recognition sequence in the primary structure, ultimately saving

time and resources.

Regarding the homolog $SrtA_{pneu}$, its particularly broad substrate tolerance presents an interesting issue to be examined with respect to the mechanism of substrate recognition. It is known that $SrtA_{staph}$ undergoes structural rearrangement upon correct substrate binding, which facilitates catalysis of transpeptidation.[70, 71] The wide substrate tolerance described above for $SrtA_{pneu}$ suggests that it may harbor a unique mode of substrate recognition. A thorough understanding this substrate recognition at the molecular level would benefit not only the continued development of $SrtA_{pneu}$ for use in protein engineering, but would also contribute to our fundamental understanding of sortase enzymology. Unfortunately, no structure of $SrtA_{pneu}$ has been published to date, which precludes a more in depth analysis of its substrate recognition and catalytic mechanism.

1.3 Substrate Binding and Structure of SrtA Homologs

Published solution NMR structures have been determined for $SrtA_{staph}$[71] and sortase A from *Bacillus anthracis* ($SrtA_{anth}$)[72] with a bound substrate analog, referred to as LPAT*, which utilizes a disulfide bond forming linker on the C-terminus that irreversibly binds to the enzyme active site.[71] These structures have been instrumental in developing an understanding of the reverse protonation mechanism of catalysis[73] and the enzyme-substrate interactions of the first four amino acids within the sorting sequence.

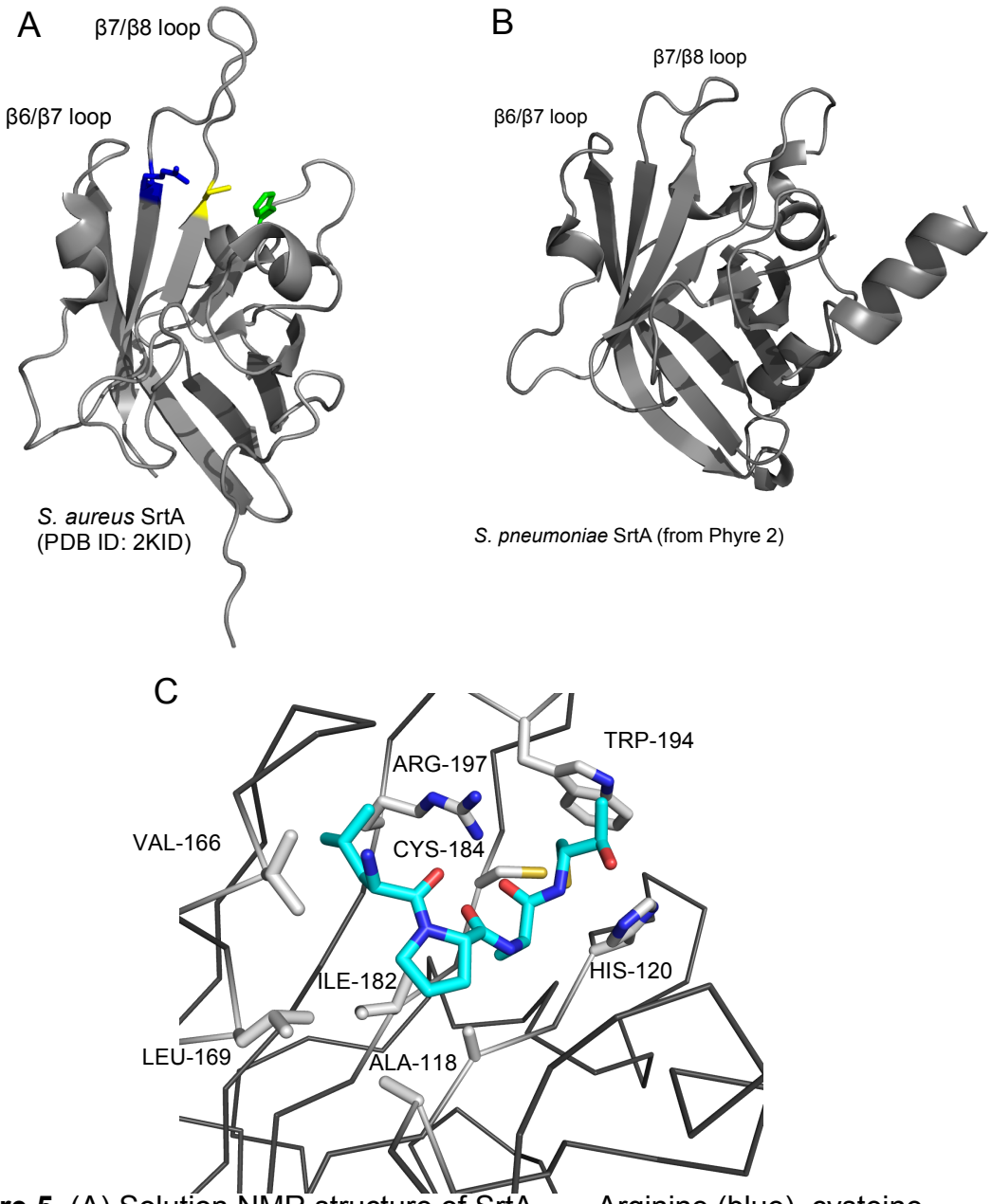


Figure 5. (A) Solution NMR structure of $SrtA_{staph}$. Arginine (blue), cysteine (yellow), and histidine (green) stick structures represent the key catalytic residues. (B) Predicted structure of $SrtA_{pneu}$ from the Phyre 2 structural prediction server[74]. (C) Active site architecture of the solution NMR structure of $SrtA_{staph}$ with the bound substrate analog LPAT* (PDB ID: 2KID). Side chains of residues surrounding the active are shown as stick structures, highlighting the number of hydrophobic interactions stabilizing the substrate-bound state.

In terms of overall structure, sortases include an 8-stranded beta barrel fold considered to be unique to this family of enzymes.[20] This β barrel is surrounded by combinations of α and 3_{10} helices and loops that vary in size and conformation between classes and homologs.[21, 23, 72, 75-77] In general, SrtA homologs contain three catalytic residues: cysteine, arginine, and histidine, within the active site.[78, 79] In $SrtA_{staph}$, the active site floor, or binding groove, is formed from strands β4 and β7 (**Figures 5** and **6**). The walls of this binding pocket are composed of loops joining other β strands and α helices in the structure. Within the active site, bound substrates are found in an "L" shape, facilitated by the proline residue, which angles the scissile bond towards the active site cysteine. The leucine sidechain makes hydrophobic contacts with residues in the β6/β7 loops and proline is buried in a hydrophobic groove formed by residues of the β4 and β7 strands. The alanine residue displays a hydrophobic interaction with the H1 helix, however, it is important to note that space exists for projection of larger sidechains away from the enzyme, partially explaining the low specificity for amino acids in this position. The threonine in the 4^{th} position of the amino acid sorting motif functions forces the sidechain of a tryptophan residue up and away from the active site, moving the catalytic cysteine residue towards the scissile peptide bond.[71] This is considered to be one of the most important interactions in substrate recognition, as substrates with glycine at this position are completely inactive[63].

Binding of the sorting motif in the active site induces further

conformational changes in some homologs of SrtA. In $SrtA_{staph}$, loop β6/β7 undergoes a large disordered to ordered transition which is key to substrate binding, as the recognition of the sorting sequence leucine residue relies on multiple interactions in this strand.[71, 80-82] The nucleophilic attack of the scissile peptide bond by the active site cysteine shifts the β7/β8 loop to reveal the backside of the active site to incoming nucleophiles.[71] This also facilitates recognition of the threonine residue, as mentioned above. The size of the β7/β8 loop is predicted to be important for recognition of the 5th position in the amino acid sorting sequence (Nikghalb and Antos, unpublished data). $SrtA_{staph}$ displays a particularly large loop extending from the β7/β8 strands (**Figure 5**), which may account for the high selectivity for glycine in the 5th position of the sorting sequence.

Apart from $SrtA_{staph}$, NMR data for $SrtA_{anth}$ bound to the LPAT* substrate analog[72] shows a number of similar features, including the cys-his-arg catalytic triad and the structure of the substrate binding groove. The binding pocket in this enzyme appears to be mostly preformed, however, in contrast to that of $SrtA_{staph.}$[72] Other notable differences between these structures exist, particularly the direction of the threonine residue, which points down towards the guanidino group of arginine. This supports the presence of an oxyanion hole to stabilize the tetrahedral transition state in the catalytic cycle, as initially hypothesized from the structure of $SrtB_{anth.}$[83] Additionally, the β7/β8 loop of this homolog demonstrates a similar structural transition upon substrate binding to that of $SrtA_{staph}$ β6/β7

loop, whereas the β6/β7 loop of is ordered in both the apo and substrate-bound forms of the enzyme. An interesting N-terminal appendage, similar to that seen in the SrtC subclass,[84] is thought to protect the active site from hydrolysis of the substrate, facilitating reactivity with the intended N-containing nucleophile[72].

While a structure of $SrtA_{pneu}$ bound to a substrate analog has not been reported, structure prediction tools reveal features that may play a role in the expanded substrate tolerance of this enzyme. One-to-one threading by the Phyre2[74] structural prediction server has allowed for prediction of the structure of $SrtA_{pneu}$ (**Figure 5**). $SrtA_{pneu}$ is predicted to have a smaller β7/β8 loop than that of $SrtA_{staph}$, which could account for its increased substrate promiscuity, as this region should be involved in recognition of amino acids in the 4^{th} and 5^{th} position of the sorting sequence.

Another variable aspect of SrtA structure and substrate recognition can be seen in metal ion sensitivity. Many sortases[20, 85-87] are Ca^{2+} dependent, where the bound ion improves the catalytic activity of the enzyme. $SrtA_{staph}$ binds calcium with a set of acidic residues in the β6/β7 loop, stabilizing the substrate bound conformation[70, 71, 80]. The presence of Ca^{2+} ion in reactions with this homolog significantly increases the reaction rate[20]. Again, in contrast to $SrtA_{staph}$, sortases have also been discovered which are independent of Ca^{2+} ion[88, 89]. A recently determined example of this is *Streptococcus suis* SrtA ($SrtA_{suis}$), where Ca^{2+} ion has no effect on reaction rate or total conversion. Our lab has previously determined that Ca^{2+} ion has little to no effect on the reaction rates of $SrtA_{pneu}$,

which indicates it may recognize and bind substrates in a mechanism fundamentally different than $SrtA_{staph}$.

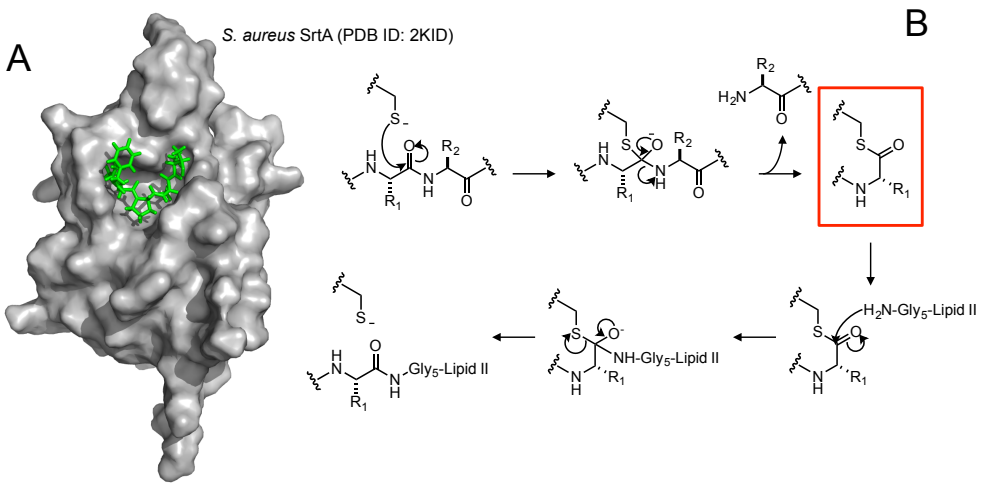


Figure 6. (A) Surface structure representation of $SrtA_{staph}$ with the bound substrate analog LPAT* (green). (B) Mechanism of transpeptidation within the active site of SrtA. The red box denotes the acyl-enzyme intermediate that closely resembles the structural state of the structure at left.

Unfortunately, these structures lack a complete description of substrate binding and recognition, as they do not illustrate interactions between the fifth position of the sorting sequence and the enzyme. The substrate analog used to collect the data from the studies by Suree et al.[71] and Chan et al.[72] closely mimics the acyl-enzyme intermediate, designated in **Figure 6** by the red box, which involves only the first four amino acids (LPAT) in the sorting motif. That said, additional work by Suree and coworkers has revealed some regions of $SrtA_{staph}$ implicated in recognition of the nucleophile and fifth amino acid positions[71] (**Figure 7**). ^{15}N-HSQC monitored titration of triglycine resulted in three regions displaying perturbed backbone resonances, presumably from the binding

of triglycine to these regions. These amino acids primarily lie around the sites where the 5th position amino acid should contact the binding pocket. Interestingly, sequence alignment of various SrtA homologs displays a marked difference in primary structure in these regions (**Figure 7**). This correlates with data from our lab **(Figure 4**) demonstrating the greatest differences from canonical LPXTG sorting motif recognition occur in the fifth amino acid, and highlights the need for a more rigorous assessment of contacts between the 5th position amino acid and SrtA during substrate recognition.

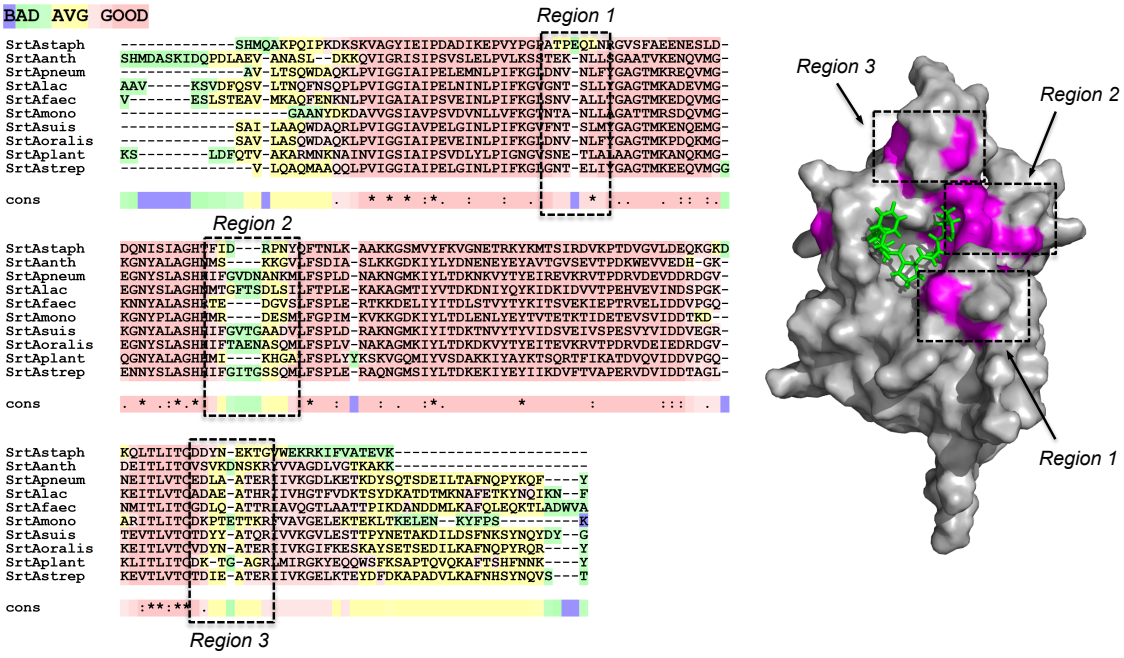


Figure 7. (A) Sequence alignment of selected SrtA homologs[90, 91]. The regions of greatest difference (boxed) correspond to the regions highlighted in the structure (B), indicating the least sequence homology lies on the structural features predicted to interact with the 5th position of the sorting sequence. (B) Surface representation of $SrtA_{staph}$ bound to substrate analog LPAT* (PDB ID: 2KID). Residues highlighted in magenta were determined to interact with the incoming nucleophile by analysis of peak perturbation during a ^{15}N-HSQC monitored titration of $SrtA_{staph}$ with G_3. These residues primarily lie around the region of the binding pocket predicted to interact with the C-terminus of the sorting signal.[71]

1.4 Enzyme Activity as a Function of Oligomeric State and Active Site Modifications

Previous studies have demonstrated the presence of both dimeric and monomeric forms of SrtA both *in vivo* and *in vitro*.[87, 92-94] *In vitro*, $SrtA_{staph}$ samples were subjected to several modes of analysis to confirm the presence of oligomers (**Figure 8**) and assess the strength of the dimer association. Specifically, analytical sedimentation equilibrium ultracentrifugation calculated an average K_d for the monomer-dimer equilibrium to be 54.6 ± 6.9 μM, considered a moderate interaction between the subunits of the complex.[92] *In vitro*, $SrtA_{staph}$ catalytic activity has been evaluated for both the monomeric and dimeric forms of the enzyme, where ligation reactions performed with monomeric preparations did not function as well as those with high concentrations of the homodimer.[92] The dimer was predicted to be as much as 8 times as active as the monomer based on steady state analysis of reactions at various concentrations (**Figure 9**).

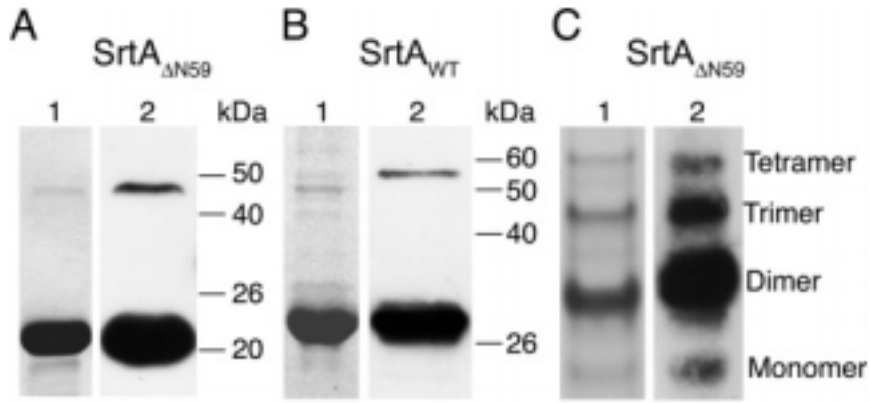


Figure 8. In each panel, lane 1 is Coomassie staining and Lane 2 is Western Blotting with anti-His6x antibody. (A) SDS-PAGE analysis of truncated $SrtA_{staph.}$ (B) SDS-PAGE analysis of wild-type $SrtA_{staph.}$ (C) Native PAGE analysis of truncated $SrtA_{staph}$. (Adapted with permission from Lu *et al*.[92] Copyright 2007 American Chemical Society)

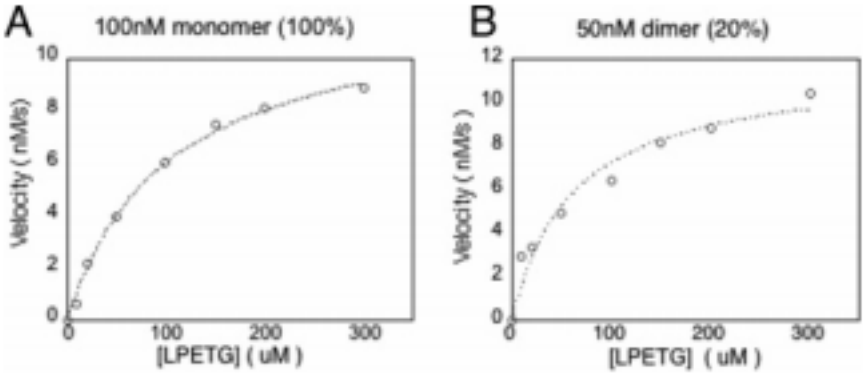


Figure 9. Steady state reaction velocity analysis with varying [LPETG] for (A) monomer and (B) dimeric $SrtA_{staph}$. (Adapted with permission from Lu *et a*l.[92] Copyright 2007 American Chemical Society)

In vivo, Zhu et al. demonstrated that higher levels of monomeric SrtA increase the amount of sortase-catalyzed surface attached protein by inserting a non-dimerizing mutant of SrtA into a knockout strain of *S. aureus.*[93, 94] This indicated greater catalytic activity for the monomeric form, in clear contrast to the *in vitro* studies. Further, the increased presence of surface attached proteins rendered bacteria with fully monomeric sortase more invasive than the wild-type strain. The dimerization observed in this homolog is anticipated to be biologically relevant as a mechanism for deactivating the enzyme when rapid attachment of proteins to the cell wall is not necessary.[92, 94] Dimerization has been observed as a mechanism of regulation in other enzymes.[95] Analysis of this monomer-dimer interplay has not been performed in other homologs of SrtA, however, based on the data above it is an essential piece of regulation which is not yet understood.

Regarding the modes of oligomerization available to $SrtA_{staph}$, it is not

certain how this enzyme associates *in vivo* or *in vitro*, as no structure has been published to demonstrate this phenomenon. The crystal structure of $SrtA_{staph}$[96] helped guide Zhu *et al.* in developing a set of mutants with diminished dimerization capacity. $SrtA_{staph}$ with any one of the mutations Y143A, K137A, or N137A demonstrated greater than 93% lower dimerization than the original truncated $SrtA_{staph}$. From these data, the authors hypothesized the Y143 and K137 mutations may disrupt a cation-π interaction which initiates dimerization. Cation-π interactions are seen in other proteins to facilitate dimerization and are often stabilized by additional charge interactions,[97-99] which is believed to be why mutation of N137 also diminishes the presence of the sortase dimer. These contacts would facilitate structural rearrangement to create further hydrophobic and hydrogen bonding, generating stable homodimers.[93]

Further information relevant to the work described here comes from a structure of $SrtA_{pneu}$ (PDB ID: 4O8L) deposited in the Protein Data Bank but not yet published in a written article. This x-ray crystallography structure shows a domain swapped dimer, where a large portion of the anticipated active site associates with the complimentary enzyme (**Figure 10**). More specifically, strands β7 and β8 and loops β6/β7 and β7/β8 are rotated out of the enzyme core structure to form contacts with the opposite monomer in the domain swapped dimer. Enzymes in this state would most likely be inactivated, as the catalytic cysteine residue lies on β7 loop, and is therefore buried in the domain swapped form. While it is not known if this structure is an artifact of enzyme preparation or

an authentic representation of the enzyme *in vitro*, it does demonstrate the ability of $SrtA_{pneu}$ to form stable, potentially inactivated, dimer constructs.

