Spring 2020

# Sorting Out Sortases: Designing a Better Enzyme and Synthesizing Trapping Peptides Towards Capturing a Bound-State Structure

Katherine Johnston
*Western Washington University*

# Sorting Out Sortases: Designing a Better Enzyme and Synthesizing Trapping Peptides Towards Capturing a Bound-State Structure

By

Katherine Johnston

June 2020

A Thesis Presented to

The Honors Department of

Western Washington University

In Partial Fulfillment

Of the Requirements for

University Honors

## Abstract

Sortase A is a powerful protein engineering tool that cleaves proteins and attaches them to an acyl acceptor of choice. However, the most active wild type variants of sortase A only show high activity at a limited number of cleavage motifs, and so work is underway to create a variant of sortase A that shows high activity at a greater variety of cleavage sites. Additionally, current studies attempting to optimize the enzyme require a way to stabilize an unstructured loop for the crystallization, in order to collect X-ray diffraction structural data. Here, preliminary results from the design of an “loop-swapped” sortase A enzyme variants are discussed, as well as the synthesis of two different thiol-containing “trapping” peptides intended to freeze sortase A in its bound conformation via a covalent disulfide bond. The first peptide is a previously described unnatural peptide, and was discontinued before completion, and the second peptide is a natural peptide which was completed and purified. Future work will involve the co-crystallization of this peptide with “loop-swapped” sortase A variants.

## Table of Contents

## List of Figures

## Introduction

Sortase A (SrtA) is a transpeptidase enzyme found in gram-positive bacteria which cleaves proteins at the LPXTG motif and ligates the C-terminal of the cleaved protein onto an acyl acceptor, lipid II (Figure 1).[1,2] Other acyl acceptors such as the N-terminus of another protein or peptide are also able to accept the cleaved protein.[3] This function has been successfully exploited by researchers as a protein engineering tool to modify protein sequences post-translationally, attach protein tags to cells, create cyclic structures, and more. *Staphylococcus aureus* sortase A ($SrtA_{staph}$) represents the most commonly used sortase A, in particular due to its high cleavage activity. Despite its many uses, $SrtA_{staph}$ is limited by its specificity to the LPXTG cleavage motif. The *Streptococcus pneumoniae* sortase A ($SrtA_{strep}$) recognizes more cleavage motifs, including LPATS and LPATA[3], however is less active than the *S. aureus* version, with approximately 1/3 cleavage activity when compared to $SrtA_{staph}$ (Figure 2).

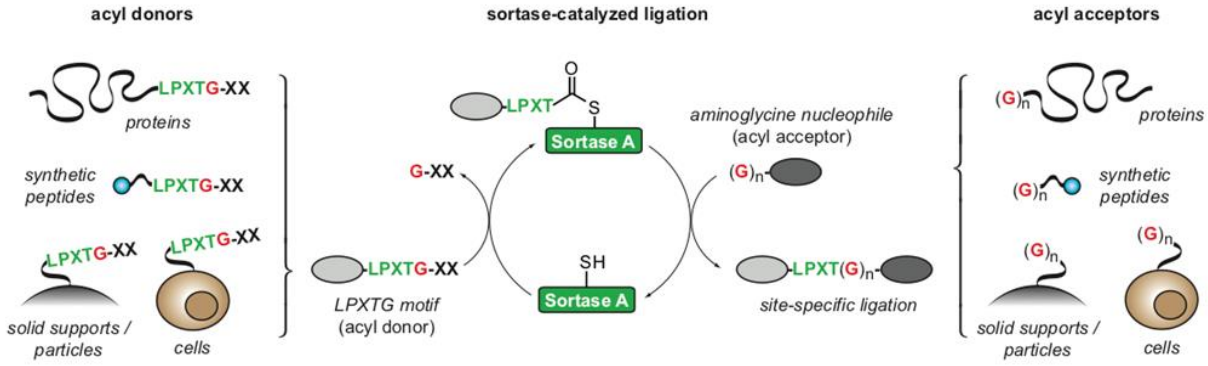



**Figure 1.** Graphical representation of sortase A catalyzed protein ligation. Figure obtained from Antos, *et al.*, (2016).

Designing a sortase variant which has the promiscuity (recognition of additional cleavage sites) of $SrtA_{strep}$ while maintaining the high activity of $SrtA_{staph}$ will be beneficial to the research community as a more flexible protein engineering tool. The activity and promiscuity of new sortase variants can be tested against peptide variants of the LPXTG motif, via fluorescence

2assay (Figure 2). This improvement of Sortase A is an ongoing endeavor in the Amacher and Antos labs, and some parts of the project will be discussed in this paper.

| | G | S | A |
|---|---|---|---|
| S. pneumoniae | 35 | 31 | 24 |
| S. aureus | >95 | 1 | 0 |
| $SPS_{Aureus}$ | >95 | 4 | 17 |
| $SPS_{Suis}$ | 94 | 69 | 83 |
| $SPS_{Oralis}$ | 41 | 21 | 38 |
| $SPS_{Faecalis}$ | >95 | 94 | 75 |
| $SPS_{Lactis}$ | 93 | 52 | 73 |
| $SPS_{Anthracis}$ | 3 | 1 | 0 |
| $SPS_{Monocytogenes}$ | 10 | 1 | 1 |

**Figure 2.** Heat map of relative cleavage by sortase A variants as measured by fluorescent activity. The single letter labels represent the amino acid in the X position of the fluorescent peptide Abz-LPATXG-K(DNP). Figure created by Izzi Piper.