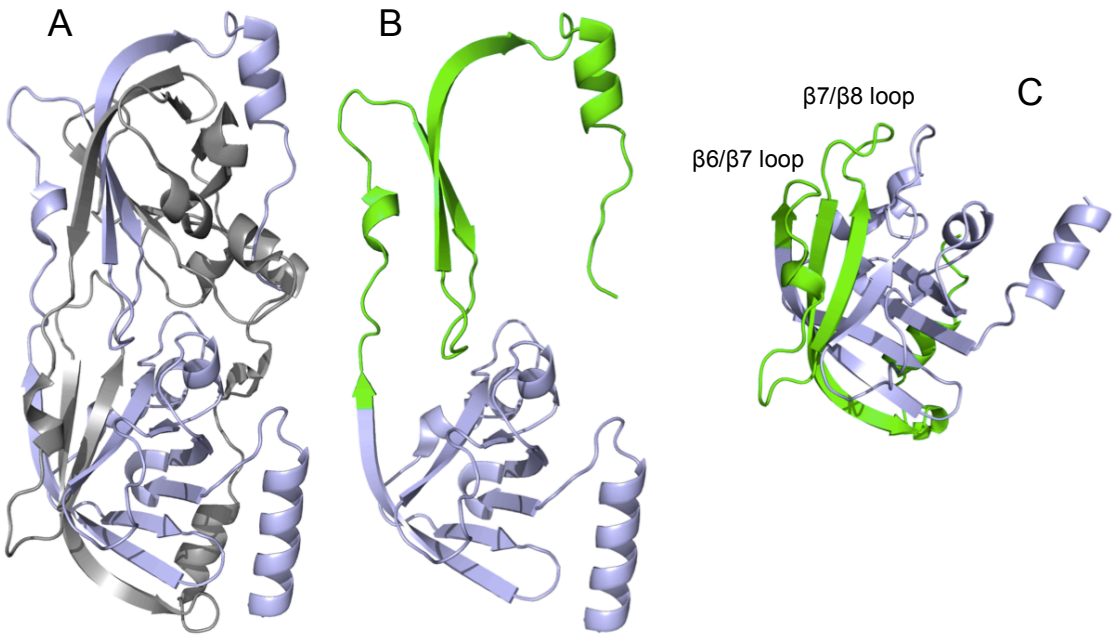


Figure 10. (A) The domain swapped dimer of $SrtA_{pneu}$ (PDB ID: 4O8L) (B) A domain swapped monomer from the structure of $SrtA_{pneu}$ in (A). (C) Predicted structure of $SrtA_{pneu}$ from the Phyre2[74] structural prediction server. In both structures, the green colored areas pertain to the domain swapped portion of the structure shown in (A).

An additional factor that may contribute to sortase activity comes from the crystal structure of $SrtA_{pyogenes}$, where a stable sulfenic acid modification was discovered in the active site of a crystal structure.[75] The sulfenic acid modification discovered was anticipated to be of a non-active form of the enzyme, as the active site cysteine would no longer be available to function as the primary nucleophile in catalysis. Interestingly, the modified cysteine stimulated conformational change in the enzyme similar to what is anticipated to occur with

a bound substrate, suggesting that structural rearrangement is an essential aspect to the mechanism of transpeptidation in $SrtA_{pyogenes}$.[75] The authors predicted that this could be biologically relevant as a potential method of inactivation by host immune systems releasing reactive oxygen species. McCafferty et al. determined that the reverse protonation mechanism of $SrtA_{staph}$ confers protection from oxidation onto the active site cysteine, as the high pKa of this residue side chain resists oxidation.[100] More recently, use of reactive Cu chelating peptides has demonstrated the potential for the active site cysteine of $SrtA_{staph}$ to be oxidized to sulfenic acid, which the researchers in this study suggested would function as a form of irreversible activation[101].

1.5 Project Goals and Overview

Given the importance of sortases as tools for protein engineering and as targets for therapeutic development, it is critical to establish a detailed understanding 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 the promiscuous enzyme $SrtA_{pneu}$ covalently docked with substrate analog. As described in this thesis, our approach involves the construction of ketomethylene-based sorting signal analogs, which we anticipate to be non-cleavable by the sortase enzyme (**Figure 11**). Importantly, the use of ketomethylene dipeptide isosteres will mimic amino acids in the 4^{th} and 5^{th} position of the sortase substrate motif, allowing for elucidation of key interactions between the enzyme and these residues. In addition, we believe this can provide

a more complete picture of substrate recognition in SrtA, and more specifically, the homolog from *S. pneumoniae,* as there is currently no published data regarding this enzyme's structure. Here we describe progress toward synthesizing the requisite ketomethylene building blocks, as well as preliminary evidence indicating that these substrate analogs are able to interact with the enzyme active site. We also describe an unanticipated oligomerization of recombinant $SrtA_{pneu}$, and the development of methods for refolding, purifying, reproducibly generating preparations of $SrtA_{pneu}$ with consistent activity. These studies provide an important foundation both for utilizing $SrtA_{pneu}$ as a tool for protein engineering, and for further structural characterization of this enzyme.

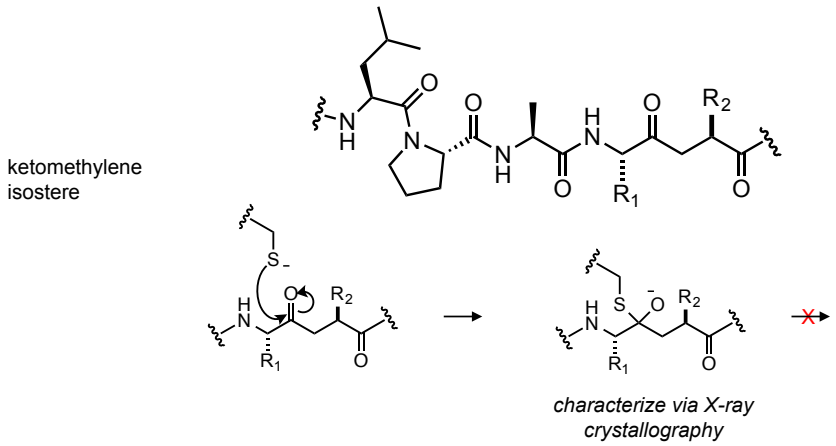


Figure 11. Anticipated mechanism of trapping for the proposed ketomethylene substrate analogs.

2. Preparation and Characterization of Recombinant *Streptococcus pneumoniae* sortase A ($SrtA_{pneu}$)

2.1 Expression, Refolding and Purification of $SrtA_{pneu}$

Initial work on this project was aimed at producing a stock of $SrtA_{pneu}$ to be used for structural characterization. To this end, an expression vector for $SrtA_{pneu}$ was obtained via commercial gene synthesis from DNA 2.0, which encoded a truncated version of $SrtA_{pneu}$ fused to an N-terminal His_6 tag. In this construct, the first 80 amino acids, corresponding to the hydrophobic transmembrane region, were removed to increase solubility in aqueous buffers. Initially, $SrtA_{pneu}$ was expressed from *E. coli* BL21(DE3) using standard molecular biology techniques. After cell lysis, the enzyme was isolated from the soluble fraction using immobilized metal affinity chromatography (IMAC). Prior work in the laboratory had succeeded in producing an active preparation of this enzyme, hereafter referred to as **batch 1**. An additional enzyme stock was generated for this work (**batch 2**) using a batch purification protocol for initial binding of enzyme to nickel nitriloacetic acid (Ni- NTA) resin. The eluted material from the first round of purification was re-purified through standard gravity-flow IMAC protocols to improve the purity of these samples. The sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis of **batch 2** (**Figure 12**) showed an intense band near 26 kDa, in reasonable agreement with the calculated molecular weight of 20.1 kDa for $SrtA_{pneu}$ This gel also demonstrated the good purity and high concentration of the collected samples. We also observed what

appeared to be $SrtA_{pneu}$ dimer at ~50 kDa that persisted in the sample despite the reducing and denaturing conditions of the SDS-PAGE loading buffer.

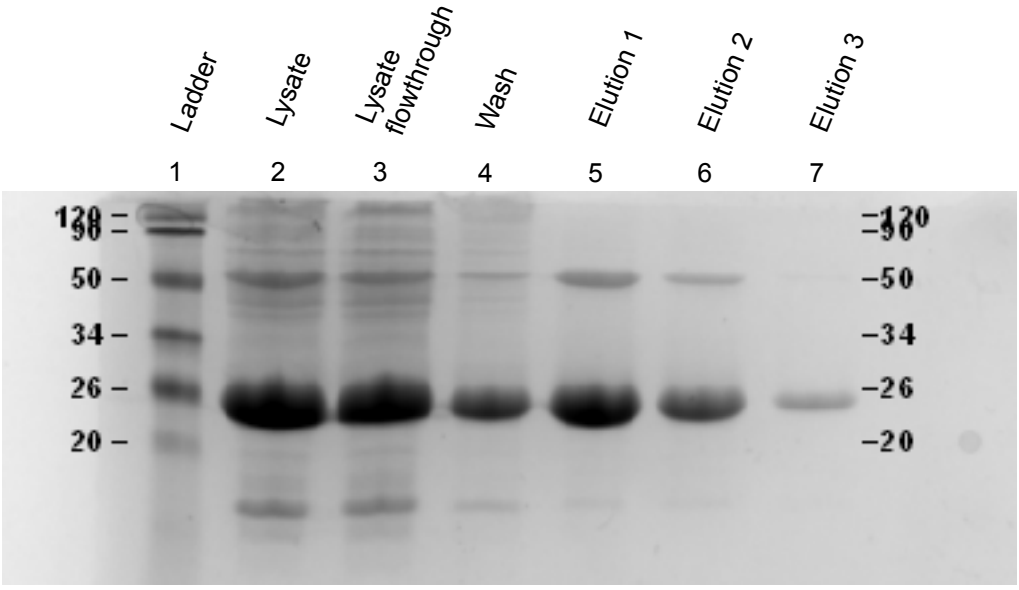


Figure 12. SDS-PAGE gel showing samples from the expression and purification of **batch 2** $SrtA_{pneu}$ prepared from IMAC purification using Ni-NTA resin.

Model sortase-mediated ligation (SML) reactions using a known peptide substrate (Abz-LPATA-GK(Dnp)) and the strong nucleophile H_2NOH were used to assess the *in vitro* reactivity of **batch 2** (**Figure 13**). After a 24 hour incubation at room temperature, the reactions were analyzed by reverse phase high performance liquid chromatography (RP-HPLC). Surprisingly, the **batch 2** enzyme stocks produced only minimal product formation (4.8% reaction conversion) when compared to the **batch 1** positive control (95% conversion).

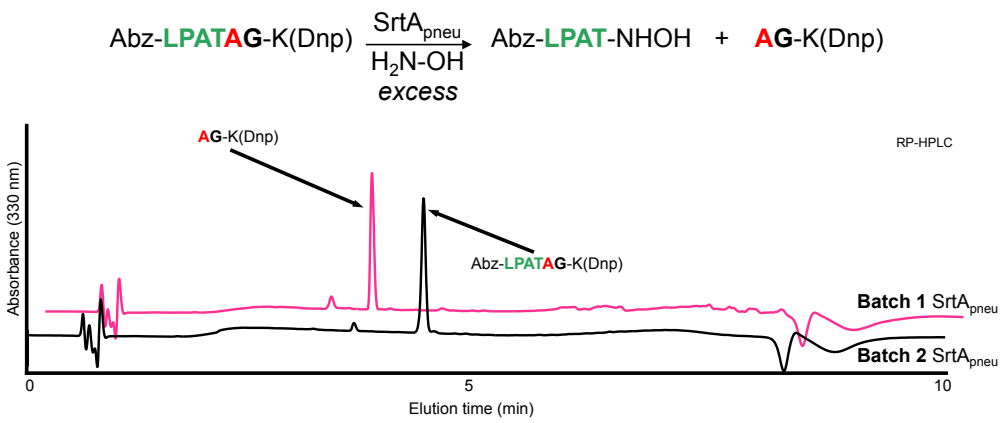


Figure 13. RP-HPLC analysis of model SML reactions demonstrating the difference in activity between **batches 1** and **2** of $SrtA_{pneu.}$ Reactions were analyzed after 24 hrs of incubation at room temperature. 40 μM Abz-LPATA-GK(Dnp) was used in this reaction.

As a result of the low observed reactivity for **batch 2,** a new batch of $SrtA_{pneu}$ (**batch 3**) was expressed and purified by standard IMAC protocols from newly transformed *E. coli* to eliminate the possibility of issues with the previous cell line (**Figure 14**). As expected, an intense band was seen around 26 kDa, matching closely to the expected molecular mass of $SrtA_{pneu}$ from SDS-PAGE analysis. We further analyzed these samples through mass spectrometry. Reconstruction of the enzyme molecular mass from the observed charge ladder (**Figure 15**) returned an uncharged mass of 20145 Da (expected MW = 20145 Da). Model SML reactions prepared from the **batch 3** $SrtA_{pneu}$ also demonstrated low reactivity (**Figure 16**), where only 21% conversion of the peptide substrate was observed.

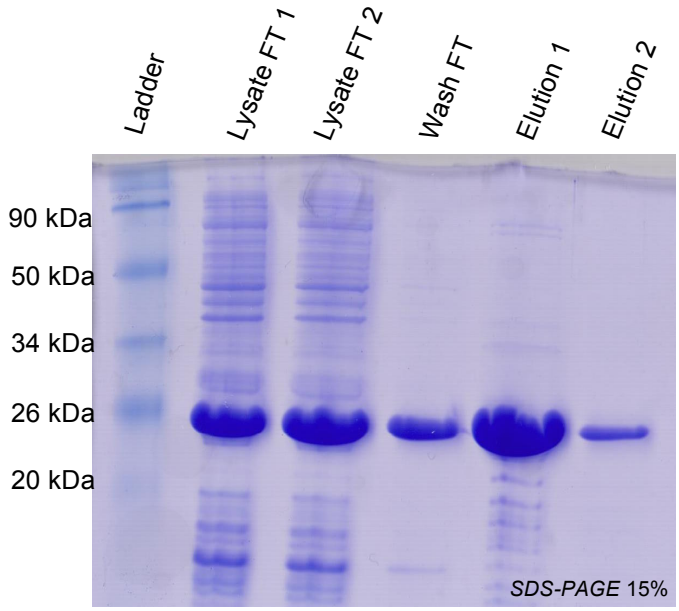


Figure 14. SDS-PAGE gel representative of $SrtA_{pneu}$ **batch 3** samples prepared from IMAC purification using Ni-NTA resin.

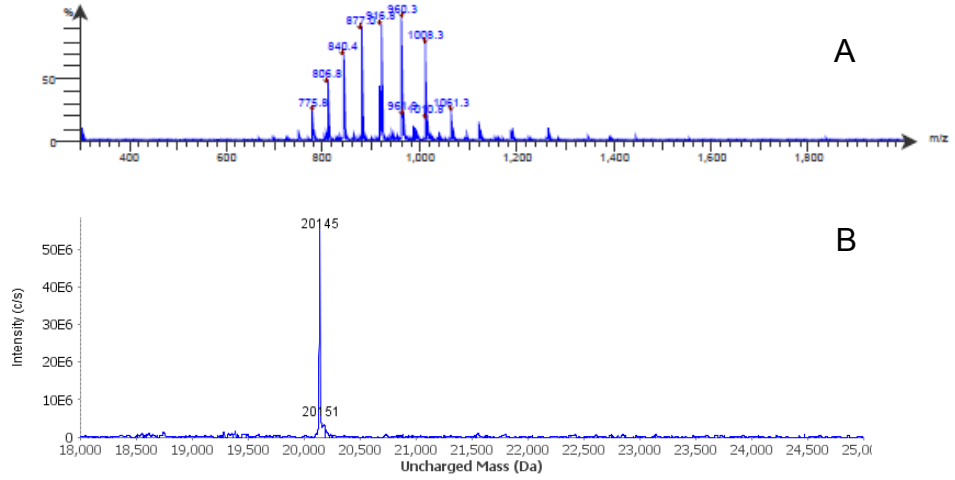


Figure 15. (A) ESI-MS spectrum of **batch 3** $SrtA_{pneu}$ (B) Deconvoluted mass spectrum of $SrtA_{pneu}$ **batch 3**. The primary mass shown here, 20145 Da, is identical to that predicted for this $SrtA_{pneu}$ truncation.

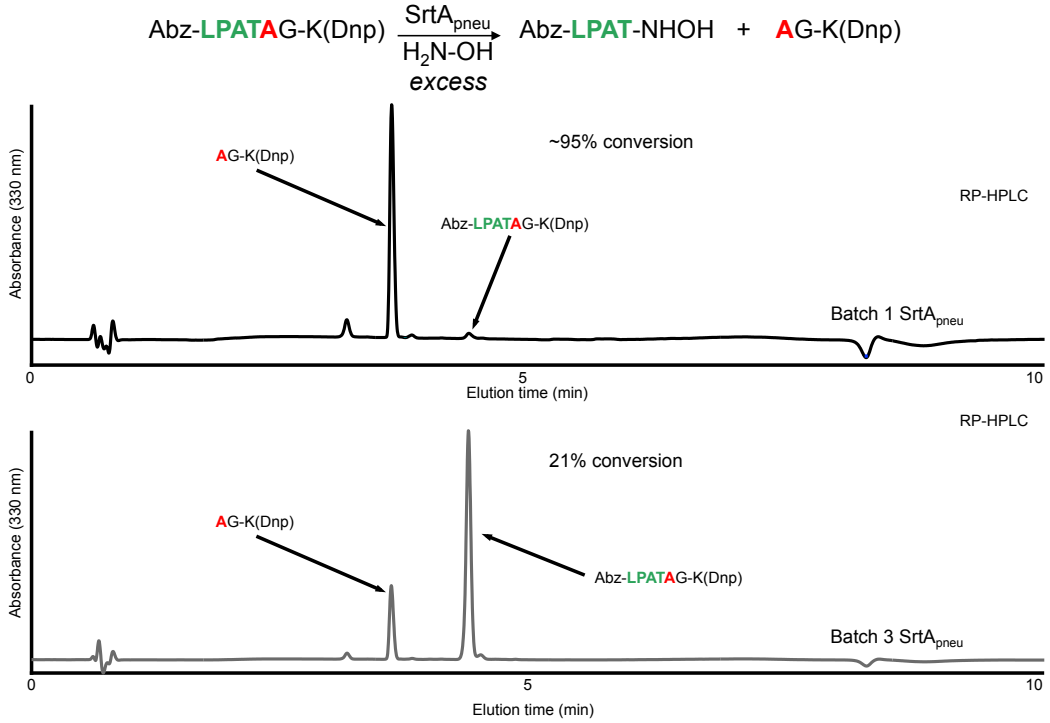


Figure 16. RP-HPLC analysis of model SML reactions demonstrating the difference in activity between **batches 1** and **3** of $SrtA_{pneu}$. Reactions were analyzed after 24 hrs of incubation at room temperature.

To further investigate the observed discrepancy in enzyme reactivity, we analyzed our catalog of $SrtA_{pneu}$ samples by native PAGE (**Figure 17**). In contrast to reducing SDS-PAGE, which suggested that all enzyme stocks were monomeric and identical, native PAGE revealed some striking differences. In particular, native PAGE revealed a number of higher molecular weight bands in all samples of $SrtA_{pneu}$. Based on the absence of these prominent bands under denaturing conditions, we hypothesized that they represented oligomers of the enzyme, a phenomenon known to occur in other sortase homologs.[87, 92, 94]

Interestingly, **batch 1** $SrtA_{pneu}$, which exhibited good *in vitro* reactivity and was used as a positive control in model SML reactions, showed a stark difference in the oligomer/monomer ratio when compared to newly prepared samples (**batch 2** and **batch 3**). Specifically, **batch 1** in lane 1 contained a lower concentration of oligomerized sortase bands, particularly in the area that we propose corresponds to trimeric and tetrameric $SrtA_{pneu}$, in addition to a higher concentration of monomer than **batch 2** and **3** samples in lanes 2 and 3, respectively. While both **batch 2** and **batch 3** samples demonstrated oligomerization in this analysis, we elected to continue our analysis with the **batch 3** $SrtA_{pneu}$ samples, as these had been developed using identical IMAC purification procedures used in the lab.

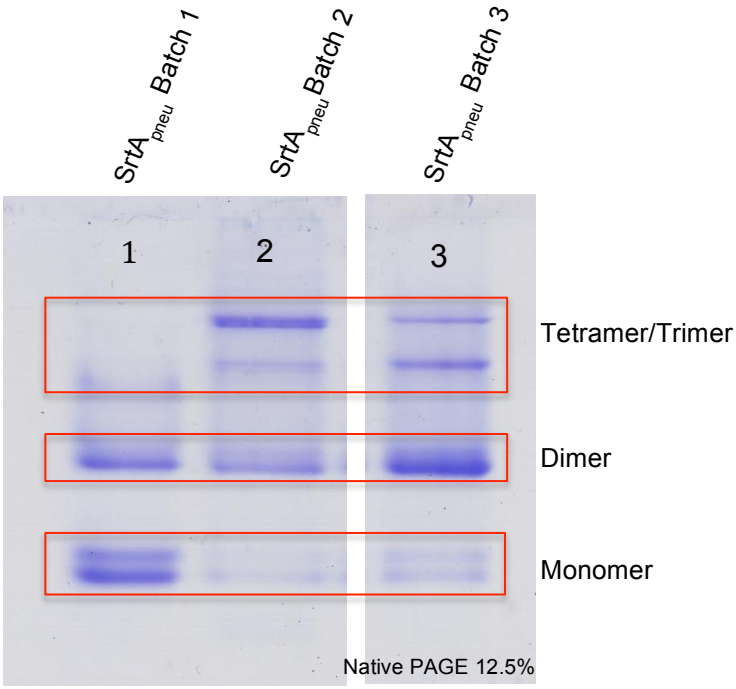


Figure 17. Native PAGE gel demonstrating the presence of potential oligomeric states in stocks of $SrtA_{pneu.}$ Samples were loaded in equimolar amounts. The predicted subunit count is given at right for each boxed region. Of particular interest here are lanes 1 and 3, which demonstrate variation in the band patterns between **batch 1** and **batch 3** $SrtA_{pneu}$. Specifically, **batch 1** shows a smaller amount of higher molecular mass bands than **batch 3**, and a greater concentration of protein anticipated to be the $SrtA_{pneu}$ monomer. Lane 2 came from the **batch 2** $SrtA_{pneu}$ expression and also displays the oligomerized band pattern, which differs from that of **batch 1**.

The presence of $SrtA_{pneu}$ in solution was further confirmed by size exclusion fast protein liquid chromatography (SE-FPLC), which was able to resolve samples of $SrtA_{pneu}$ into its assembled components (**Figure 18**). Resolution of different molecular weight species was clearly seen in FPLC traces, and native PAGE analysis of fractions from SE-FPLC separation showed

the same band pattern as unprocessed samples of the enzyme (**Figure 19**). Based on these data, it was concluded that $SrtA_{pneu}$ exists as about 93% oligomer in **batch 3**, which we reasoned was interfering with the enzymatic activity of the new preparation. In addition, given that the most active prep of $SrtA_{pneu}$ (**batch 1**) appeared to have the highest monomer content (**Figures 16** and **17**), we hypothesized that the monomer was the active form of the enzyme.

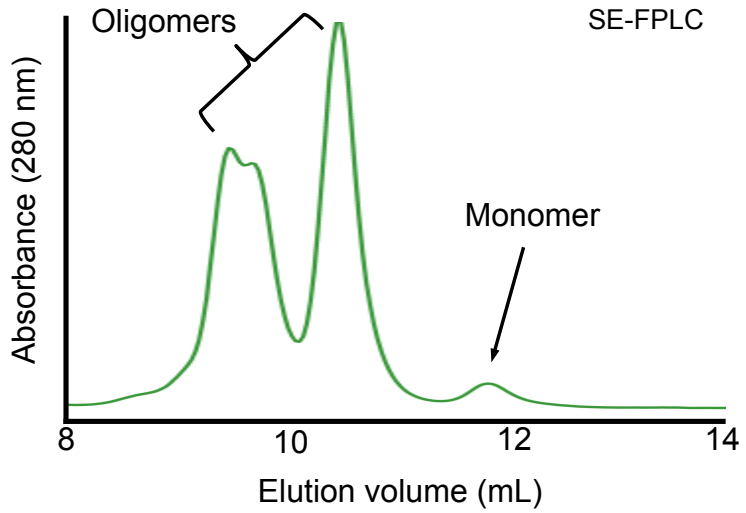


Figure 18. SE-FPLC trace demonstrating the presence of multiple oligomeric species in a pure sample of **batch 3** $SrtA_{pneu}$. This sample was separated at a 0.75 mL/min flowrate.

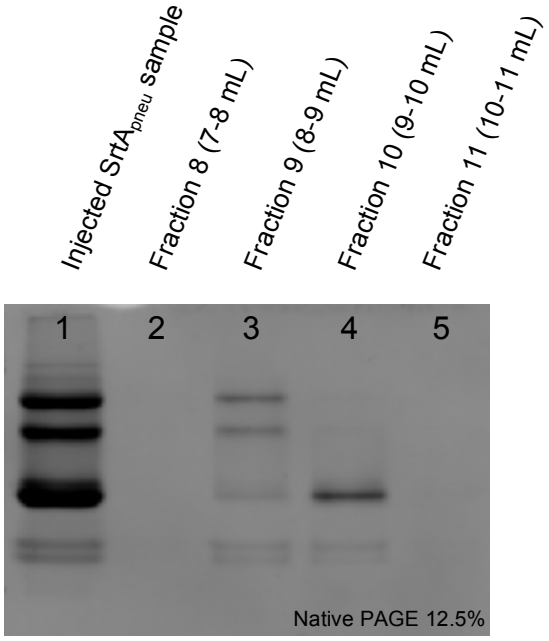


Figure 19. Native PAGE of the SE-FPLC fractions 8-11 from the separation in **Figure 18**. Similar band patterns can be seen in the fractions as in a sample of the injected **batch 3** SrtA$_{pneu}$ in lane 1.

To begin to explore this phenomenon and understand how to generate fully active, monomeric SrtA$_{pneu}$ preparations, samples were subjected to conditions that potentiated disassembly of the enzyme oligomers. First attempts at this involved incubation of serial diluted samples of the **batch 3** stock at room temperature to evaluate the possibility of high stock concentrations forcing the monomer-oligomer equilibrium towards the oligomerized state. Unfortunately, the presence of high molecular weight bands seen in native PAGE analysis of these diluted samples was unchanged (**Figure 20**). Notably, this observation did not align with previous accounts of sortase dimerization, as the dissociation constant

determined for $SrtA_{staph}$ dimer/monomer equilibrium (55 μM) suggested that the low affinity interactions between monomers would favor dissociation at low concentrations[92]. This indicated that $SrtA_{pneu}$ oligomerization could be fundamentally different from that found in $SrtA_{staph}$.

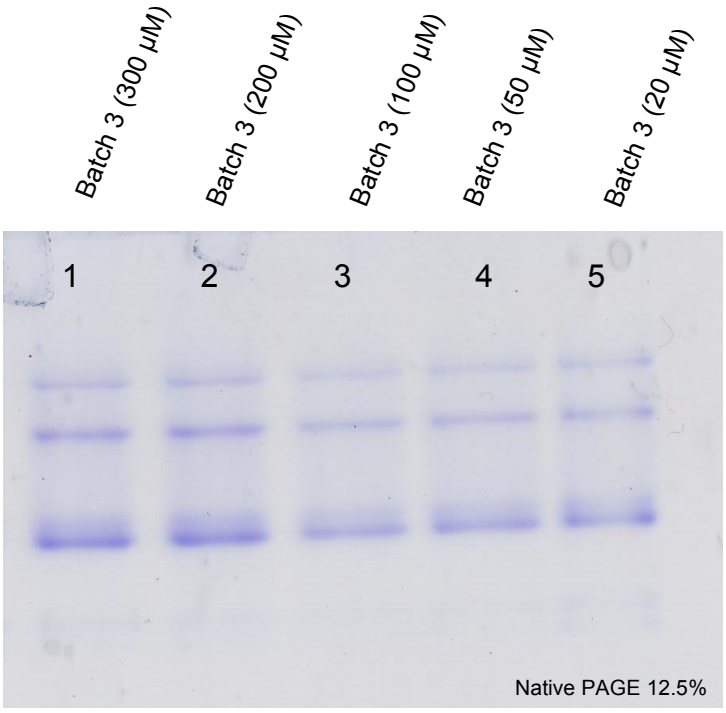


Figure 20. Native PAGE gel of room temperature incubated serial dilutions of **batch 3** $SrtA_{pneu}$.

Fortunately, further attempts at disassembling $SrtA_{pneu}$ were more successful. We next reasoned that oligomerization could be dependent on temperature. To evaluate this, samples of **batch 3** $SrtA_{pneu}$ were incubated at 37 ^{o}C and room temperature for 24 hours at 50 μM, 100 μM, and 300 μM concentrations. The relative disassembly of the putative oligomers was analyzed

via native PAGE. It was immediately apparent from native PAGE that incubation at 37 ^{o}C for 24 hrs degraded the oligomers into monomeric enzyme, generating a band pattern that was similar to the active **batch 1** preparation (**Figure 21**). In contrast, room temperature incubation yielded material that was identical to the original **batch 3** stock. Lower incubation concentrations also appeared to slightly improve the monomer content, although this did not appear to be as substantial. SE-FPLC analysis of undiluted samples of **batch 3** $SrtA_{pneu}$ heated for more than 24 hours also demonstrated an increase in peak area where the monomer was expected, mirroring the results of native PAGE experiments for shorter incubation times (**Figure 22)**. Use of "heat-disassembled" **batch 3** samples in a model SML reaction resulted in greater product formation. In particular, enzyme samples incubated at 37 °C for 1 week (**Figure 23)** produced conversion percentages of 95% after 48 hours, very near the **batch 1** positive control reactions, which exhibited 97% conversion in the same period of time (**Figure 24**). Reactions using **batch 3** $SrtA_{pneu}$ produced only 35% conversion in this reaction set. This indicated that the oligomerized states of the enzyme most likely exhibit some control over the rate and effectiveness of SML. While incubation of **batch 3** samples for extended periods of time (>24 hrs) resulted in somewhat greater dissociation of high molecular weight bands (**Figure 23**), aggregation was also observed in SE-FPLC traces of samples incubated for these extended periods of time (**Figure 22**).

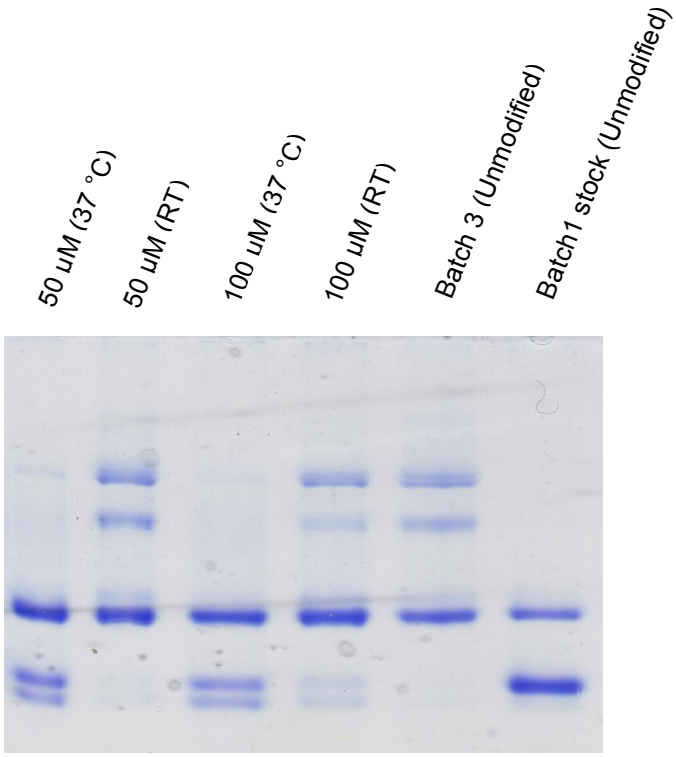


Figure 21. Native PAGE gel demonstrating the effect of heating samples of $SrtA_{pneu.}$ Lanes 1 and 3 contain diluted samples of **batch 3** enzyme heated for 24 hrs at 37 °C. The band patterns of the heated samples more closely resemble that of the batch 1 samples in lane 6 than the **batch 3** stock in lane 5.

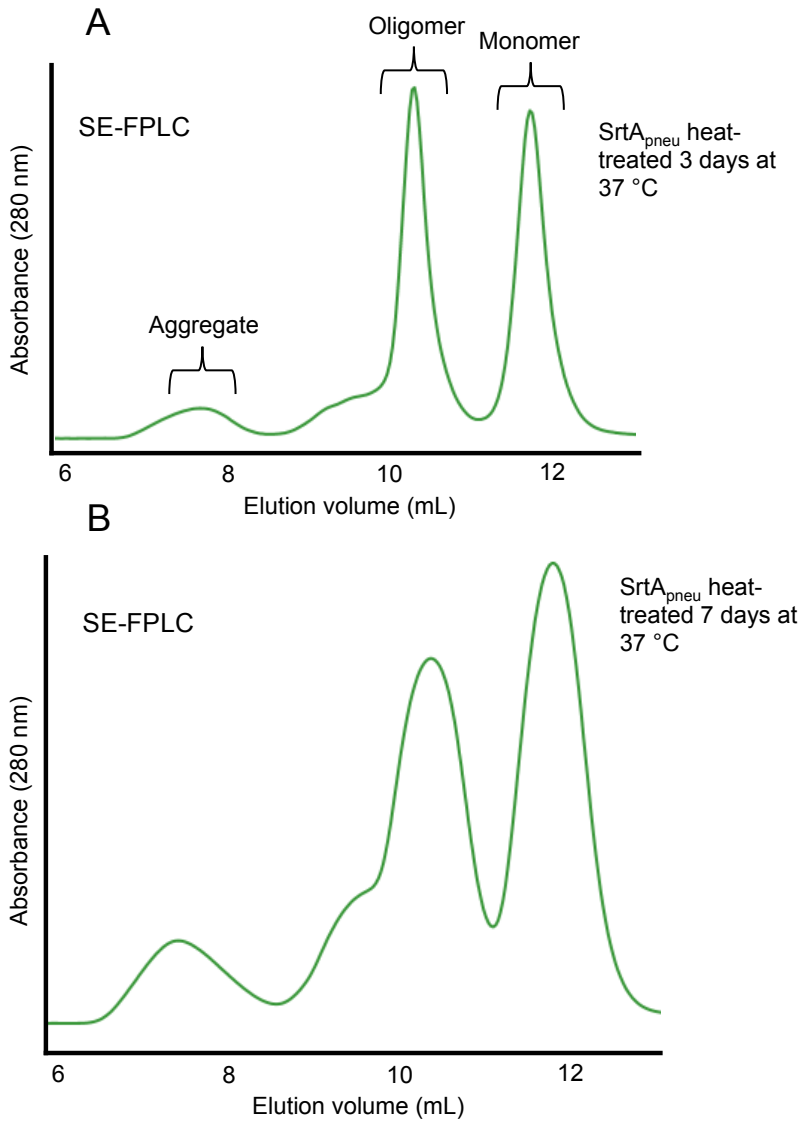


Figure 22. SE-FPLC traces showing the oligomer distribution in samples of $SrtA_{pneu}$ heated for (A) 3 days and (B) 7 days at 37 °C. A peak centered around 7.5 mL developed after longer periods of heating, anticipated to be aggregation of the enzyme.

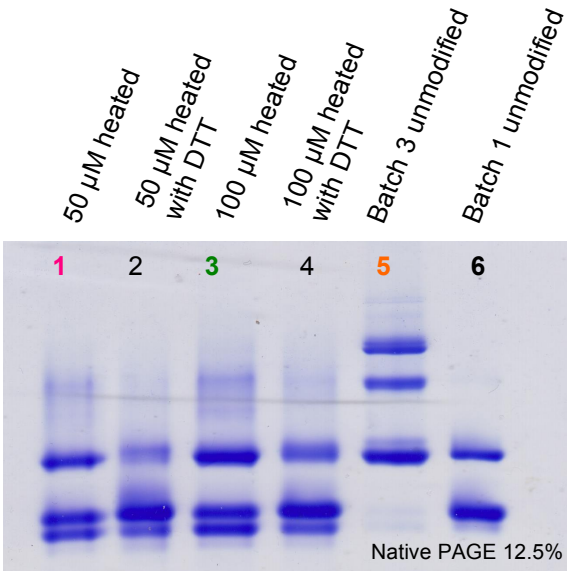


Figure 23. Native PAGE gel demonstrating the effects of 7 days of incubation at 37 °C on diluted samples of **batch 3** $SrtA_{pneu}$ with and without 10 mM dithiothreitol (DTT). As in **Figure 21**, the band pattern of the heat-treated **batch 3** samples looks more similar to the **batch 1** $SrtA_{pneu}$ than untreated **batch 3** $SrtA_{pneu}$.

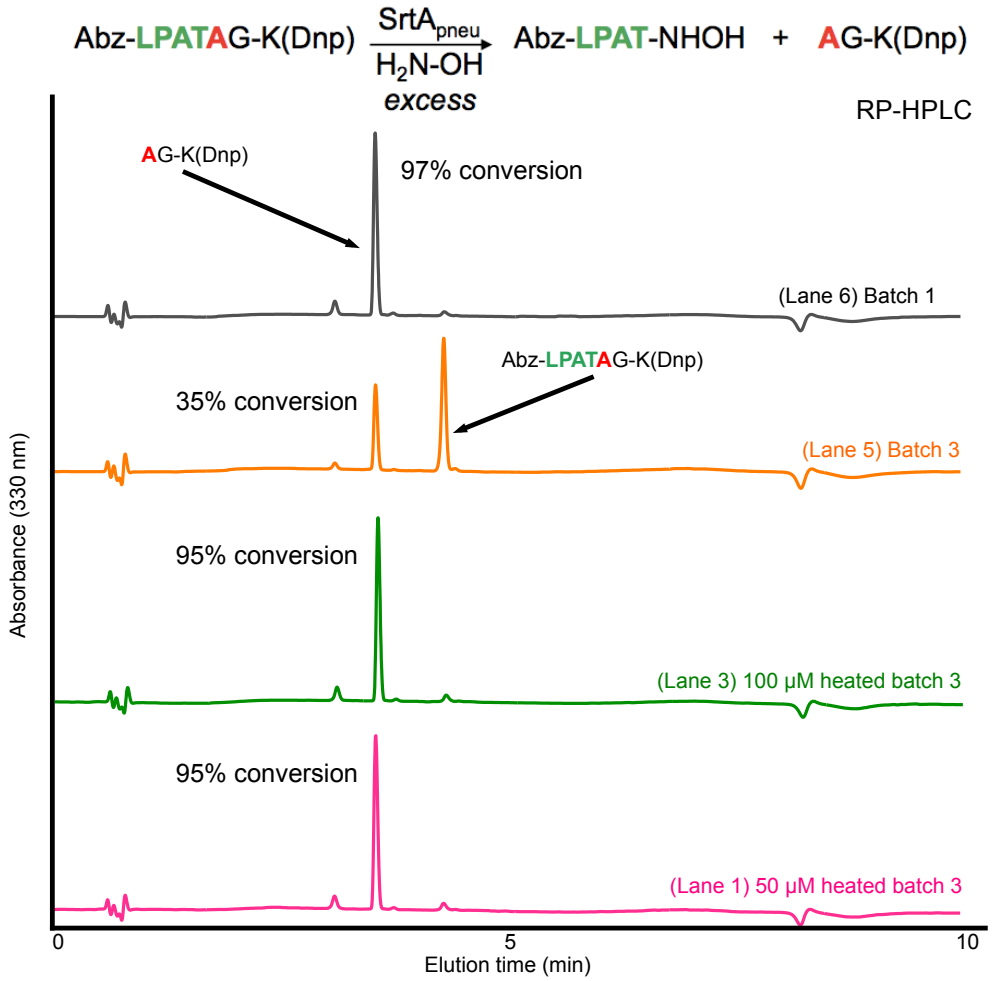


Figure 24. RP-HPLC analysis of reactions assembled with preparations of $SrtA_{pneu}$ from the gel in **Figure 23**. Reactions were incubated at room temperature for 48 hrs. The enzyme preparation used in the reaction is denoted by the lane number at the right of the trace. Heated samples of **batch 3** $SrtA_{pneu}$ produced similar conversion to the **batch 1** stock reactions, which matches the gel showing similar band patterns between the three samples.

Having succeeded in recovering some enzyme activity by heat treatment, we next evaluated the behavior of heat-disassembled $SrtA_{pneu}$ after prolonged incubation at room temperature, 4 ^{o}C, or -80 ^{o}C, (**Figure 25**). None of these treatments resulted in reversion of the native PAGE band pattern back to the

initial assembled state, and the desired monomeric $SrtA_{pneu}$ remained present. Overall, this indicated the $SrtA_{pneu}$ assembly process potentially involves an equilibrium that favors the monomeric form of the enzyme. As mentioned previously, however, dilution of the **batch 3** stocks without heating did not appear to diminish the presence of high molecular weight bands. Along with the information collected from the heating experiments, this suggested the presence of an activation barrier for converting between the monomeric and oligomeric forms of the enzyme, which can be overcome by heating the samples.

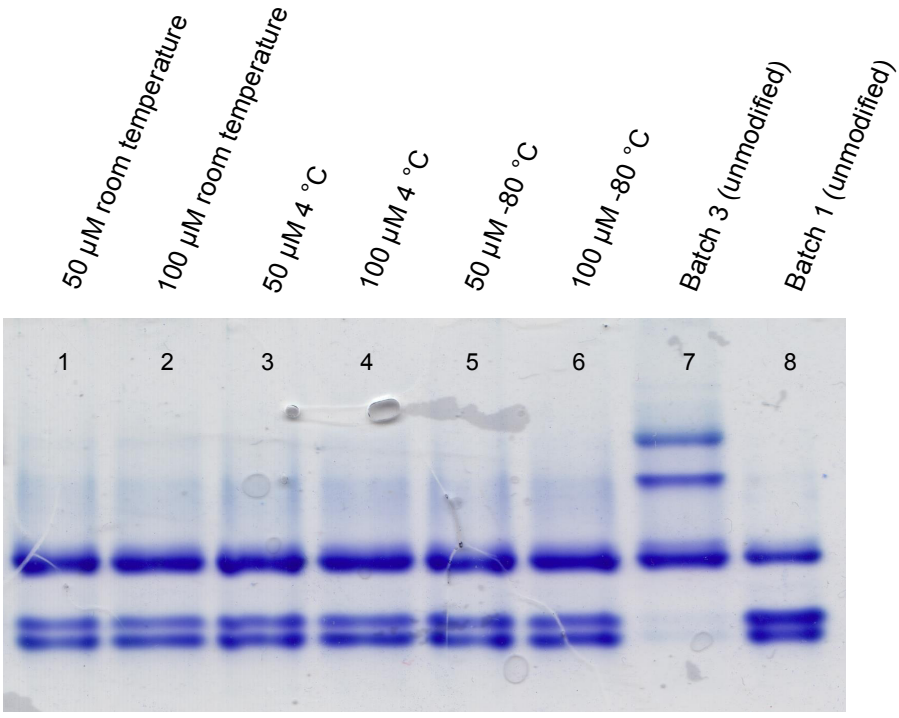


Figure 25. Native PAGE gel showing **batch 3** $SrtA_{pneu}$ heat-disassembled for 24 hours at variable concentrations and incubated for 24 hours at room temperature, 4 °C, and -80 °C. All samples were loaded in equimolar amounts.

To apply the information gained from the heating experiments to production of active $SrtA_{pneu}$ samples, alternative expression temperatures were tested to determine if production of the protein at lower temperatures would minimize the oligomer concentrations. After expressions attempts at 16, 25, and 37 ^{o}C, analysis by SE-FPLC determined there was no appreciable difference in the oligomer content following IMAC purification (**Figure 26**), as high relative oligomer concentrations could be seen in each injected sample.

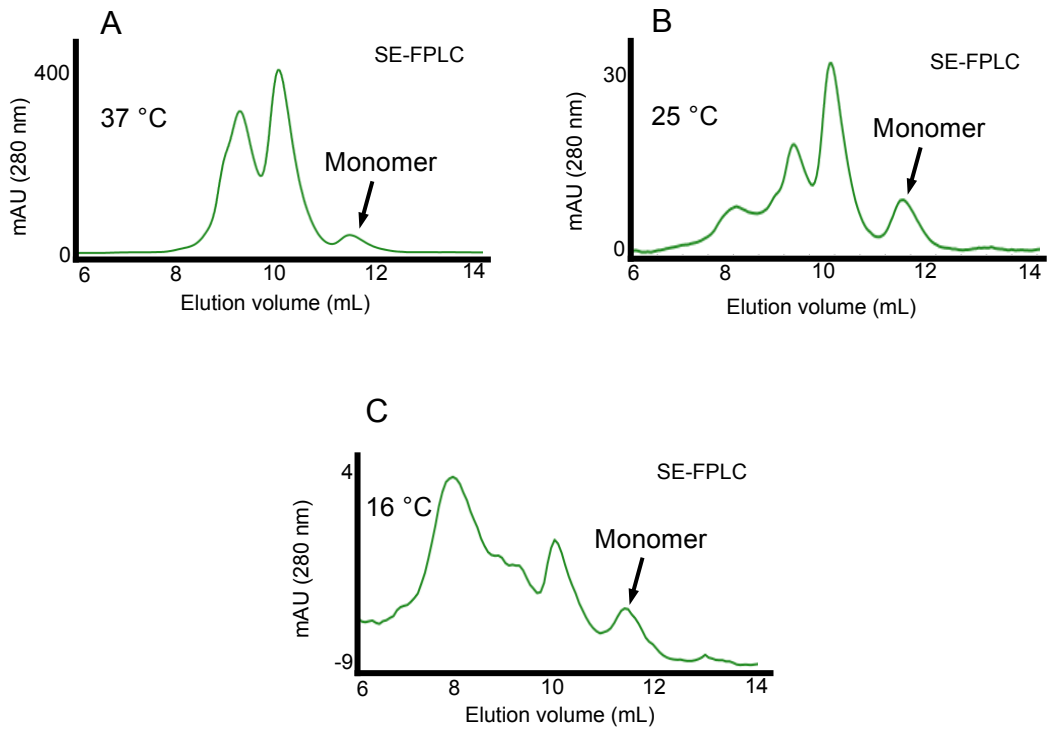


Figure 26. SE-FPLC traces demonstrating relative monomer concentration in samples of $SrtA_{pneu}$ expressed at variable temperatures and purified by standard IMAC protocols. (A) Expression at 37 °C, (B) expression at 25 °C, and (C) expression at 16 °C did not appear to alter the relative concentration of monomeric $SrtA_{pneu}$ by this analysis.

Based on the information gathered from the experiments described above, we next sought to explore a redesigned protocol for purifying $SrtA_{pneu}$ that would maximize monomer content. After considering the data on heating to increase monomer concentration, a denaturing purification step followed by refolding was anticipated to provide a similar role as heating post-purification. Denaturation during purification would presumably eliminate any intermolecular contacts or domain swapping interactions that could hold together multiple units of the enzyme. Therefore, the original purification protocol was modified to include an initial denaturing purification with 8 M urea, where the *E. coli* were first lysed in a denaturing buffer, the protein purified from clarified lysate in denaturing buffer via IMAC, followed by the denatured sample being refolded with a non-denaturing buffer. **Figure 27** shows the analysis of $SrtA_{pneu}$ refolding via a rapid dilution of denatured $SrtA_{pneu}$ into non-denaturing buffer (50 mM Tris pH 8.0, 150 mM NaCl) followed by an additional IMAC step to concentrate the refolded, soluble enzyme. Elution fractions from the rapid dilution refolding protocol were further purified by SE-FPLC to separate the sample into monomeric and oligomeric forms (**Figure 27).** Analysis of SE-FPLC fractions by native PAGE analysis demonstrated that $SrtA_{pneu}$ fractions with high monomer content could be isolated, however, a high relative concentration of enzyme was still trapped in the inactivated forms. Unfortunately, when refolded, monomeric $SrtA_{pneu}$ tested in a model SML reaction continued to show diminished reactivity when compared to positive control reactions employing **batch 1** $SrtA_{pneu}$ (**Figure 28**). This was surprising, as we had

anticipated the pure monomer samples to perform as well as or better than **batch 1** $SrtA_{pneu}$.

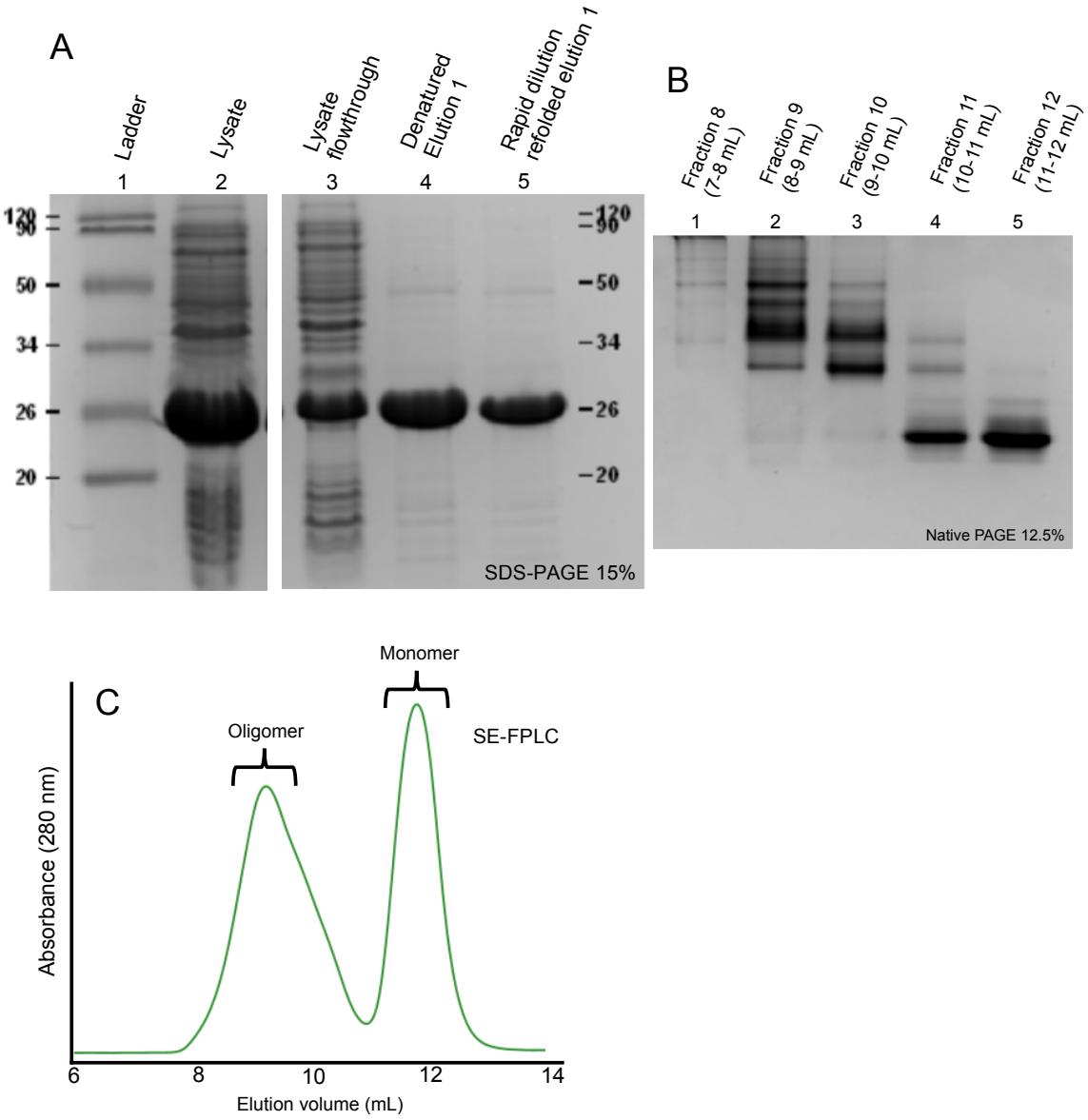


Figure 27. (A) SDS-PAGE gel showing the purity of $SrtA_{pneu}$ samples obtained from Ni-NTA purification after denaturing and refolding steps. (B) Native PAGE gel of SE-FPLC fractions 8-12 from the separation of $SrtA_{pneu}$ purified by the rapid dilution protocol. A sample from lane 7 of the (A) was injected to produce the samples in this gel. (C) SE-FPLC trace of an injected sample from lane 7 in (A). The fractions shown in (B) come from this separation.

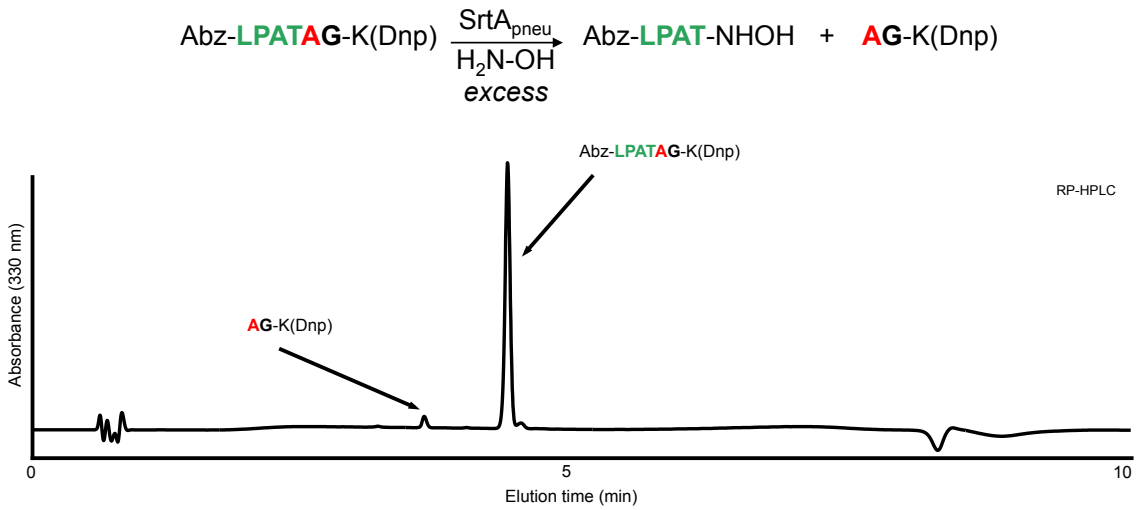


Figure 28. RP-HPLC analysis of a model SML reaction prepared with SE-FPLC purified monomeric $SrtA_{pneu}$ from the denaturing purification protocol. This reaction was incubated for 24 hrs at room temperature before analysis. Only minimal processing of the substrate was observed, corresponding to 3.5% conversion of the substrate.

To continue improvement of the $SrtA_{pneu}$ preparations, we next explored the use of reducing agents to promote oligomer disassembly and to ensure that the active site cysteine residue remained in its fully reduced form. Preliminary experiments with dithiothreitol (DTT) had demonstrated some efficacy for restoring $SrtA_{pneu}$ activity *in vitro* after a 22 hour incubation period at 37 ^{o}C (**Figure 29**), where $SrtA_{pneu}$ incubated with 10 mM DTT produced more product (72% conversion) when compared to both unheated (21% conversion) and heated (47% conversion) samples. To build on this initial result, DTT was combined with prolonged heating at 37 ^{o}C. This resulted in even greater dissociation of oligomerized $SrtA_{pneu}$, in addition to increased intensity of the top band of the monomer doublet observed in native PAGE for $SrtA_{pneu}$ (**Figure 30**).

While these native PAGE results were promising, the enzyme treated with these conditions failed to show any activity *in vitro* (**Figure 29**), despite the presence of significant monomer content in the enzyme sample. This prompted us to replace DTT with the non-sulfrous reducing agent tris(caboxyethyl)phosphine (TCEP). To our satisfaction, we observed that addition of 10 mM TCEP to SE-FPLC purified monomeric $SrtA_{pneu}$ from the denaturing purification protocol greatly improved the percent conversion observed in model SML reactions by RP-HPLC analysis (**Figure 31**). Specifically, in the absence of TCEP, reactions reached only 3.5% conversion, whereas in the presence of TCEP reaction conversion increased to 60%.

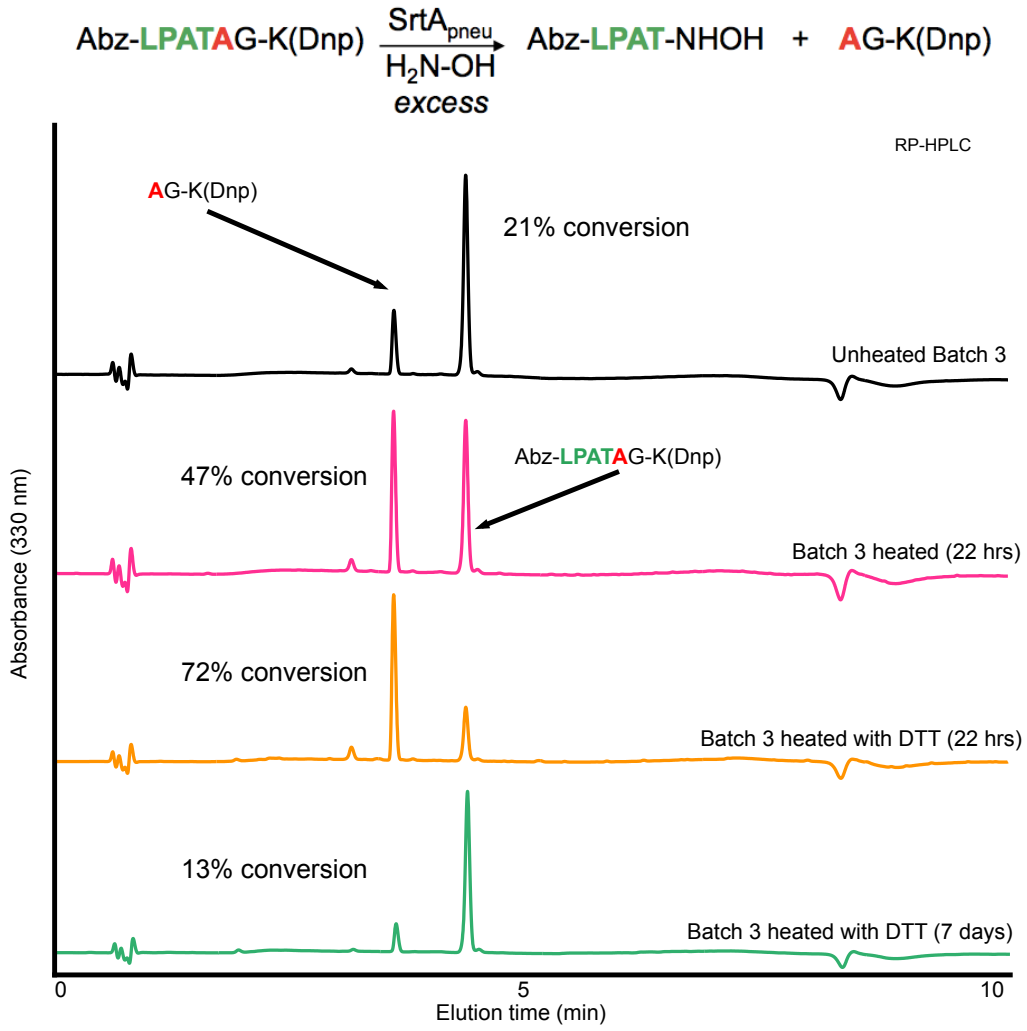


Figure 29. RP-HPLC traces demonstrating the changes in product conversion produced by including 10 mM DTT during heat treatment of $SrtA_{pneu}$ **batch 3**. Reactions were incubated at room temperature for 24 hrs before analysis. Heat-treatment of $SrtA_{pneu}$ for 22 hours at 37 °C increased substrate processing when 10 mM DTT was added compared to samples which were not heated or heated without 10 mM DTT. Unfortunately, incubation of $SrtA_{pneu}$ for 1 week with DTT at 37 °C resulted in samples with diminished activity when compared to shorter heating times.

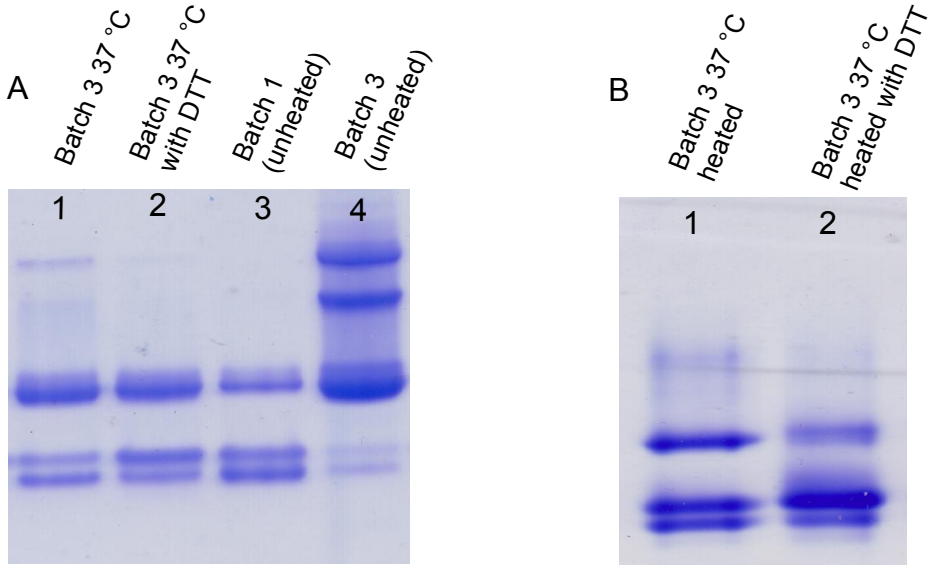
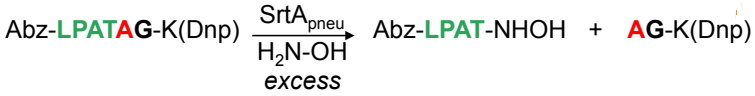


Figure 30. (A) Native PAGE gel demonstrating the effects of DTT on 22 hour incubations of **batch 3** $SrtA_{pneu}$ at 37 °C. (B) Native PAGE gel demonstrating the effects of DTT on 7 day incubations of **batch 3** $SrtA_{pneu}$ at 37 °C. Samples with DTT are denoted in the lane labels above. DTT appears to increase the intensity of the top band in the monomer doublet at the bottom of the gel, in addition to increasing overall monomer concentration.



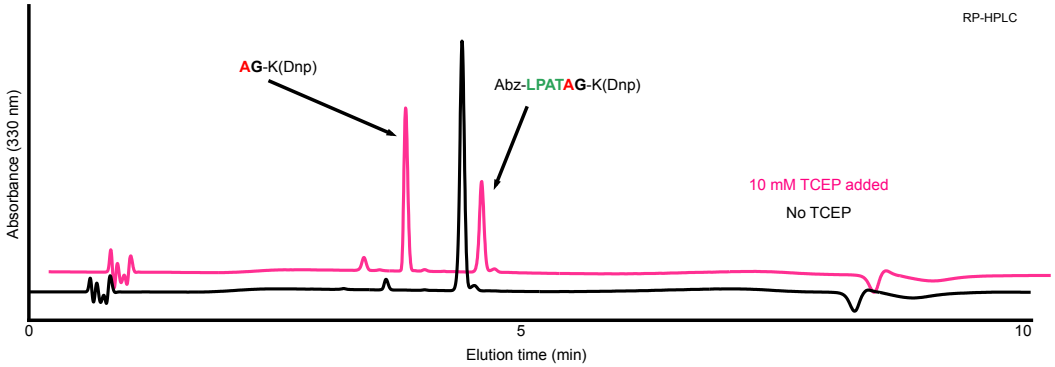


Figure 31. RP-HPLC traces displaying the increase in SML activity with 10 mM TCEP added to SE-FPLC purified monomeric $SrtA_{pneu}$ the denaturing purification after it was isolated. Reaction with TCEP was incubated 23 hrs at room temperature and the reaction without TCEP was incubated 24 hrs at room temperature before analysis.

In an attempt to combine the incremental improvements gained by enzyme refolding and the addition of TCEP, we ultimately settled on a revised $SrtA_{pneu}$ purification/refolding scheme in which TCEP was added to 1 mM in all purification and storage buffers. First, $SrtA_{pneu}$ was expressed at 37 ^{o}C, followed by IMAC purification under denaturing conditions in the presence if 1 mM TCEP. IMAC fractions were then rapidly diluted in non-denaturing buffer containing 1 mM TCEP, followed by an additional IMAC step to concentrate the enzyme. Enzyme samples purified under these conditions had higher relative monomeric concentration as compared to our original **batch 2/3** materials that were neither refolded nor treated with TCEP (**Figure 32**). TCEP addition also increased the intensity of the upper band in the monomer doublet observed in native PAGE gels of $SrtA_{pneu}$ samples, which we presume to be the active component of these enzyme preparations. Using SE-FPLC, enriched monomeric samples could be easily isolated (**Figures 32** and **33**). After isolation, we were pleased to observe that monomeric $SrtA_{pneu}$ performed equally as well as the **batch 1** positive control in model SML reactions (**Figure 34**), where 95% conversion was observed for the TCEP purified $SrtA_{pneu}$ samples.

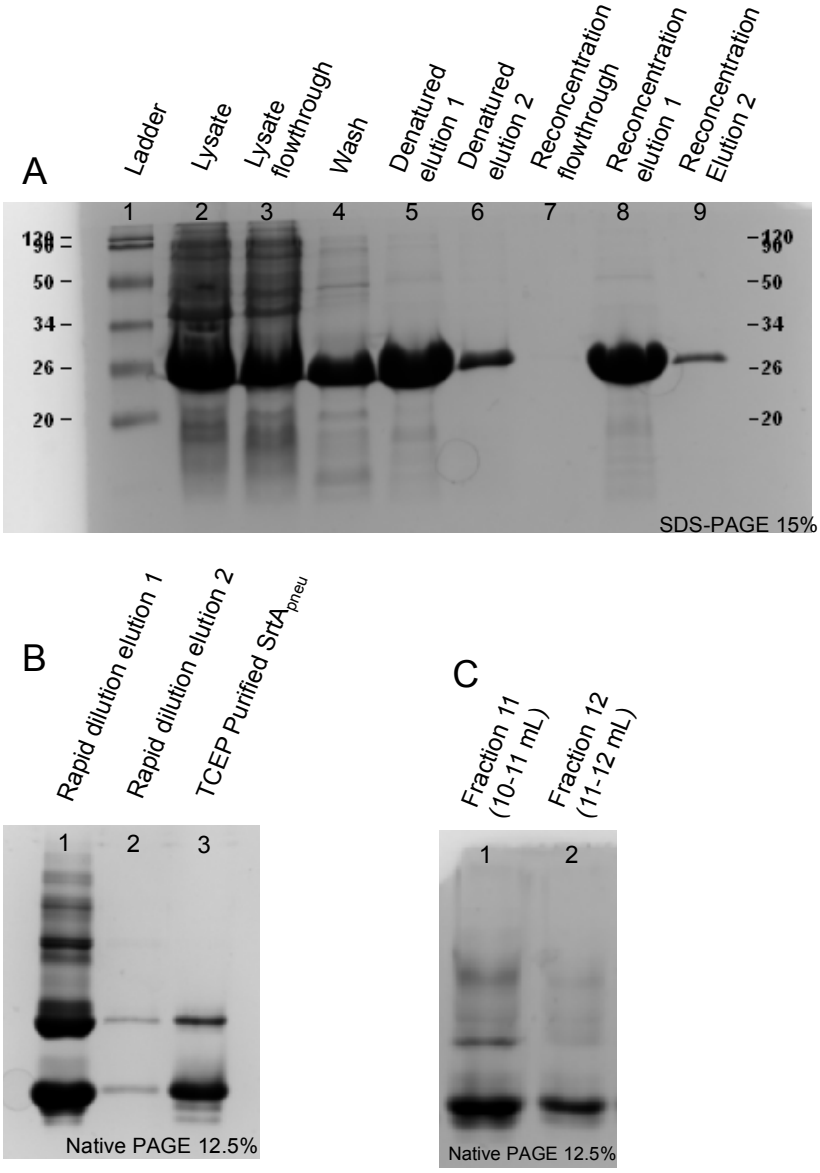


Figure 32. (A) SDS-PAGE gel showing analysis of the purification steps of $SrtA_{pneu}$ with TCEP as a buffer component. The sample in lane 5 of this gel was determined to be of good purity and concentration before rapid dilution and reconcentration to produce the sample in lane 8. (B) Native PAGE gel of reconcentrated $SrtA_{pneu}$ after refolding by rapid dilution. When compared to **batch 1** $SrtA_{pneu}$ in lane 3, the band pattern appears very similar, including the high intensity of the upper band in the monomer doublet. (C) Native PAGE gel of the $SrtA_{pneu}$ monomer fractions from SE-FPLC purification of a 1 mL injection of the sample in lane 1 of (B).

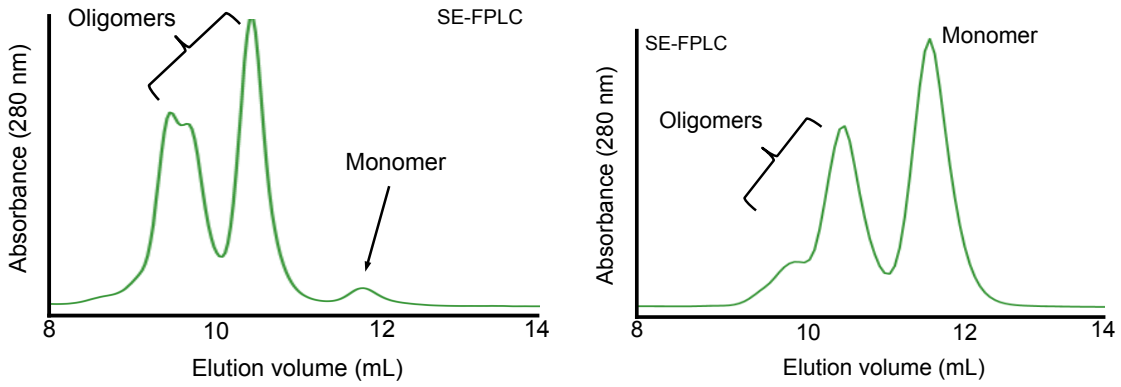


Figure 33. SE-FPLC traces from the purification of $SrtA_{pneu}$ by standard IMAC protocol separated at 0.75 mL/min (A) and with our adopted protocol using a denaturing purification step, refolding by rapid dilution, and reconcentration with 1 mM TCEP in all buffers separated at 0.5 mL/min (B) .

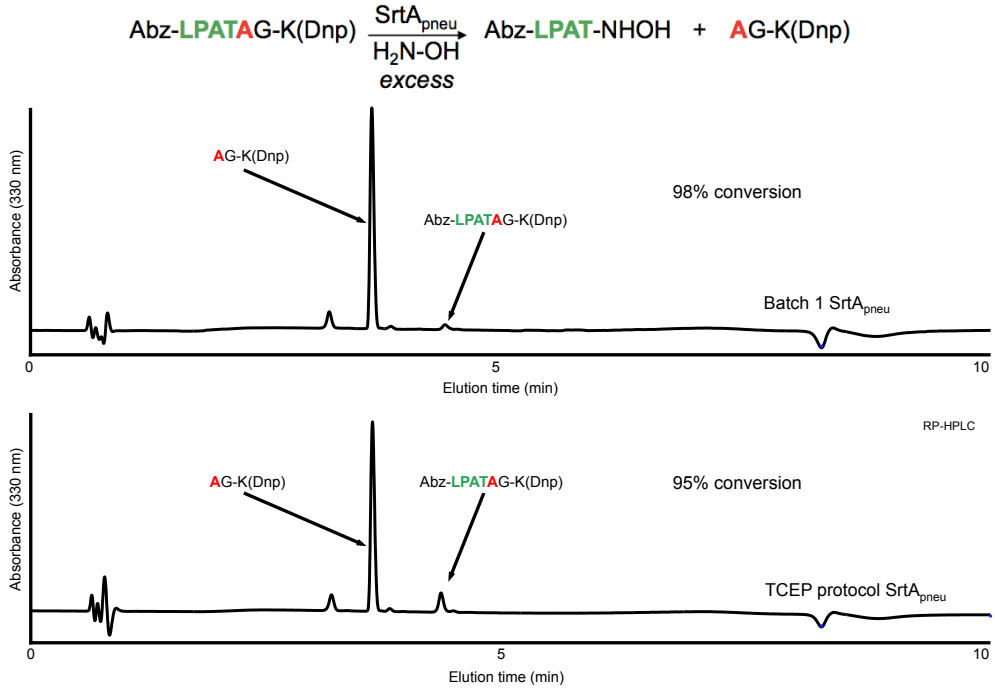


Figure 34. RP-HPLC traces demonstrating the improved reactivity of $SrtA_{pneu}$ purified with 1 mM TCEP in all buffers. The percent conversion after 24 hrs of incubation at room temperature is very similar to that of the reaction assembled with **batch 1** $SrtA_{pneu}$, indicating the activity of these samples was fully recovered by this purification process.

2.2 Analysis of Potential Modes of $SrtA_{pneu}$ Inactivation

Given that we were able to separate monomeric $SrtA_{pneu}$ from oligomeric $SrtA_{pneu}$, we next explored how the reactivity of oligomerized enzyme compared to the monomer. Using oligomeric SE-FPLC fractions in model SML reactions, we observed significantly less conversion of the peptide substrate when compared to monomeric fractions from the same purification (**Figure 35**). SDS-PAGE and native PAGE analysis of these reactions confirmed that total enzyme content was identical between reactions, and further showed that the assembly state of the enzyme was not altered by the presence of SML reaction components. Overall, these results demonstrate that monomeric $SrtA_{pneu}$ provided the best *in vitro* reactivity, while preparations enriched in enzyme oligomer were significantly less successful.

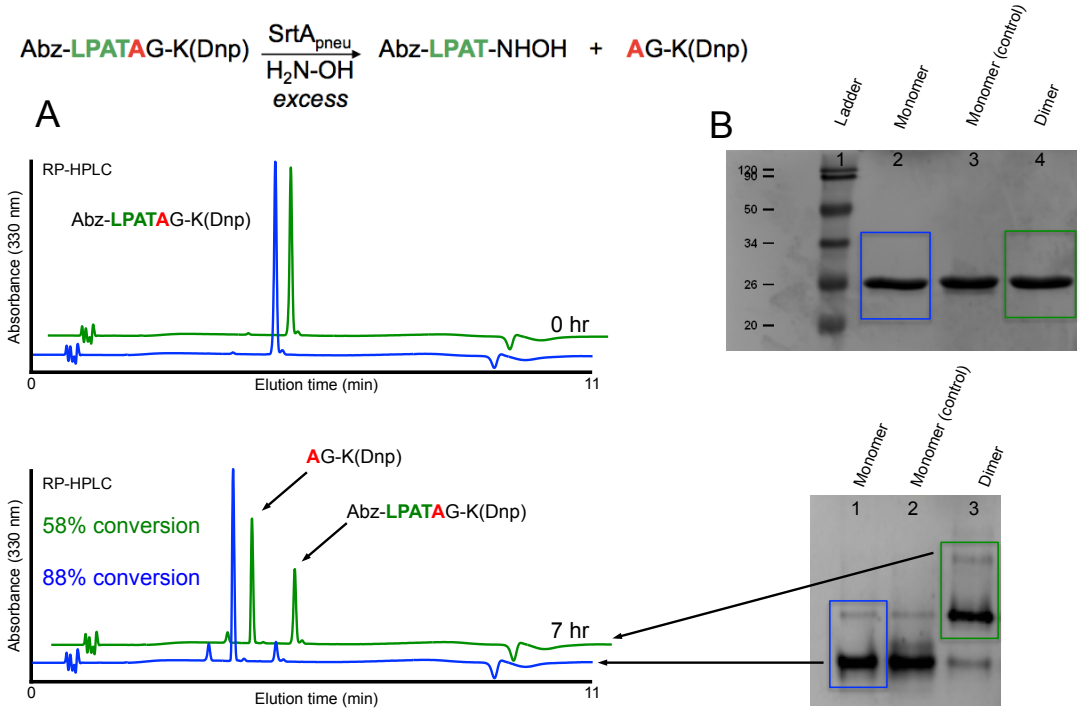


Figure 35. (A) RP-HPLC traces demonstrating the variability in reactivity between monomeric (blue) and oligomeric (green) preparations of TCEP purified $SrtA_{pneu}$. After 7 hours of incubation at room temperature, the reaction containing monomer reached 88% conversion, where as the oligomer reaction reached only 58% conversion. (B) Equimolar loading of samples from the $SrtA_{pneu}$ catalyzed peptide modification reactions in (A) separated through polyacrylamide gel electrophoresis. Gel (1) contains SDS and gel (2) does not. Identical samples were loaded in the lanes of each gel. Boxed lanes correspond to samples from reactions in HPLC traces of the same color at left.

With the knowledge that reducing conditions improved the activity of our $SrtA_{pneu}$ preparations, we decided to evaluate the possibility of the active site cysteine oxidation to sulfenic acid, which has been observed in $SrtA_{pyogenes}$ previously[75]. Commercially available, highly selective probes for sulfenic acid residues can be used to evaluate this type of oxidation in protein samples[102, 103] (**Figure 36**). For this purpose, **batch 3** $SrtA_{pneu}$ was buffer exchanged into

HEPES buffer for the labeling reactions. A 10x molar excess of the sulfenic acid probe NO_2-alkyne was added to these samples. After 24 hours of incubation at room temperature, positive control reactions assembled with 1.5 eq of H_2O_2 to **batch 3** $SrtA_{pneu}$ yielded an uncharged mass of 20427 Da upon deconvolution of the mass spectrum from the single peak in the sample (**Figure 37)**. This exactly matches the expected mass of 20427 Da for NO_2-alkyne combined with $SrtA_{pneu}$. The mass of $SrtA_{pneu}$ was also found in the sample. Deconvolution of the mass spectrum from the single peak in the experimental reaction did not display the expected uncharged mass for the NO_2-alkyne+$SrtA_{pneu}$ adduct for **batch 3** $SrtA_{pneu}$. This result suggests that **batch 3** samples, which function poorly in sortase mediated ligation reactions, do not suffer from sulfenic acid modifications to the active site cysteine.

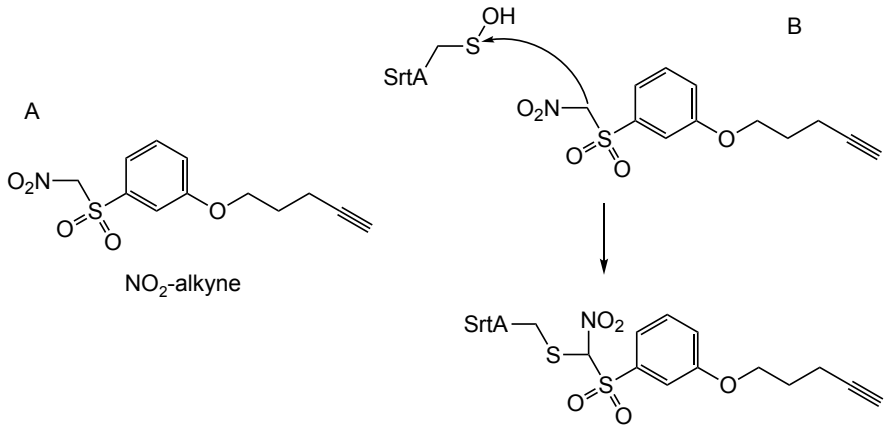


Figure 36. (A) Structure of sulfenic acid probe NO_2-alkyne used in the **batch 3** $SrtA_{pneu}$ labeling reactions. (B) Mechanism of addition to a potential sulfenic acid modification in the active site of $SrtA_{pneu}$ by the NO_2-alkyne probe. This forms a stable thioether bond, allowing the change in mass to be easily recognized by mass spectrometry.

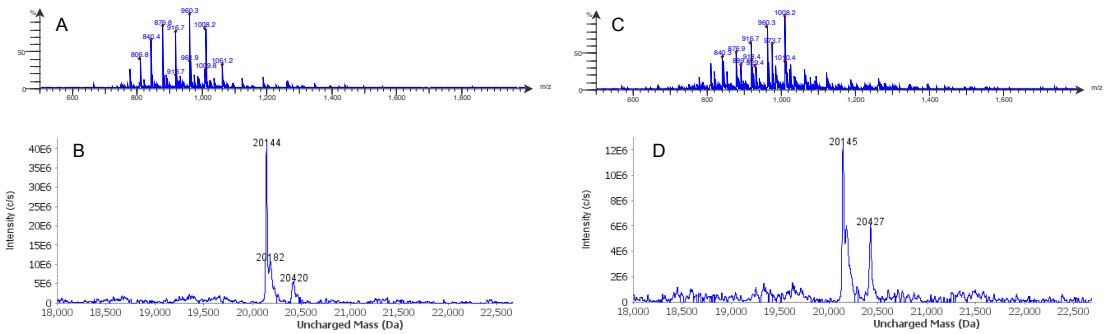


Figure 37. (A) Mass spectrum of the single peak found in labeling reactions for an unmodified sample of **batch 3** $SrtA_{pneu}$ after 24 hours of incubation with the NO_2-alkyne sulfenic acid probe. (B) Deconvolution of the mass spectrum in (A), which returns an uncharged mass of 20144 Da as the major species, matching that of unmodified $SrtA_{pneu}$. (C) Mass spectrum of the single peak found in labeling reactions for a sample of **batch 3** $SrtA_{pneu}$ oxidized with 1.5 eq of H_2O_2 and incubated for 24 hours with the NO_2-alkyne sulfenic acid probe. (D) Deconvolution of the mass spectrum in (C) which shows two uncharged species in the sample. One species represents unmodified $SrtA_{pneu}$ (20145 Da) and the other represents $SrtA_{pneu}$ with the conjugated sulfenic acid probe NO_2-alkyne (20427 Da).

2.3 Conclusions and Future Directions.

This section detailed the progress from inactive samples of $SrtA_{pneu}$ to fully active samples of the enzyme. After evaluation of inactive $SrtA_{pneu}$ samples showed a high degree of oligomerization, we began exploring methods of reducing the assembled forms of the enzyme. While heating samples of purified $SrtA_{pneu}$ returned activity to the samples and removed most of the oligomerized enzyme, we believed more efficient procedures could be developed to achieve the same result. By adding a denaturing step to the purification, we were able to decrease some oligomerization in the $SrtA_{pneu}$ samples. SE-FPLC purification of $SrtA_{pneu}$ after disassembly of oligomers proved to be an effective method to

produce monomeric enzyme stocks. While we expected this to improve the catalytic activity of the preparations, activity was only regained after addition of the reducing agent TCEP to the purification buffer. The addition of TCEP to the total purification also improved the monomer:oligomer ratio in the samples, which afforded more available monomer for isolation by SEC. This purification procedure has now consistently yielded active batches of enzyme over multiple expressions in the lab.

As noted previously, inclusion of a reducing agent in the $SrtA_{pneu}$ buffers modifies the intensity of an interesting doublet that appears in monomer region of native PAGE gels. Previously, $SrtA_{pyogenes}$ has been shown to have an altered conformation upon oxidation of the active site cysteine to sulfenic acid.[75] We hypothesize that reducing conditions reverse some form of oxidation in $SrtA_{pneu}$, which ultimately leads to differences in protein conformation. Both oxidized and unoxidized forms of the protein could be present in solution, creating two alternative conformers which would appear as separate bands in native PAGE analysis. Additionally, based on the observation that reducing conditions diminish the presence of oligomerized $SrtA_{pneu}$ in solution, oxidation of the enzyme may facilitate a conformation that is more susceptible to interactions with other enzyme units in solution.

This thesis also marks the first account of oligomerized $SrtA_{pneu}$, along with the first attempts to understand the cause and dissolution of the assembled enzyme. The oligomerization of this SrtA homolog appears to be driven by highly

favorable interactions as determined by dilution experiments of these samples, where no effect on oligomer concentration in solution was observed. Interconversion between the dimer and monomer states of the enzyme appears to have a high activation barrier, as it can only be disrupted by heating or complete unfolding of the enzyme samples. Further, recovered monomeric $SrtA_{pneu}$ does not appear to rapidly re-oligomerize, also supporting the hypothesis of a high activation barrier between the two states. Hypothetically, this behavior could be attributed to domain swapping of $SrtA_{pneu}$ monomers, evidence of which comes from the structure of $SrtA_{pneu}$ in a domain swapped conformation (PDB ID: 4O8L).

Three dimensional domain swapping is an interaction between two or more identical protein monomers, where one or more domains or segments of the monomers rotate out of the core structure and take the place of the same amino acids in the opposite monomer.[104] This interaction generally results in identical interactions in both the closed monomer and domain swapped forms of the proteins, but may also result in additional favorable interactions within the domain swapped oligomers.[104] *In vivo,* the open state that propagates oligomerization is most likely generated when monomers are transiently exposed to denaturing conditions, such as an acidic compartment within the cell.[104] The open monomers become free to form domain swapped interactions, which creates the oligomerized forms of the protein. The process of converting between either monomer or oligomer most likely has a high energy barrier, as favorable

contacts must be broken to rotate protein domains out of the core structure.[104] This allows for different states of the same protein to be present in solution. Domain swapping has been simulated *in vitro* to afford not only dimeric proteins,[105] but also trimeric[106, 107] and higher order[107, 108] assembly states, which suggests this phenomenon could produce the full range of oligomers seen in samples of $SrtA_{pneu}$. These observed qualities of domain swapped proteins, in addition to the domain swapped structure of $SrtA_{pneu}$, align with the data presented in this thesis, and suggests that $SrtA_{pneu}$ undergoes domain swapping to form oligomerized enzyme. It would be beneficial to perform experiments probing the dissociation constant of the monomer-dimer equilibrium, as these could aide our understanding of the strength of the association between monomers. Additionally, more rigorous analysis of alternative purification conditions could be done to aide in our understanding of the mechanism of domain swapping in $SrtA_{pneu}$, while also potentially simplifying purification of the enzyme.

Future directions for this aspect of the project revolve around crystallization of the monomeric preparations of $SrtA_{pneu}$. While trials towards this have already been undertaken, more work is needed to develop the conditions necessary for crystal growth.

Chapter 3. Design and Synthesis of Ketomethylene-based Substrate Analogs

3.1 Design of Ketomethylene Isosteres

Concurrent with our efforts to produce reliable preparations of active $SrtA_{pneu}$ suitable for structural characterization, we have also undertaken the synthesis of substrate mimics that can be covalently anchored in the enzyme active site. To this end, we have designed substrate analogs in which the scissile amide bond of standard sortase substrates is replaced with a non-cleavable ketomethylene dipeptide isostere (**Figure 38**).

We anticipated that ketomethylene isosteres of $SrtA_{pneu}$ substrates would serve as non-cleavable analogs for obtaining a structure of the substrate bound state. The proposed ketomethylene analogs for this purpose have been previously used as potent inhibitors of serine proteases,[109] which share mechanistic characteristics with sortases. We propose that the inserted ketone will be able to form a covalent bond with the active site cysteine which will be further stabilized by the remaining contacts present in standard sortase substrates. The synthetic processes laid out here allow for variability in the substitution present at either site in the ketomethylene-linked dipeptide. Further, the methods for developing these molecules allow for the production of fluorenylmethyloxycarbonyl (Fmoc) protected N-termini[109, 110], affording products ready for incorporation into peptides via solid phase synthesis.

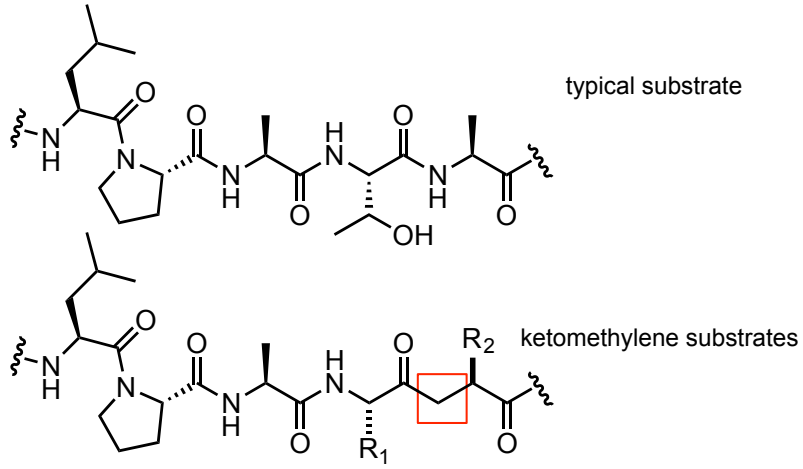


Figure 38. Comparison of a normal peptide substrate and a substrate with a ketomethylene linkage inserted between the 4th and 5th positions of the sorting sequence. The red box denotes the methylene substitution in the amide bond. R_1 and R_2 can be any amino acid functional group and are only limited by the synthetic process.

3.2 Preliminary Ketomethylene Analog Synthesis and Testing

To evaluate the ability of a ketomethylene-linked peptide interacting with the active site of $SrtA_{pneu}$, we first synthesized a substrate analog incorporating commercially available 5-aminolevulinic acid. While this unit was not anticipated to be optimal with respect to enzyme recognition due to the lack of sidechains in position 4 and 5 of the substrate mimic, we elected to start with this compound as a proof of concept study. As described by Rogers *et al.,*[111] synthesis began with solution-phase acylation of 5-aminolevulinic acid with the N-hydroxysuccinamide ester of Fmoc-Ala-OH (Fmoc-Ala-OSu) to produce an Fmoc-protected ketomethylene tripeptide analog of Ala-Gly-Gly, hereafter referred to as Fmoc-A[G(keto)G]. This material was then used in solid-phase peptide synthesis

(SPPS) to generate a full size substrate analog (**1**).[112] The identity of **1** was verified by ESI-MS, the material was purified to homogeneity by RP-HPLC, and the peptide was lyophilized prior to use **(Figure 39).**

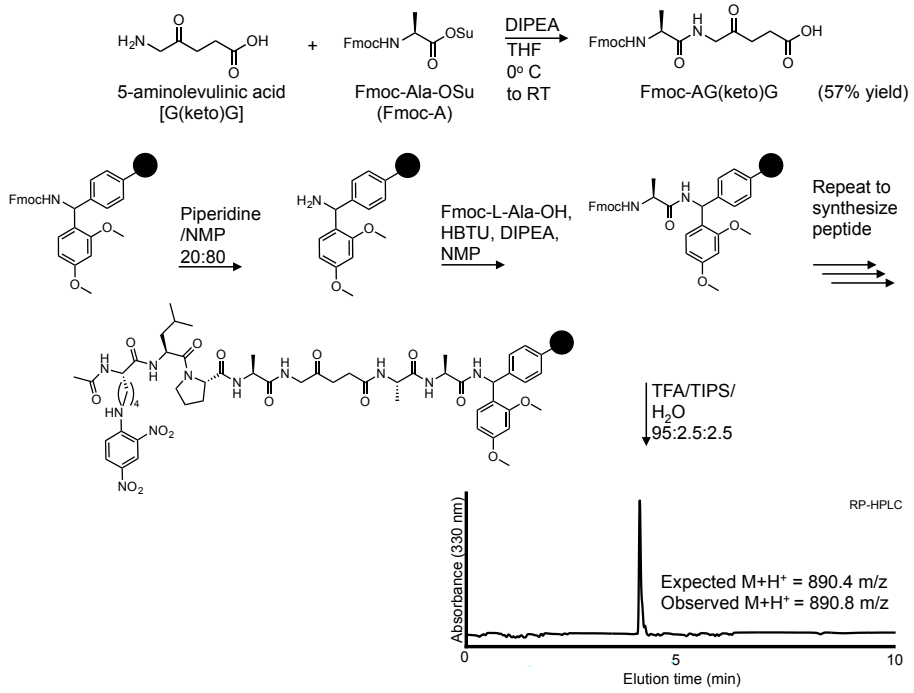


Figure 39. Synthesis and analysis of the preliminary ketomethylene substrate analog using the AG(keto)G block to install a non-cleavable linkage between the 4th and 5th residues.

With purified **1** in hand, we proceeded to test its inhibitory activity using a model SML reaction. Specifically, a heat-treated stock of $SrtA_{pneu}$ with demonstrated SML activity was combined with a standard Abz-LPATA-GK(Dnp) substrate, L-alanine amide, and varying concentrations of ketomethylene analog **1** (**Figure 40**). The nucleophile L-Alanine amide was used in these reactions to avoid the oxime formation observed with hydroxyamines and ketones, as would

be expected for the model reaction system used in Chapter 2 of this thesis. Analysis via RP-HPLC suggested that analog **1** was able to act as an inhibitor, reducing reaction conversion at as low as 0.1 eq in relation to the Abz-LPATA-GK(Dnp) substrate. Reaction progress was diminished by less than 50% at 1 eq of inhibitor, and reached as little as 30% relative conversion at 5 eq. Aside from the single expected peak from the inhibitor itself, no additional peaks developed when the inhibitor was added, demonstrating that it was not being cleaved by the enzyme. Overall, these data indicated that ketomethylenes can serve as non-cleavable isosteres in SML, which in turn could be useful for structural characterization of enzyme/substrate complexes.

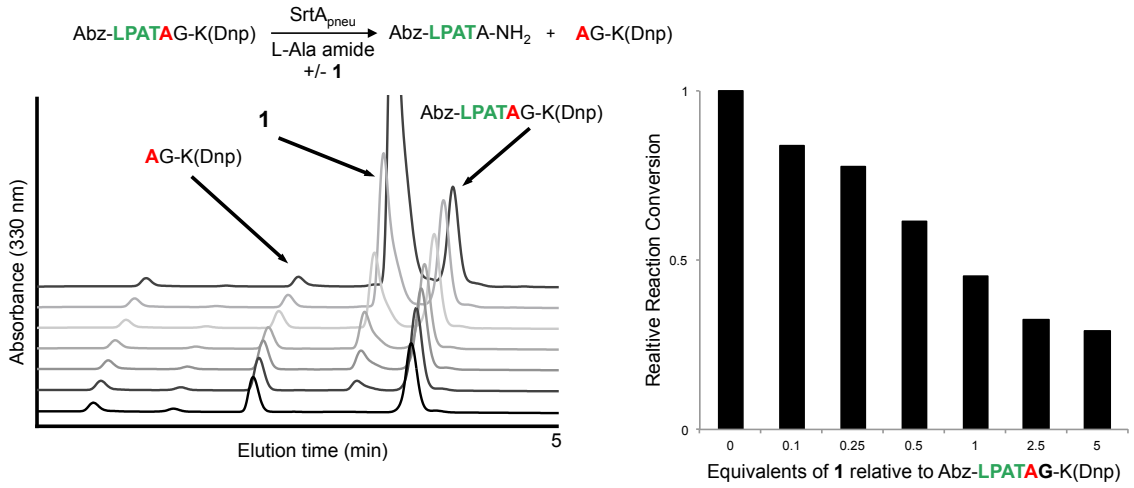


Figure 40. RP-HPLC analysis for the inhibition of transpeptidation by ketomethylene substrate analog **1**. These reactions were assembled with undiluted, 7 day heat-treated SrtA$_{pneu}$ and allowed to incubate at room temperature for 24 hrs before quenching with N-ethylmaleimide (NEM) for analysis by HPLC.

3.3 Second Generation Design of Ketomethylene Substrate Analogs

Having successfully shown that **1** was able to impede an *in vitro* SML reaction, we next sought to design derivatives with substituents that mimic the side chains of position 4 and 5 in the sortase substrate motif. Ideally, these substituents would match the side chains of threonine and alanine residues for optimal structural overlap with the known LPATA substrate of $SrtA_{pneu}$. However, for synthetic ease we opted to focus on simple methyl groups, which would mirror the side chains of alanine residues. While it was known from previous work in our group that alanine was well tolerated in position 5 (see **Figure 4**), we wanted to verify that alanine was accepted at position 4 or in combination with alanine at position 5. To test this, peptide substrates displaying LPATAG, LPAAGG, and LPAAAG were prepared via SPPS and tested in model SML reactions. As shown in **Figure 41**, all variants were accepted by $SrtA_{pneu}$, however, the LPAAGG (53% conversion) and LPAAAG (20% conversion) substrates were not processed to the same extent as LPATAG (88% conversion). Nonetheless, these results suggested that simple methyl groups in positions 4 and 5 should suffice for enzyme recognition. Based on these results, two additional ketomethylene-linked peptide isosteres were designed to serve as substrate analogs of LPAAG (**2**) and LPAAA (**3**) (**Figure 42**).

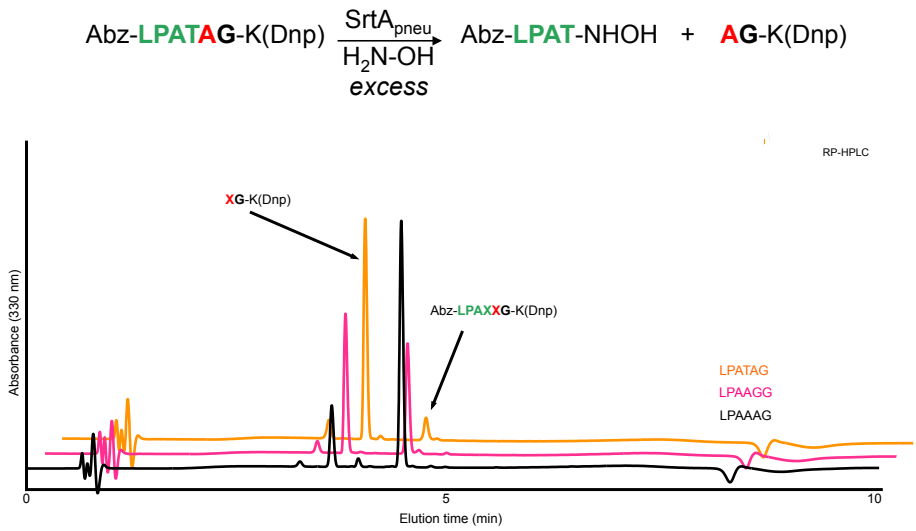
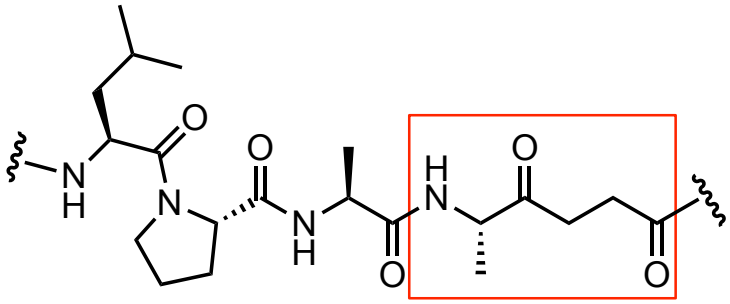
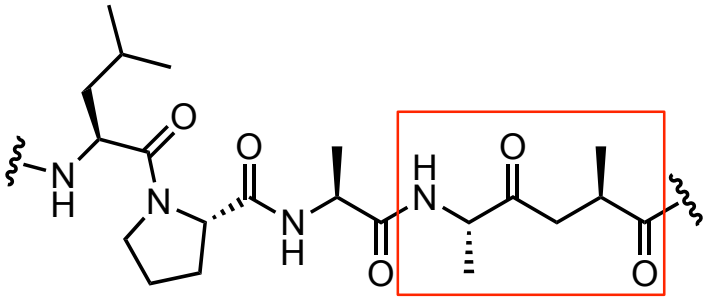


Figure 41. Analytical RP-HPLC traces showing the variable activity of $SrtA_{pneu}$ with substrates differing from its preferred sequence of LPATA. These reactions were incubated for 24 hrs at room temperature using a monomeric fraction of $SrtA_{pneu}$ from FPLC purification with TCEP.



2



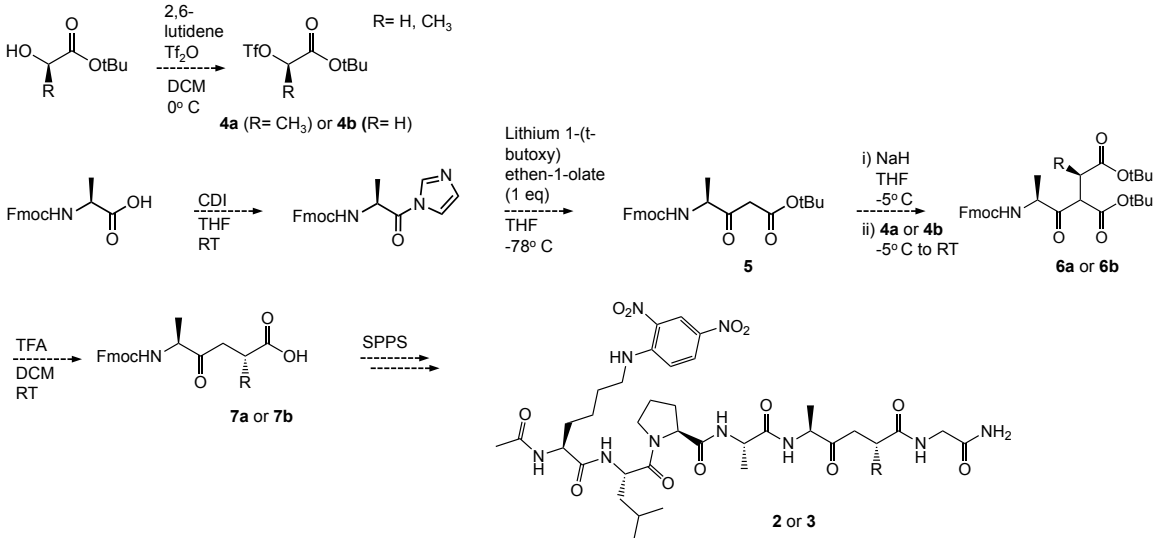
3

Figure 42. Structures of the next generation of ketomethylene-based substrate analogs for $SrtA_{pneu}$. **2** represents the ketomethylene analog of LPAAG and **3** represents the analog of LPAAA. These pseudopeptides will include substitution at the amino acid positions surrounding the non-cleavable ketomethylene linkage used to replace the scissile peptide bond.

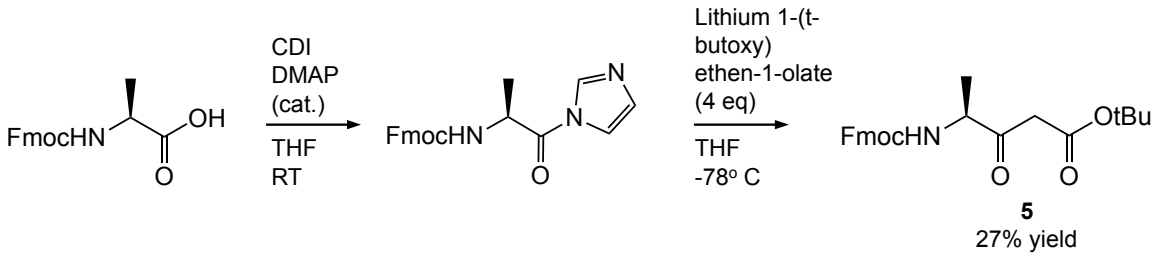
3.4 Progress Toward Synthesis of Substituted Ketomethylene Isosteres

To access structures **2** and **3**, we began with a procedure from Budnjo *et al.*[109] towards synthesizing the ketomethylene linked dipeptide isosteres A(keto)G and A(keto)A to be incorporated into full substrate analogs. As shown in **Scheme 1**, our proposed synthetic process involved addition of the lithium enolate of t-butyl acetate to carbonyldiimidazole (CDI)-activated Fmoc-Ala-OH. The resulting β-ketoester (**5**) would then serve as a nucleophile in a stereospecific substitution of triflates derived from either t-butyl *R*-lactate (**4a**) or t-butyl glycolate (**4b**).

Treatment with TFA was anticipated to remove the t-butyl ester protecting groups, followed by in situ decarboxylation to generate the desired keto methylene building blocks **7a** or **7b**.



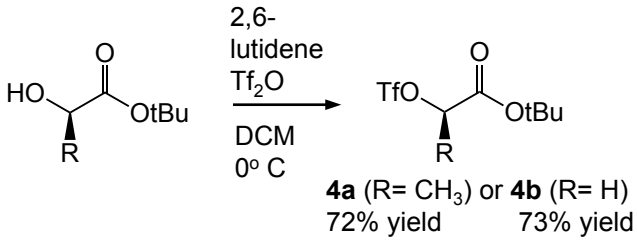
Scheme 1. Overview of the synthetic scheme from Budnjo *et al*.[109] using an Fmoc-protected amino acid starting material.



Scheme 2. Synthesis of the Fmoc-protected ketoester

Efforts to execute the plan in **Scheme 1** began with lithium enolate addition to CDI-activated Fmoc-Ala-OH following the reported protocol from Budnjo *et al*. However, in our hands this approach repeatedly failed to generate the desired product, instead leading to Fmoc deprotection. An alternative

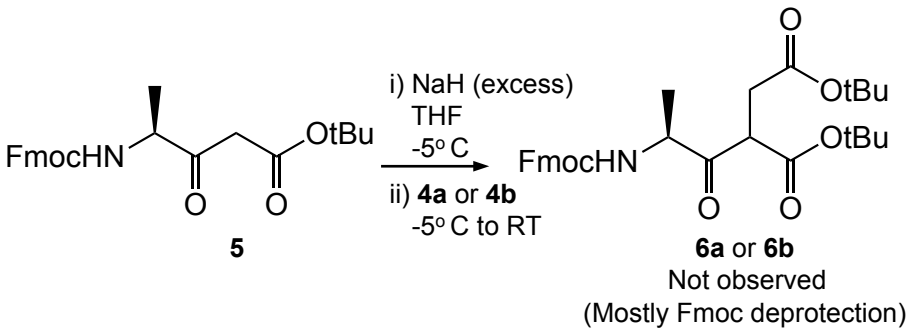
procedure from Mathieu *et al*.[110] involving catalytic 4-dimethylaminopyridine (DMAP) to activate the amino acid and excess t-butyl acetate to generate the enolate, was substituted for the procedure by Budnjo *et al*. (**Scheme 2)** which enabled us to synthesize the desired product **5** in 27% yield. With the ketoester in hand, we then prepared triflates **4a** (72% yield) and **4b** (73% yield) using triflic anhydride and 2,6-lutidine (**Scheme 3**) according to the procedure from Budnjo *et al*.



Scheme 3. Synthesis of triflate compounds from t-butyl esters with varying substitution.

Having successfully synthesized the triflate compounds, we attempted the SN2 substitution using NaH to deprotonate the ketoester (**Scheme 4**) also described by Budnjo *et al*. Despite repeated attempts using varying equivalents of NaH and triflate **4a**, the reaction proved entirely unsuccessful, with either Fmoc deprotection or unreacted starting material being the only consistent results observed. Since no product was observed, we decided to attempt the final step of the synthetic scheme with triflate **4b**. Since this compound lacks the extra methyl group of **4a**, we believed it would behave as a better electrophile due to its relative lack of steric hindrance around the reactive center. This compound still failed to produce significant amounts of product **6b**, and attempts to improve

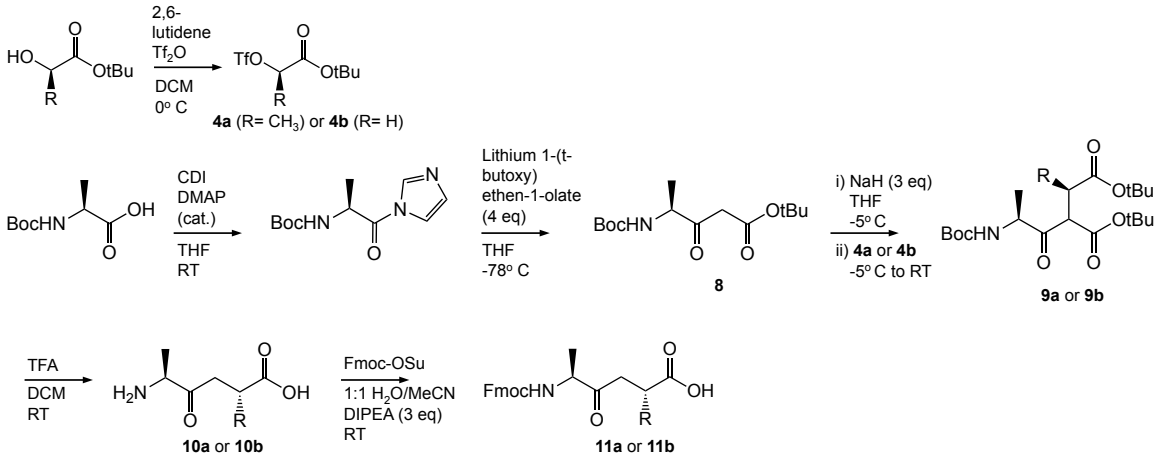
reaction turnover by including slight excesses of NaH only resulted in greater deprotection of the Fmoc-Ala t-butyl ketoester.



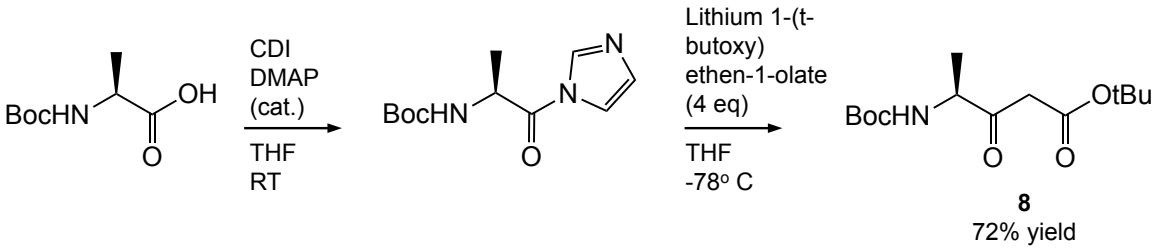
Scheme 4. Synthetic process to generate the fully protected ketomethylene

After these many failed attempts at using the Fmoc-protected starting material, we elected to begin the synthesis with *tert*-butyloxycarbonyl protected Ala-OH (Boc-Ala-OH), which was anticipated to retain the Boc protecting group throughout each set of reaction conditions. We refrained from this initially because it involved the deprotection and reprotection of the N-terminus of the ketomethylene, however, we determined that an extra step in the synthesis was more desirable than being unable to obtain the desired product.

The synthetic scheme to generate the desired A(keto)G dipeptide isostere starting with a Boc protected amino acid is described in **Scheme 5**. Boc-Ala-OH was combined with the lithium enolate of t-butyl acetate to generate the Boc-protected ketoester **8** (**Scheme 6**) which was easily purified from the extracted reaction by flash column chromatography (FCC) in 72% yield. Triflates were synthesized from the procedure used previously and used without further purification.



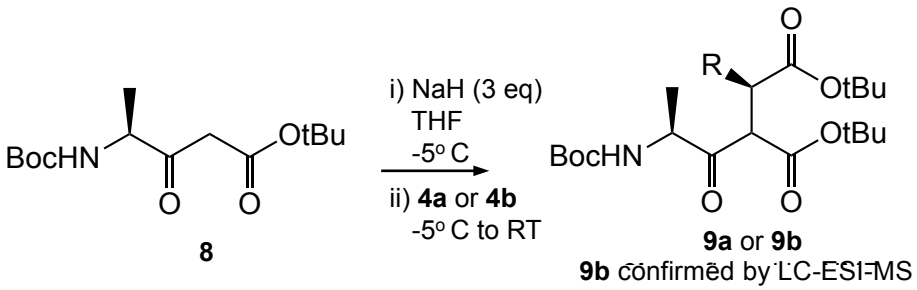
Scheme 5. Overall synthetic scheme to produce Fmoc-protected ketomethylene dipeptide isosteres using Boc amino acids as starting materials.



Scheme 6. Synthesis of the Boc-protected ketoester

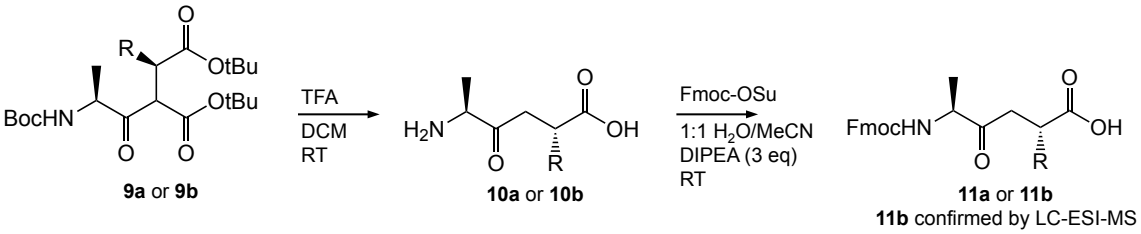
The final step of this synthesis was achieved by deprotonating ketoester **8** using excess NaH (**Scheme 7**) and combining the resulting mixture with triflates **4a** or **4b**. Initial attempts at this reaction using triflate **4b** failed without the addition of excess NaH, where substantial amounts of starting materials remained unreacted even after 24 hours of stirring. After determining the necessity for excess NaH, the reaction succeeded and the **product 9b** was easily purified from the reaction mixture. Unfortunately, ^{1}H nuclear magnetic

resonance (NMR) spectroscopy of this product was too convoluted to accurately assign peaks, however, LC-ESI-MS was able to confirm the mass of the Na^+ adduct of this material.



Scheme 7. Synthesis of the fully protected Boc-ketomethylene.

After removal of the Boc and t-butyl groups with trifluoroacetic acid (TFA) and *in situ* decarboxylation to afford the deprotected ketomethylene **10b**, reprotection of the N-terminus was carried out using Fmoc-OSu (**Scheme 8**). The use of excess diisopropylethylamine (DIPEA) was necessary to maintain the deprotonated state of the ketomethylene, as we were unable to completely remove the TFA from the samples, which we anticipated consumed much of the base before it was able to increase the pH of the reaction mixture. The mass of the N-protected ketomethylene **11b** was confirmed by mass spectrometry of the crude mixture. A ^1H NMR spectrum was also obtained (see Appendix I), but was not analyzed due to contaminants, as these made it impossible to confidently assign peaks.



Scheme 8. Synthesis of disubstituted Fmoc protected ketomethylene

Current efforts on this project involve the producing larger quantities of **11b**, followed by incorporation of the material into SPPS procedures to generate the aforementioned substrate analog **2**. After synthesis of this analog, analysis on its inhibitory capacity with $SrtA_{pneu}$ will be performed, which we anticipate to be more effective than that displayed by the preliminary substrate analog **1.**

4. Experimental

Protein Expression and Purification

Full sequence of $SrtA_{pneu}$ protein used in this study:

MESSHHHHHHAVLTSQWDAQKLPVIGGIAIPELEMNLPIFKGLDNVNLFYGAGT
MKREQVMGEGNYSLASHHIFGVDNANKMLFSPLDNAKNGMKIYLTDKNKVYTY
EIREVKRVTPDRVDEVDDRDGVNEITLVTCEDLAATERIIVKGDLKETKDYSQTS
DEILTAFNQPYKQFY

Expression and purification of $SrtA_{pneu}$ batches 1 and 3. The $SrtA_{pneu}$ expression vector was obtained by commercial gene synthesis in a pJ414 expression vector from DNA2.0 (Menlo Park, CA). The plasmid containing $SrtA_{pneu}$ was transformed into BL21 (DE3) chemically competent *E. coli* by heat shock. Following addition of 1 mL Luria-Bertani (LB) broth, the cells were incubated for 1 hour at 37 ^{o}C with shaking. Cells were the plated on LB-agar plates containing 100 µg/mL ampicillin, and incubated overnight at 37 ^{o}C. A single isolated colony was selected and used to inoculate 50 mL of LB broth (containing 100 µg/mL ampicillin), which was then placed in a 37 ^{o}C shaking incubator to grow overnight. Roughly 30 mL of this starter culture were then added to 1 L of LB broth (containing 100 µg/mL ampicillin) to initiate large-scale growth. This culture was allowed to grow to an OD_{600} reading of 0.7-0.8 at 37 ^{o}C in a shaking incubator. Protein expression was then induced by the addition of isopropyl-β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cells remained at 37 ^{o}C with shaking for 3 hours to express $SrtA_{pneu}$, and were then isolated by centrifugation at 10000xg. The cell pellets were then resuspended in 30 mL lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.5 mM

EDTA), treated with lysozyme at a final concentration of 10 mg/mL, and incubated at room temperature for 1 hour. This mixture was sonicated for 2, 30 second intervals at 50% output and the lysate was then clarified by centrifugation at 20000xg. This clarified lysate was added to a 5 mL His-Bind (Thermo-Fisher) Ni-NTA column equilibrated in wash buffer (20 mM Tris pH 8.0, 150 mM NaCl, 20 mM imidazole). Bound protein was further washed with 10 column volumes of wash buffer, then eluted in two, 1 column volume aliquots of elution buffer (20 mM Tris pH 8.0, 150 mM NaCl, 300 mM imidazole). These preparations were dialyzed against dialysis buffer (20 mM Tris pH 8.0, 150 mM NaCl) to remove imidazole. Glycerol was added to 10% v/v before the samples were stored at -80 ^{o}C. Collected fractions were analyzed by native and SDS-PAGE. The activity of these samples was evaluated by model sortase-mediated ligation (SML) reactions with the known substrate LPATA. Were described, FPLC analysis of these samples was performed Samples of $SrtA_{pneu}$ were subjected to various incubation temperatures for various times before purification and analysis on an NGC FPLC system (Bio-Rad) by size-exclusion chromatography using an Enrich SEC 70 column (Bio-Rad) with 50 mM Tris pH 8.0, 150 mM NaCl buffer as the eluent at 0.5 or 0.75 mL/min. Variable temperature expressions described in Chapter 2 were carried out by this procedure, except during expression, where the temperature was adjusted to 16 or 25 °C.

Expression and purification of $SrtA_{pneu}$ batch 2. The $SrtA_{pneu}$ expression vector was obtained by commercial gene synthesis in a pJ414 expression vector

from DNA2.0 (Menlo Park, CA). The plasmid containing $SrtA_{pneu}$ was transformed into BL21 (DE3) chemically competent *E. coli* by heat shock. Following addition of 1 mL LB broth, the cells were incubated for 1 hour at 37 ^{o}C with shaking. Cells were the plated on LB-agar plates containing 100 µg/mL ampicillin, and incubated overnight at 37 ^{o}C. A single isolated colony was selected and used to inoculate 50 mL of LB broth (containing 100 µg/mL ampicillin), which was then placed in a 37 ^{o}C shaking incubator to grow overnight. Roughly 30 mL of this starter culture were then added to 1 L of LB broth (containing 100 µg/mL ampicillin) to initiate large-scale growth. This culture was allowed to grow to an OD_{600} reading of 0.7-0.8 at 37 ^{o}C in a shaking incubator. Protein expression was then induced by the addition of IPTG to a final concentration of 1 mM. Cells remained at 37 ^{o}C with shaking for 3 hours to express $SrtA_{pneu}$, and were then isolated by centrifugation at 6000xg. The cell pellets were then resuspended in 30 mL lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.5 mM EDTA), treated with lysozyme at a final concentration of 10 mg/mL, and incubated at RT for 1 hour. This mixture was sonicated for 2, 30 second intervals at 50% output and the lysate was then clarified by centrifugation at 18000xg. This clarified lysate was added to 5 mL of His-Bind (Thermo-Fisher) Ni-NTA resin in a conical tube and imidazole was added to 20 mM before the contents were mixed for 30 min at room temperature. The slurry was centrifuged at 700xg for 2 min and the supernatant removed. The resin pellet was resuspended in 25 mL of wash buffer (20 mM Tris, pH 8.0, 150 mM NaCl, 20 mM imidazole) and transferred to an

empty column. After settling, the wash buffer was drained to the resin bed and 25 mL of wash buffer was passed through the column. Bound protein was then eluted in two, 5 mL aliquots of elution buffer (20 mM Tris pH 8.0, 150 mM NaCl, 300 mM imidazole). To further purify the eluted protein, elution 1 of the initial purification was diluted 10-fold with dilution buffer (20 mM Tris, pH 8.0, 150 mM NaCl) and recirculated through a 5 mL Ni-NTA column. The bound protein was further washed with 10 mL of wash buffer and eluted in 3, 5 mL portions of elution buffer. The loading flowthrough of this purification was also reconcentrated using a 5 mL Ni-NTA column. After loading, the bound protein was washed with 15 mL wash buffer and and eluted in two, 5 mL aliquots of elution buffer. These preparations from the 2nd and 3rd purification steps were dialyzed against dialysis buffer (20 mM Tris pH 8.0, 150 mM NaCl) to remove imidazole. Glycerol was added to 10% v/v before the samples were stored at -80 ^{o}C. Collected fractions were analyzed by native and SDS-PAGE. The activity of these samples was evaluated by model sortase-mediated ligation (SML) reactions with the known $SrtA_{pneu}$ known substrate LPATA.

$SrtA_{pneu}$ expression, non-reducing/denaturing purification and refolding. 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 ^{o}C overnight. Roughly 30 mL of culture were then added to one 1 L of LB broth containing 100 µg/mL ampicillin to initiate large-scale growth. This culture was allowed to grow to an OD_{600} reading of 0.7-0.8 at

37 ^{o}C in a shaking incubator before induction with 1 mM IPTG. Cells remained at 37 ^{o}C with shaking for three hours to express $SrtA_{pneu}$, and were then isolated by centrifugation at 6000xg. Cell pellets were resuspended in 30 mL denaturing lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.5 mM EDTA, 8 M urea). The resuspended cells were sonicated for 2, 30 second intervals at 50% power output and the lysate was clarified by centrifugation at 18000xg. This clarified lysate was added to a 5 mL His-Bind resin (Thermo-Fisher) column pre-equilibrated in denaturing wash buffer (50 mM Tris pH 8.0, 150 mM NaCl, 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 8.0, 150 mM NaCl, 300 mM imidazole, 8 M urea). The first eluted fraction was then rapidly diluted by addition to a 100x volume of dilution buffer (50 mM Tris pH 8.0, 150 mM NaCl). This material was then recirculated through a 5 mL Ni-NTA column equilibrated in wash buffer (50 mM Tris pH 8.0, 150 mM NaCl, 20 mM imidazole). Bound protein was further washed with 10 column volumes of wash buffer, then eluted in two 1 column volume aliquots of elution buffer (50 mM Tris pH 8.0, 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 50 mM Tris pH 8.0, 150 mM NaCl buffer as the eluent at 0.5 mL/min. Monomeric protein fractions were pooled, and if necessary, concentrated using centrifugal concentrators. The activity of these samples was

evaluated by model sortase-mediated ligation (SML) reactions with the known substrate LPATA. These samples were stored at 4 ^{o}C or -20 ^{o}C.

$SrtA_{pneu}$ expression, reducing/denaturing purification and refolding. The optimized procedure for the expression and purification of $SrtA_{pneu}$ is identical to that for the non-reducing/denaturing purification described above, with the exception of 1 mM TCEP being included in all buffers used.

Evaluation of protein concentration. UV-vis spectroscopy for determining concentrations of the prepared samples was performed on a Nanodrop^{TM} ND-1000 spectrophotometer (Thermo Scientific) at 280 nm using 17,440 $M^{-1}cm^{-1}$ 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 Å C4 column (2.0 x 100 mm) with the method (Method A): MeCN (0.1% formic acid)/95% H_2O, 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 (0.5-6.0 min), hold 90% MeCN (6.0-7.0 min), 90% MeCN to 10% MeCN (7.0-7.1 min), re-equilibrate to 10% MeCN (7.1-10.0 min). Data analysis was done in Advion Data Express software version 3.0. Mass spectrum deconvolution was achieved through a max entropy

goodness of fit algorithm to determine uncharged masses of samples.

Analysis of $SrtA_{pneu}$ Transpeptidation Activity

Evaluation of enzymatic activity with various peptides containing the fluorophore-quencher pair 2-aminobenzoyl (Abz) and 2,4-dinitrophenol (Dnp) were carried out at room temperature with the concentrations shown in the table below for the model sortase mediated ligation reactions. Specific additives are noted where necessary within the presentation of the data. Reactions were prepared by combining all components shown in **Table 4** except sortase, which was added to initiate the reaction. Conversion was analyzed by UV/VIS analysis of analytical RP-HPLC using a Dionex Ultimate 3000 HPLC system (Thermo Scientific) with a Phenomenex Kinetex 2.6 µm, 100 Å C18 column (3.0 x 100 mm) with the method (Method B): MeCN (0.1% TFA)/H_2O (0.1% TFA) mobile phase. Flow rate = 0.7 mL/min. Gradient = 10% MeCN (0.0-0.5 min), 10% MeCN to 90% MeCN (0.5-6.0 min), hold 90% MeCN (6.0-7.0 min), 90% MeCN to 10% MeCN (7.0-7.1 min), re-equilibrate to 10% MeCN (7.1-10.0 min).

Percent conversion was calculated by dividing the area of the product peak by the addition of total 2,4-DNP containing reactant and product peak areas at 330 nm.

Table 2. Reaction conditions for sortase mediated ligation. Water was added to 50 uL total reaction volume unless otherwise noted.

Stock Solution	Reaction Concentration
Buffer (500 mM Tris, pH 7.5, 150 mM NaCl)	50 mM Tris, pH 7.5, 150 mM NaCl
Substrate (1:1 DMSO/Water) (0.5-4 mM)	40 or 100 μM
Nucleophile (Water) (10 mM)	10 mM
$SrtA_{pneu}$ (50-300 μM)	10 μM

Peptide Synthesis

All chemicals were obtained from commercial sources and were used without further purification. 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 API 2000 Triple Quadrupole mass spectrometer (Applied Biosystems) and an Agilent 1100 HPLC system equipped with a Phenomenex Aeris Widepore 3.6 μm 200 Å XB-C8 column (4.6 x 150 mm). All samples were analyzed using the following method (Method C): H_2O (0.1% formic acid) / organic (95% MeCN, 5% isopropanol, 0.1% formic acid) mobile phase. Flow rate = 1.25 mL/min. Gradient = 10% organic (0.0-1.0 min), 10% organic to 90% organic (1.0-10.0 min).

Analytical separations for UV/vis analysis were performed with a Phenomenex Kinetex 2.6 μm, 100 Å C18 column (3.0 x 100 mm) with the method (Method B): MeCN (0.1% TFA)/H_2O (0.1% TFA) mobile phase. Flow rate = 0.7 mL/min. Gradient = 10% MeCN (0.0-0.5 min), 10% MeCN to 90% MeCN (0.5-6.0

min), hold 90% MeCN (6.0-7.0 min), 90% MeCN to 10% MeCN (7.0-7.1 min), re-equilibrate to 10% MeCN (7.1-10.0 min).

Semi-preparative separations for 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 one of the following methods: (Method D): MeCN (0.1% TFA)/H_2O (0.1% TFA) mobile phase. Flow rate = 0.5 to 4.0 mL/min (0.0-2.0 min), hold 4.0 mL/min (2.0-17.01 min), 4.0 to 0.5 mL/min (17.01-19.0 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) or (Method E): MeCN (0.1% TFA)/H_2O (0.1% TFA) mobile phase. Flow rate = 0.5 to 4.0 mL/min (0.0-2.0 min), hold 4.0 mL/min (2.0-17.01 min), 4.0 to 0.5 mL/min (17.01-19.0 min). Gradient = 30% MeCN (0.0-2.0 min), 30% MeCN to 60% MeCN (2.0-15.0 min), 60 to 90% MeCN (15.0-15.01 min), hold 90% MeCN (15.01-17.0 min), 90% MeCN to 10% MeCN (17.0-17.01 min), re-equilibrate to 10% MeCN (17.01-19.0 min)

Solid-phase peptide synthesis. 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 hr-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. Acetyl capping of the N-terminus was achieved by combining acetic anhydride (0.3 mmol), DIPEA, and NMP, 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) fluorphore-quencher pair to simplify analysis by UV/vis spectroscopy. Dnp was added as a conjugate to ε-amine of a lysine side chain [Fmoc-K(Dnp)]. 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 H_2O, 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 5000xg for 5 min and the ether discarded to afford a peptide pellet, which was dried under vacuum for 24 hrs. Peptides were solubilized by 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 either method D or E and the molecular mass of the peptides verified via LC-ESI-MS method C. 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 HPLC analysis using method B. Concentrations were estimated by UV/Vis spectroscopy on a Nanodrop^{TM} ND-1000 spectrophotometer (Thermo

Scientific) at 365 nm using the molar extinction coefficient 17,400 $M^{-1}cm^{-1}$ for the Dnp chromophore.

Synthesis of Small Molecules

All chemicals were obtained from commercial sources and were used without further purification. NMR spectra were collected with a Brüker Avance spectrometer at 500 MHz for ^{1}H. FID processing and figure generation was done using Mestrelab MestReNova software version 10.0.2-15465. 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 Å C18 column (2.0 x 100 mm) with the method (Method F): MeCN (0.1% formic acid)/95% H_2O, 5% MeCN (0.1% formic acid) mobile phase. Flow rate = 0.4 mL/min. Gradient = 5% MeCN (0.0-0.5 min), 5% MeCN to 90% MeCN (0.5-6.0 min), hold 90% MeCN (6.0-7.0 min), 90% MeCN to 10% MeCN (7.0-7.1 min), re-equilibrate to 10% MeCN (7.1-12.0 min).

Ala-5-ALA synthesis. Fmoc-Ala-OSu (1.2 g, 3.0 mmol) and 5-aminolevulinic hydrochloride (0.50 g, 3.0 mmol) were added to dry THF (40 mL) and cooled to 0 ^{o}C. DIPEA (0.52 mL, 3.0 mmol) was dissolved in dry THF (20 mL) and then added dropwise to the stirred suspension across 120 minutes. The

reaction was allowed to return to room temperature and stirred for 23 days after which the mixture was filtered, concentrated by rotary evaporation and dissolved into ethyl acetate (50 mL). The organic layer was washed with water (3x, 50 mL) and extracted into 5% $NaHCO_3$ (3x, 25 mL). The aqueous layer was then collected and acidified with 1M HCl until no further precipitation occurred. This heterogeneous mixture was then extracted with ethyl acetate (3x, 75 mL) and dried over $MgSO_4$. The solvent was removed by rotary evaporation to produce a white solid which was used without further purification. IUPAC Name: (*S*)-5-(2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)propanamido)-4-oxopentanoic acid (0.72 g, 57% yield). ^{1}H NMR (500 MHz, D_6-DMSO): δ 12.10 (br. s, 1H), 8.13 (m, 1H), 7.88 (d, *J* = 7.6, 2H), 7.72 (t, *J* = 7.7, 2H), 7.55 (d, *J* = 7.7, 1H), 7.40 (t, *J* = 7.3, 2H), 7.32 (t, *J* = 7.5, 2H), 4.26-4.18 (m, 3H), 4.11-4.05 (m, 1H), 3.98-3.87 (m, 2H), 2.63 (t, *J* = 6.2, 2H), 2.39 (t, *J* = 6.5, 2H), 1.23 (d, *J* = 7.2, 3H).

Ketoester synthesis. Boc-Ala-OH (7.0 mmol) was dissolved in dry THF (20 mL), to which CDI (1.08 g, 7.7 mmol) was added under stirring in three portions, resulting in bubble formation. Within five minutes of CDI addition, 5 mol% DMAP was added to the reaction mixture. This was left to stir for one hour, during which 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 ^{o}C under stirring across 5-10 minutes. This reaction was left to stir for 10-15 min at -78 °C, removed from cooling and stirred at room temperature for 10 minutes, then cooled to -78 °C and stirred for 20 additional minutes before the activated amino acid was added dropwise at -78 °C

across 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. $NaHCO_3$ (30 mL) and sat. NaCl (3x, 30 mL) before being dried over $MgSO_4$. 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. IUPAC Name: *tert*-butyl (*S*)-4-((*tert*-butoxycarbonyl)amino)-3-oxopentanoate (1.44 g, 72% yield). ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.19-5.09 (m, 1H), 4.42-4.32 (m, 1H), 3.49, 3.42 (ABq, J_{AB} = 16, 2H), 1.46 (s, 9H), 1.44 (s, 9H), 1.35 (d, *J* = 7.2, 3H).

Triflate synthesis. To a solution of t-butyl R-lactate or t-butyl 2-hydroxyacetate (5.0 mmol) in dry DCM (20 mL) was added 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 across 70 minutes, during which 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 1M HCl/sat. NaCl (3x, 50 mL), and dried over $MgSO_4$. The extract was concentrated by rotary evaporation and dried under vacuum to afford the product as a red or orange oil which was used without further purification.

IUPAC Name: *tert*-butyl (*R*)-2-(((Trifluoromethyl)sulfonyl)oxy)propanoate (1.01 g, 73% yield). ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.09 (q, *J* = 7.0, 1H), 1.66 (d, *J* = 7.0, 3H), 1.50 (s, 9H). ^{1}H NMR consistent with previously reported spectra.[109]

IUPAC Name: *tert*-butyl 2-(((Trifluoromethyl)sulfonyl)oxy)acetate (0.96 g, 72% yield). ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.78 (s, 2H), 1.51 (s, 9H).

Ketomethylene synthesis. Boc-Ala ketoester (0.50 g, 1.7 mmol) was solvated in dry DCM (10 mL) and added dropwise to a stirred suspension of NaH (60% in mineral oil, 0.21 g, 5.2 mmol) in dry THF (10 mL) at -5 °C. This mixture was allowed to stir for 20 min, after which triflate (1.74 mmol) in dry THF (5 mL) was added at -5 °C. The resulting contents were allowed to stir overnight at room temperature before being quenched with 10% w/v citric acid (20 mL). The quenched reaction was extracted with EtOAc (3x, 20 mL) washed with sat. NaCl (60 mL), and dried over $MgSO_4$ before being concentrated via rotary evaporation to yield a yellow oil. This residue was purified by flash column chromatography with 1:3 EtOAc/hexane and the desired product fractions identified, pooled, and concentrated by rotary evaporation. This residue was solvated in 10% TFA/DCM and allowed to stir overnight at room temperature. After concentrating the resulting mixture by rotary evaporation, the residue was dissolved in DCM and reconcentrated by rotary evaporation (3x) after which the remaining solvent was removed *in vacuo*. Data for protected ketomethylene: IUPAC Name: di-*tert*-butyl 2-((*tert*-butoxycarbonyl)-*L*-alanyl)succinate LC-ESI-MS: *m*/*z* calculated ($M+Na^+$): 424.5, observed: 424.3.

Fmoc protection. The vacuum dried deprotected ketomethylene was dissolved in 1:1 water/MeCN (10 mL) and DIPEA (0.91 mL, 3 eq) was administered to bring the solution to pH ~8. Fmoc-OSu (0.59 g, 1 eq.) was 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 EtOAc (3x, 10 mL), washed with sat. NaCl (30 mL) and dried over $MgSO_4$ before being concentrated under rotary evaporation. The residue was solubilized in 1:1 EtOAc/n-hexane (0.1% AcOH) and purified by flash column chromatography using 1:1 EtOAc/n-hexane (0.1% AcOH) until the product began to elute, at which point 3:1 EtOAc/n-hexane (0.1% AcOH) was used as the solvent. IUPAC Name: (*S*)-5-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-4-oxohexanoic acid. LC-ESI-MS: *m*/*z* calculated ($M+Na^+$) 390.4, observed 390.1.

Inhibition of *In vitro* Transpeptidation Reactions.

Transpeptidation inhibition assays. Reactions for analysis of the preliminary inhibitor substrate analog **1** were prepared from 7 day 37 °C heat-treated **batch 3** $SrtA_{pneu}$ and incubated for 24 hours at room temperature before quenching with N-ethylmaleimide. Reactions were prepared by combining all components shown in **Table 4** except sortase, which was added to initiate the reaction. HPLC analysis was performed using a Dionex Ultimate 3000 HPLC system, which allowed for monitoring by UV/Vis analysis of the separations using a Phenomenex Kinetex 2.6 µm, 100 Å C18 column (3.0 x 100 mm) with the method (Method B): MeCN (0.1% TFA)/H_2O (0.1% TFA) mobile phase. Flow rate = 0.7 mL/min. Gradient = 10% MeCN (0.0-0.5 min), 10% MeCN to 90% MeCN (0.5-6.0 min), hold 90% MeCN (6.0-7.0 min), 90% MeCN to 10% MeCN (7.0-7.1

min), re-equilibrate to 10% MeCN (7.1-10.0 min). Percent conversion was calculated by dividing the area of the product peak by the addition of total 2,4-DNP containing reactant and product peak areas at 330 nm.

Table 3. Reaction conditions for sortase mediated ligation inhibition assays. Volume of components used are given in µL. Water was added to 50 uL total reaction volume unless otherwise noted.

Reaction	1	2	3	4	5	6	7
Buffer (500 mM Tris, pH 7.5, 150 mM NaCl)	5	5	5	5	5	5	5
$SrtA_{pneu}$ (300 µM)	1.7	1.7	1.7	1.7	1.7	1.7	1.7
L-Alanine amide HCl (100 mM)	5	5	5	5	5	5	0
Abz-LPATA-GK(Dnp) (3.7 mM)	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Ac-K(Dnp)LPAG(keto)GAA (1.0 mM)	0	0.5	2.5	5	10	25	0

Attempted detection of Sulfenic Acid Modification

Sulfenic acid labeling reactions. **Batch 3** $SrtA_{pneu}$ in HEPES buffer (20 mM HEPES, 150 mM NaCl) was diluted to 50 µM before NO_2-alkyne was added to 500 µM. For the positive control reaction with H_2O_2, $SrtA_{pneu}$ in HEPES buffer was diluted to 10 µM and H_2O_2 was added to 15 µM before NO_2-alkyne was added to 100 µM. After mixing the solutions, they were allowed to stand for 24 hours at room temperature before analysis by LC-MS Dionex Ultimate 3000 HPLC system (Thermo Scientific) connected to an expression L compact mass spectrometer

(Advion, Inc.) using a Phenomenex Kinetex 2.6 µm, 100 Å C4 column (2.0 x 100 mm) with the method (Method A): MeCN (0.1% Formic acid)/95% H_2O, 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 (0.5-6.0 min), hold 90% MeCN (6.0-7.0 min), 90% MeCN to 10% MeCN (7.0-7.1 min), re-equilibrate to 10% MeCN (7.1-10.0 min). Data analysis was done in Advion Data Express software version 3.0. Mass spectrum deconvolution was achieved through a max entropy goodness of fit algorithm to determine uncharged masses of samples

Table 4. Stock solutions and reaction concentrations of the components used in sulfenic acid labeling reactions for $SrtA_{pneu}$. All unlabeled values are in µL. * denotes the use of a 10 mM stock of NO_2-alkyne probe, prepared from 10:1 dilution of the 50 mM stock with HEPES buffer.

		Reaction		
Stock Solution	Reaction Concentration	1	2	3
HEPES Buffer (1x)	1x	66	66	92.6
Batch 3 $SrtA_{pneu}$ (150 µM)	50 or 10 µM	33	33	6.7
NO_2-alkyne (50 or 10 mM)	500 µM	0.5	-	1*
H_2O_2 (2.0 mM)	15 um	-	-	0.75

5. Literature Cited

[1] Wals, K., and Ovaa, H. (2014) Unnatural amino acid incorporation in E. coli: current and future applications in the design of therapeutic proteins, *Front Chem 2*, 15.

[2] Rashidian, M., Dozier, J. K., and Distefano, M. D. (2013) Enzymatic labeling of proteins: techniques and approaches, *Bioconjug Chem 24*, 1277-1294.

[3] Koniev, O., and Wagner, A. (2015) Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation, *Chem Soc Rev 44*, 5495-5551.

[4] Brannigan, J. A., and Wilkinson, A. J. (2002) Protein engineering 20 years on, *Nat Rev Mol Cell Biol 3*, 964-970.

[5] Lang, K., and Chin, J. W. (2014) Bioorthogonal reactions for labeling proteins, *ACS Chem Biol 9*, 16-20.

[6] Mao, H., Hart, S. A., Schink, A., and Pollok, B. A. (2004) Sortase-mediated protein ligation: a new method for protein engineering, *J Am Chem Soc 126*, 2670-2671.

[7] Beerli, R. R., Hell, T., Merkel, A. S., and Grawunder, U. (2015) Sortase Enzyme-Mediated Generation of Site-Specifically Conjugated Antibody Drug Conjugates with High In Vitro and In Vivo Potency, *PLoS One 10*, e0131177.

[8] Chen, Q., Sun, Q., Molino, N. M., Wang, S. W., Boder, E. T., and Chen, W. (2015) Sortase A-mediated multi-functionalization of protein nanoparticles, *Chem Commun (Camb) 51*, 12107-12110.

[9] Antos, J. M., Chew, G. L., Guimaraes, C. P., Yoder, N. C., Grotenbreg, G. M., Popp, M. W., and Ploegh, H. L. (2009) Site-specific N- and C-terminal labeling of a single polypeptide using sortases of different specificity, *J Am Chem Soc 131*, 10800-10801.

[10] Roeser, D., Preusser-Kunze, A., Schmidt, B., Gasow, K., Wittmann, J. G., Dierks, T., von Figura, K., and Rudolph, M. G. (2006) A general binding mechanism for all human sulfatases by the formylglycine-generating enzyme, *Proc Natl Acad Sci U S A 103*, 81-86.

[11] Rush, J. S., and Bertozzi, C. R. (2008) New aldehyde tag sequences identified by screening formylglycine generating enzymes in vitro and in vivo, *J Am Chem Soc 130*, 12240-12241.

[12] Wu, P., Shui, W., Carlson, B. L., Hu, N., Rabuka, D., Lee, J., and Bertozzi, C. R. (2009) Site-specific chemical modification of recombinant proteins produced in mammalian cells by using the genetically encoded aldehyde tag, *Proc Natl Acad Sci U S A 106*, 3000-3005.

[13] Fernandez-Suarez, M., Baruah, H., Martinez-Hernandez, L., Xie, K. T., Baskin, J. M., Bertozzi, C. R., and Ting, A. Y. (2007) Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes, *Nat Biotechnol 25*, 1483-1487.

[14] Slavoff, S. A., Liu, D. S., Cohen, J. D., and Ting, A. Y. (2011) Imaging protein-protein interactions inside living cells via interaction-dependent fluorophore ligation, *J Am Chem Soc 133*, 19769-19776.

[15] Beckett, D., Kovaleva, E., and Schatz, P. J. (1999) A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation, *Protein Sci 8*, 921-929.

[16] Algar, W. R., Prasuhn, D. E., Stewart, M. H., Jennings, T. L., Blanco-Canosa, J. B., Dawson, P. E., and Medintz, I. L. (2011) The controlled display of biomolecules on nanoparticles: a challenge suited to bioorthogonal chemistry, *Bioconjug Chem 22*, 825-858.

[17] Howarth, M., Takao, K., Hayashi, Y., and Ting, A. Y. (2005) Targeting quantum dots to surface proteins in living cells with biotin ligase, *Proc Natl Acad Sci U S A 102*, 7583-7588.

[18] Ton-That, H., Liu, G., Mazmanian, S. K., Faull, K. F., and Schneewind, O. (1999) Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of Staphylococcus aureus at the LPXTG motif, *Proc Natl Acad Sci U S A 96*, 12424-12429.

[19] Novick, R. P. (2000) Sortase: the surface protein anchoring transpeptidase and the LPXTG motif, *Trends Microbiol 8*, 148-151.

[20] Ilangovan, U., Ton-That, H., Iwahara, J., Schneewind, O., and Clubb, R. T. (2001) Structure of sortase, the transpeptidase that anchors proteins to

the cell wall of Staphylococcus aureus, *Proc Natl Acad Sci U S A 98*, 6056-6061.

[21] Bradshaw, W. J., Davies, A. H., Chambers, C. J., Roberts, A. K., Shone, C. C., and Acharya, K. R. (2015) Molecular features of the sortase enzyme family, *FEBS J 282*, 2097-2114.

[22] Dramsi, S., Trieu-Cuot, P., and Bierne, H. (2005) Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria, *Res Microbiol 156*, 289-297.

[23] Ton-That, H., and Schneewind, O. (1999) Anchor structure of staphylococcal surface proteins. IV. Inhibitors of the cell wall sorting reaction, *J Biol Chem 274*, 24316-24320.

[24] Mazmanian, S. K., Liu, G., Ton-That, H., and Schneewind, O. (1999) Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall, *Science 285*, 760-763.

[25] Mazmanian, S. K., Liu, G., Jensen, E. R., Lenoy, E., and Schneewind, O. (2000) Staphylococcus aureus sortase mutants defective in the display of surface proteins and in the pathogenesis of animal infections, *Proc Natl Acad Sci U S A 97*, 5510-5515.

[26] Mazmanian, S. K., Ton-That, H., and Schneewind, O. (2001) Sortase-catalysed anchoring of surface proteins to the cell wall of Staphylococcus aureus, *Mol Microbiol 40*, 1049-1057.

[27] Marraffini, L. A., Dedent, A. C., and Schneewind, O. (2006) Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria, *Microbiol Mol Biol Rev 70*, 192-221.

[28] Patti, J. M., Jonsson, H., Guss, B., Switalski, L. M., Wiberg, K., Lindberg, M., and Hook, M. (1992) Molecular characterization and expression of a gene encoding a Staphylococcus aureus collagen adhesin, *J Biol Chem 267*, 4766-4772.

[29] Lalioui, L., Pellegrini, E., Dramsi, S., Baptista, M., Bourgeois, N., Doucet-Populaire, F., Rusniok, C., Zouine, M., Glaser, P., Kunst, F., Poyart, C., and Trieu-Cuot, P. (2005) 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 73*, 3342-3350.

[30] Signas, C., Raucci, G., Jonsson, K., Lindgren, P. E., Anantharamaiah, G. M., Hook, M., and Lindberg, M. (1989) 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 U S A 86*, 699-703.

[31] Vanier, G., Sekizaki, T., Domínguez-Punaro, M. C., Esgleas, M., Osaki, M., Takamatsu, D., Segura, M., and Gottschalk, M. (2008) Disruption of srtA gene in Streptococcus suis results in decreased interactions with

endothelial cells and extracellular matrix proteins, *Vet Microbiol 127*, 417-424.

[32] Gomez, M. I., Lee, A., Reddy, B., Muir, A., Soong, G., Pitt, A., Cheung, A., and Prince, A. (2004) Staphylococcus aureus protein A induces airway epithelial inflammatory responses by activating TNFR1, *Nat Med 10*, 842-848.

[33] Cossart, P., and Jonquières, R. (2000) Sortase, a universal target for therapeutic agents against gram-positive bacteria?, *Proc Natl Acad Sci U S A 97*, 5013-5015.

[34] Bierne, H., Mazmanian, S. K., Trost, M., Pucciarelli, M. G., Liu, G., Dehoux, P., Jänsch, L., Garcia-del Portillo, F., Schneewind, O., Cossart, P., and Consortium, E. L. G. (2002) Inactivation of the srtA gene in Listeria monocytogenes inhibits anchoring of surface proteins and affects virulence, *Mol Microbiol 43*, 869-881.

[35] Garandeau, C., Réglier-Poupet, H., Dubail, I., Beretti, J. L., Berche, P., and Charbit, A. (2002) The sortase SrtA of Listeria monocytogenes is involved in processing of internalin and in virulence, *Infect Immun 70*, 1382-1390.

[36] Chen, S., Paterson, G. K., Tong, H. H., Mitchell, T. J., and DeMaria, T. F. (2005) Sortase A contributes to pneumococcal nasopharyngeal colonization in the chinchilla model, *FEMS Microbiol Lett 253*, 151-154.

[37] Paterson, G. K., and Mitchell, T. J. (2006) The role of Streptococcus pneumoniae sortase A in colonisation and pathogenesis, *Microbes Infect 8*, 145-153.

[38] Maresso, A. W., and Schneewind, O. (2008) Sortase as a target of anti-infective therapy, *Pharmacol Rev 60*, 128-141.

[39] Gaspar, A. H., Marraffini, L. A., Glass, E. M., Debord, K. L., Ton-That, H., and Schneewind, O. (2005) Bacillus anthracis sortase A (SrtA) anchors LPXTG motif-containing surface proteins to the cell wall envelope, *J Bacteriol 187*, 4646-4655.

[40] Ton-That, H., Marraffini, L. A., and Schneewind, O. (2004) Protein sorting to the cell wall envelope of Gram-positive bacteria, *Biochim Biophys Acta 1694*, 269-278.

[41] Mazmanian, S. K., Ton-That, H., Su, K., and Schneewind, O. (2002) An iron-regulated sortase anchors a class of surface protein during Staphylococcus aureus pathogenesis, *Proc Natl Acad Sci U S A 99*, 2293-2298.

[42] Ton-That, H., Marraffini, L. A., and Schneewind, O. (2004) Sortases and pilin elements involved in pilus assembly of Corynebacterium diphtheriae, *Mol Microbiol 53*, 251-261.

[43] Ton-That, H., and Schneewind, O. (2004) Assembly of pili in Gram-positive bacteria, *Trends Microbiol 12*, 228-234.

[44] Ton-That, H., and Schneewind, O. (2003) Assembly of pili on the surface of Corynebacterium diphtheriae, *Mol Microbiol 50*, 1429-1438.

[45] Budzik, J. M., Marraffini, L. A., and Schneewind, O. (2007) Assembly of pili on the surface of Bacillus cereus vegetative cells, *Mol Microbiol 66*, 495-510.

[46] Duong, A., Capstick, D. S., Di Berardo, C., Findlay, K. C., Hesketh, A., Hong, H. J., and Elliot, M. A. (2012) Aerial development in Streptomyces coelicolor requires sortase activity, *Mol Microbiol 83*, 992-1005.

[47] Fischetti, V. A., Pancholi, V., and Schneewind, O. (1990) Conservation of a hexapeptide sequence in the anchor region of surface proteins from gram-positive cocci, *Mol Microbiol 4*, 1603-1605.

[48] Schneewind, O., Model, P., and Fischetti, V. A. (1992) Sorting of protein A to the staphylococcal cell wall, *Cell 70*, 267-281.

[49] Theile, C. S., Witte, M. D., Blom, A. E., Kundrat, L., Ploegh, H. L., and Guimaraes, C. P. (2013) Site-specific N-terminal labeling of proteins using sortase-mediated reactions, *Nat Protoc 8*, 1800-1807.

[50] Petrache, A. I., Machin, D. C., Williamson, D. J., Webb, M. E., and Beales, P. A. (2016) Sortase-mediated labelling of lipid nanodiscs for cellular tracing, *Mol Biosyst 12*, 1760-1763.

[51] Witte, M. D., Cragnolini, J. J., Dougan, S. K., Yoder, N. C., Popp, M. W., and Ploegh, H. L. (2012) Preparation of unnatural N-to-N and C-to-C protein fusions, *Proc Natl Acad Sci U S A 109*, 11993-11998.

[52] Tanaka, T., Yamamoto, T., Tsukiji, S., and Nagamune, T. (2008) Site-specific protein modification on living cells catalyzed by Sortase, *Chembiochem 9*, 802-807.

[53] Popp, M. W., Antos, J. M., Grotenbreg, G. M., Spooner, E., and Ploegh, H. L. (2007) Sortagging: a versatile method for protein labeling, *Nat Chem Biol 3*, 707-708.

[54] Proft, T. (2010) Sortase-mediated protein ligation: an emerging biotechnology tool for protein modification and immobilisation, *Biotechnol Lett 32*, 1-10.

[55] Matsumoto, T., Takase, R., Tanaka, T., Fukuda, H., and Kondo, A. (2012) Site-specific protein labeling with amine-containing molecules using Lactobacillus plantarum sortase, *Biotechnol J 7*, 642-648.

[56] Guimaraes, C. P., Witte, M. D., Theile, C. S., Bozkurt, G., Kundrat, L., Blom, A. E., and Ploegh, H. L. (2013) Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions, *Nat Protoc 8*, 1787-1799.

[57] Strijbis, K., Spooner, E., and Ploegh, H. L. (2012) Protein ligation in living cells using sortase, *Traffic 13*, 780-789.

[58] Chen, I., Dorr, B. M., and Liu, D. R. (2011) A general strategy for the evolution of bond-forming enzymes using yeast display, *Proc Natl Acad Sci U S A 108*, 11399-11404.

[59] Piotukh, K., Geltinger, B., Heinrich, N., Gerth, F., Beyermann, M., Freund, C., and Schwarzer, D. (2011) Directed evolution of sortase A mutants with altered substrate selectivity profiles, *J Am Chem Soc 133*, 17536-17539.

[60] Schmohl, L., Bierlmeier, J., Gerth, F., Freund, C., and Schwarzer, D. (2017) Engineering sortase A by screening a second-generation library using phage display, *J Pept Sci*.

[61] Hirakawa, H., Ishikawa, S., and Nagamune, T. (2012) Design of Ca2+-independent Staphylococcus aureus sortase A mutants, *Biotechnol Bioeng 109*, 2955-2961.

[62] Hirakawa, H., Ishikawa, S., and Nagamune, T. (2015) Ca2+ -independent sortase-A exhibits high selective protein ligation activity in the cytoplasm of Escherichia coli, *Biotechnol J 10*, 1487-1492.

[63] Kruger, R. G., Otvos, B., Frankel, B. A., Bentley, M., Dostal, P., and McCafferty, D. G. (2004) Analysis of the substrate specificity of the Staphylococcus aureus sortase transpeptidase SrtA, *Biochemistry 43*, 1541-1551.

[64] Perry, A. M., Ton-That, H., Mazmanian, S. K., and Schneewind, O. (2002) 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 277*, 16241-16248.

[65] Huang, X., Aulabaugh, A., Ding, W., Kapoor, B., Alksne, L., Tabei, K., and Ellestad, G. (2003) Kinetic mechanism of Staphylococcus aureus sortase SrtA, *Biochemistry 42*, 11307-11315.

[66] David Row, R., Roark, T. J., Philip, M. C., Perkins, L. L., and Antos, J. M. (2015) Enhancing the efficiency of sortase-mediated ligations through nickel-peptide complex formation, *Chem Commun (Camb) 51*, 12548-12551.

[67] Williamson, D. J., Webb, M. E., and Turnbull, W. B. (2014) Depsipeptide substrates for sortase-mediated N-terminal protein ligation, *Nat Protoc 9*, 253-262.

[68] Boekhorst, J., de Been, M. W., Kleerebezem, M., and Siezen, R. J. (2005) Genome-wide detection and analysis of cell wall-bound proteins with LPxTG-like sorting motifs, *J Bacteriol 187*, 4928-4934.

[69] Kruger, R. G., Dostal, P., and McCafferty, D. G. (2004) Development of a high-performance liquid chromatography assay and revision of kinetic parameters for the Staphylococcus aureus sortase transpeptidase SrtA, *Anal Biochem 326*, 42-48.

[70] Naik, M. T., Suree, N., Ilangovan, U., Liew, C. K., Thieu, W., Campbell, D. O., Clemens, J. J., Jung, M. E., and Clubb, R. T. (2006) Staphylococcus aureus Sortase A transpeptidase. Calcium promotes sorting signal binding by altering the mobility and structure of an active site loop, *J Biol Chem 281*, 1817-1826.

[71] Suree, N., Liew, C. K., Villareal, V. A., Thieu, W., Fadeev, E. A., Clemens, J. J., Jung, M. E., and Clubb, R. T. (2009) The structure of the Staphylococcus aureus sortase-substrate complex reveals how the universally conserved LPXTG sorting signal is recognized, *J Biol Chem 284*, 24465-24477.

[72] Chan, A. H., Yi, S. W., Terwilliger, A. L., Maresso, A. W., Jung, M. E., and Clubb, R. T. (2015) Structure of the Bacillus anthracis Sortase A Enzyme Bound to Its Sorting Signal: A FLEXIBLE AMINO-TERMINAL APPENDAGE MODULATES SUBSTRATE ACCESS, *J Biol Chem 290*, 25461-25474.

[73] Frankel, B. A., Kruger, R. G., Robinson, D. E., Kelleher, N. L., and McCafferty, D. G. (2005) Staphylococcus aureus sortase transpeptidase SrtA: insight into the kinetic mechanism and evidence for a reverse protonation catalytic mechanism, *Biochemistry 44*, 11188-11200.

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

[75] 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., and Banfield, M. J. (2009) Crystal structure of Streptococcus pyogenes sortase A: implications for sortase mechanism, *J Biol Chem 284*, 6924-6933.

[76] Khare, B., Krishnan, V., Rajashankar, K. R., I-Hsiu, H., Xin, M., Ton-That, H., and Narayana, S. V. (2011) Structural differences between the Streptococcus agalactiae housekeeping and pilus-specific sortases: SrtA and SrtC1, *PLoS One 6*, e22995.

[77] Chambers, C. J., Roberts, A. K., Shone, C. C., and Acharya, K. R. (2015) Structure and function of a Clostridium difficile sortase enzyme, *Sci Rep 5*, 9449.

[78] Ton-That, H., Mazmanian, S. K., Alksne, L., and Schneewind, O. (2002) Anchoring of surface proteins to the cell wall of Staphylococcus aureus. Cysteine 184 and histidine 120 of sortase form a thiolate-imidazolium ion pair for catalysis, *J Biol Chem 277*, 7447-7452.

[79] Marraffini, L. A., Ton-That, H., Zong, Y., Narayana, S. V., and Schneewind, O. (2004) 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 279*, 37763-37770.

[80] Kappel, K., Wereszczynski, J., Clubb, R. T., and McCammon, J. A. (2012) The binding mechanism, multiple binding modes, and allosteric regulation of Staphylococcus aureus Sortase A probed by molecular dynamics simulations, *Protein Sci 21*, 1858-1871.

[81] Moritsugu, K., Terada, T., and Kidera, A. (2012) Disorder-to-order transition of an intrinsically disordered region of sortase revealed by multiscale enhanced sampling, *J Am Chem Soc 134*, 7094-7101.

[82] Pang, X., and Zhou, H. X. (2015) Disorder-to-Order Transition of an Active-Site Loop Mediates the Allosteric Activation of Sortase A, *Biophys J 109*, 1706-1715.

[83] 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., and Clubb, R. T. (2014) Structural and computational studies of the Staphylococcus aureus sortase B-substrate complex reveal a substrate-stabilized oxyanion hole, *J Biol Chem 289*, 8891-8902.

[84] Neiers, F., Madhurantakam, C., Fälker, S., Manzano, C., Dessen, A., Normark, S., Henriques-Normark, B., and Achour, A. (2009) Two crystal structures of pneumococcal pilus sortase C provide novel insights into catalysis and substrate specificity, *J Mol Biol 393*, 704-716.

[85] Persson, K. (2011) Structure of the sortase AcSrtC-1 from Actinomyces oris, *Acta Crystallogr D Biol Crystallogr 67*, 212-217.

[86] Cozzi, R., Malito, E., Nuccitelli, A., D'Onofrio, M., Martinelli, M., Ferlenghi, I., Grandi, G., Telford, J. L., Maione, D., and Rinaudo, C. D. (2011) Structure analysis and site-directed mutagenesis of defined key residues and motives for pilus-related sortase C1 in group B Streptococcus, *FASEB J 25*, 1874-1886.

[87] Chen, J., Dong, H., Murfin, K. E., Feng, C., Wu, S., and Zheng, B. (2016) Active site analysis of sortase A from Staphylococcus simulans indicates

function in cleavage of putative cell wall proteins, *Biochem Biophys Res Commun 478*, 1653-1659.

[88] Chen, F., Xie, F., Yang, B., Wang, C., Liu, S., and Zhang, Y. (2017) Streptococcus suis sortase A is Ca2+ independent and is inhibited by acteoside, isoquercitrin and baicalin, *PLoS One 12*, e0173767.

[89] Das, S., Pawale, V. S., Dadireddy, V., Singh, A. K., Ramakumar, S., and Roy, R. P. (2017) Structure and specificity of a new class of Ca(2+)-independent housekeeping sortase from Streptomyces avermitilis provide insights into its non-canonical substrate preference, *J Biol Chem 292*, 7244-7257.

[90] Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment, *J Mol Biol 302*, 205-217.

[91] Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J. M., Taly, J. F., and Notredame, C. (2011) T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension, *Nucleic Acids Res 39*, W13-17.

[92] Lu, C., Zhu, J., Wang, Y., Umeda, A., Cowmeadow, R. B., Lai, E., Moreno, G. N., Person, M. D., and Zhang, Z. (2007) Staphylococcus aureus sortase A exists as a dimeric protein in vitro, *Biochemistry 46*, 9346-9354.

[93] Zhu, J., Lu, C., Standland, M., Lai, E., Moreno, G. N., Umeda, A., Jia, X., and Zhang, Z. (2008) Single mutation on the surface of Staphylococcus aureus Sortase A can disrupt its dimerization, *Biochemistry 47*, 1667-1674.

[94] Zhu, J., Xiang, L., Jiang, F., and Zhang, Z. J. (2016) Equilibrium of sortase A dimerization on Staphylococcus aureus cell surface mediates its cell wall sorting activity, *Exp Biol Med (Maywood) 241*, 90-100.

[95] Dekker, N., Tommassen, J., Lustig, A., Rosenbusch, J. P., and Verheij, H. M. (1997) Dimerization regulates the enzymatic activity of Escherichia coli outer membrane phospholipase A, *J Biol Chem 272*, 3179-3184.

[96] Zong, Y., Bice, T. W., Ton-That, H., Schneewind, O., and Narayana, S. V. (2004) Crystal structures of Staphylococcus aureus sortase A and its substrate complex, *J Biol Chem 279*, 31383-31389.

[97] Gallivan, J. P., and Dougherty, D. A. (1999) Cation-pi interactions in structural biology, *Proc Natl Acad Sci U S A 96*, 9459-9464.

[98] Axelrod, H. L., Abresch, E. C., Okamura, M. Y., Yeh, A. P., Rees, D. C., and Feher, G. (2002) X-ray structure determination of the cytochrome c2: reaction center electron transfer complex from Rhodobacter sphaeroides, *J Mol Biol 319*, 501-515.

[99] Crowley, P. B., and Golovin, A. (2005) Cation-pi interactions in protein-protein interfaces, *Proteins 59*, 231-239.

[100] Melvin, J. A., Murphy, C. F., Dubois, L. G., Thompson, J. W., Moseley, M. A., and McCafferty, D. G. (2011) Staphylococcus aureus sortase A contributes to the Trojan horse mechanism of immune defense evasion with its intrinsic resistance to Cys184 oxidation, *Biochemistry 50*, 7591-7599.

[101] Fidai, I., Hocharoen, L., Bradford, S., Wachnowsky, C., and Cowan, J. A. (2014) Inactivation of sortase A mediated by metal ATCUN complexes, *J Biol Inorg Chem 19*, 1327-1339.

[102] Gupta, V., Paritala, H., and Carroll, K. S. (2016) Reactivity, Selectivity, and Stability in Sulfenic Acid Detection: A Comparative Study of Nucleophilic and Electrophilic Probes, *Bioconjug Chem 27*, 1411-1418.

[103] Gupta, V., and Carroll, K. S. (2016) Rational design of reversible and irreversible cysteine sulfenic acid-targeted linear C-nucleophiles, *Chem Commun (Camb) 52*, 3414-3417.

[104] Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping: a mechanism for oligomer assembly, *Protein Sci 4*, 2455-2468.

[105] Liu, Y., Hart, P. J., Schlunegger, M. P., and Eisenberg, D. (1998) The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-A resolution, *Proc Natl Acad Sci U S A 95*, 3437-3442.

[106] Zegers, I., Deswarte, J., and Wyns, L. (1999) Trimeric domain-swapped barnase, *Proc Natl Acad Sci U S A 96*, 818-822.

[107] Nenci, A., Gotte, G., Bertoldi, M., and Libonati, M. (2001) Structural properties of trimers and tetramers of ribonuclease A, *Protein Sci 10*, 2017-2027.

[108] Gotte, G., Laurents, D. V., and Libonati, M. (2006) Three-dimensional domain-swapped oligomers of ribonuclease A: identification of a fifth tetramer, pentamers and hexamers, and detection of trace heptameric, octameric and nonameric species, *Biochim Biophys Acta 1764*, 44-54.

[109] Budnjo, A., Narawane, S., Grauffel, C., Schillinger, A. S., Fossen, T., Reuter, N., and Haug, B. E. (2014) Reversible ketomethylene-based inhibitors of human neutrophil proteinase 3, *J Med Chem 57*, 9396-9408.

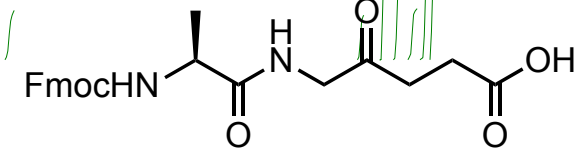
[110] Mathieu, L. B., C.; Masurier, N.; Maillard, L.; Martinez, J.; and Lisowski, V. (2015) Cross-Claisen Condensation of N-Fmoc-Amino Acids – A Short Route to Heterocyclic γ-Amino Acids, *Eur. J. Org. Chem. 2015*, 2262-2270.

[111] Rogers, L. M.-A., McGivern, P. G., Butler, A. R., MacRobert, A. J., and Eggleston, I. M. (2005) An efficient synthesis of 5-aminolevulinic acid (ALA)-containing peptides for use in photodynamic therapy, *Tetrahedron 61*, 6951-6958.

[112] Dixon, M. J., Bourre, L., MacRobert, A., and Eggleston, I. M. (2007) Novel prodrug approach to photodynamic therapy: Fmoc solid-phase synthesis of a cell permeable peptide incorporating 5-aminolevulinic acid, *Bioorg. Med. Chem. Lett 17*, 4518-4522.

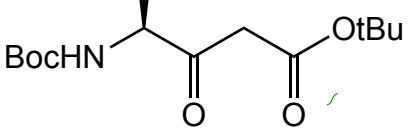
6. Appendix I: NMR Spectra of Synthesized Compounds

(*S*)-5-(2-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)propanamido)-4 oxopentanoic acid

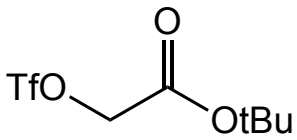


Crude product (residual solvent contamination)

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

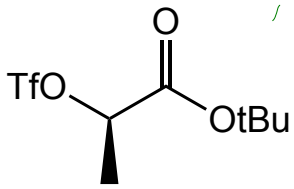


tert-butyl 2-(((Trifluoromethyl)sulfonyl)oxy)acetate



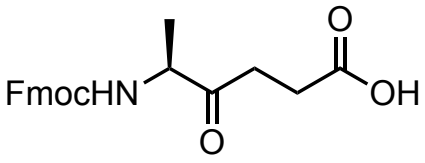
Crude product

tert-butyl (*R*)-2-(((Trifluoromethyl)sulfonyl)oxy)propanoate



Crude product

(*S*)-5-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-4-oxohexanoic acid



Contains solvent and residual starting material