Additionally, structural characterization of Sortase A while it is in its bound conformation will provide insight into the role of the unstructured β-7/β-8 loop which appears to play a role in substrate specificity and cleavage activity. The catalytic residues for $SrtA_{strep}$ are C184, H118, and R192. The cysteine, histidine, and arginine catalytic triad is conserved in SrtA enzymes of other species. Our hypothesis is that a disulfide bond between the catalytic cysteine and the LPXTG peptide involved in crystallization will facilitate the collection of accurate bound-state structural data by both trapping the enzyme in its bound conformation and preventing peptide shifting. A previous structure of *S. aureus* SrtA, which lacked such a covalent bond, revealed a shifted LPETG peptide out of the active site, while a structure that did contain a covalent bond between its peptide and the active site of sortase did not have this issue (Figure 3A, B).[4,5]

The structure PDB ID: 2KID, which is the current best structure of a bound-state sortase A enzyme, includes the unnatural peptide LPAT*, where T* is an unnatural, thiol-containing threonine analogue. The authors of Jung, *et al.*, 2005 developed the LPAT* peptide for the purpose of forming a disulfide bond between sortase and a peptide ligand.[1] They showed that a

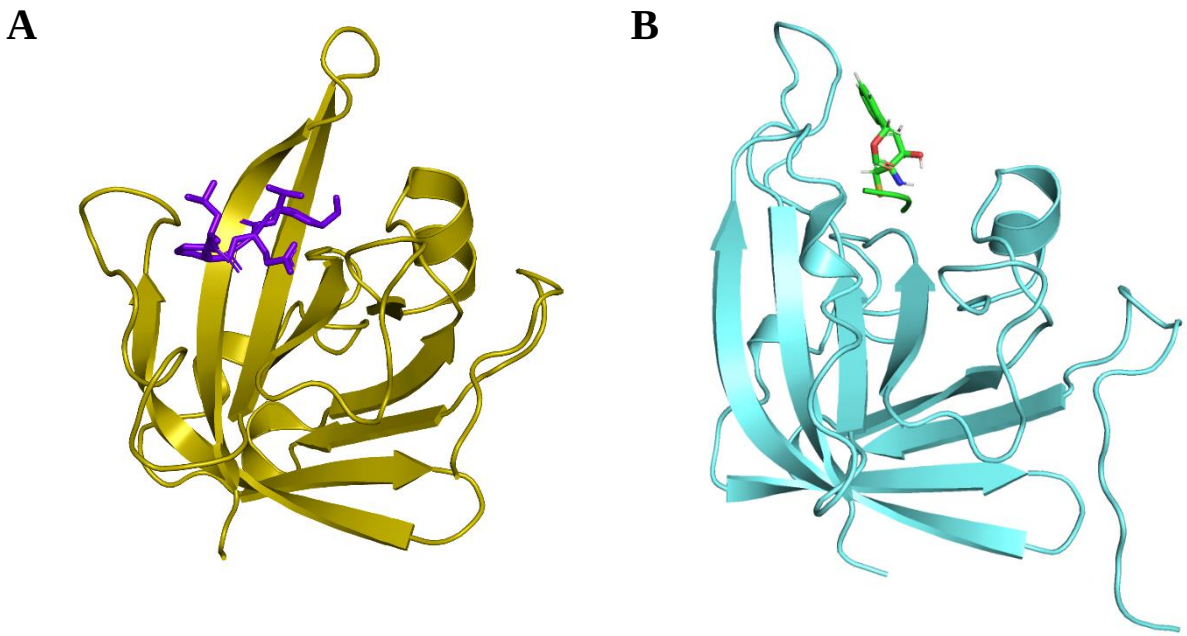



**Figure 3. A.** X-ray crystallography structure of $SrtA_{staph}$ complexed with shifted LPETG peptide, PDB 1T2W. **B.** NMR structure of $SrtA_{staph}$ complexed with (PHQ)LPA(B27) peptide, PDB 2KID.

disulfide bond did indeed form between the two when the unnatural peptide was added in excess.[1] The goal of the disulfide bond was to allow for the collection of structural data on the bound enzyme. They had previously synthesized other versions of this threonine analogue towards the same goal, and found that the resulting bound enzyme was not optimal for structural studies.[1] While their original synthesis of T* began with L-threonine and required nine synthetic steps, involving the use of highly toxic hydrogen sulfide gas ($H_2S$), the authors designed a new route in 2012 that was both shorter, requiring only four synthetic steps, and did not require the use of $H_2S$.[1,6] I initially set out to recreate this unnatural peptide following the 2012 synthetic route, with the ultimate goal of co-crystallizing it with our "loop-swapped" sortase variants in order to solve X-ray crystallography structures of the bound enzymes. When the synthesis proved to be unreasonable to complete with the resources available, I created a similar natural peptide to use towards the same goal.

## Activity of the β-7/β-8 Loop

Previous computational work identified the β-7/β-8 unstructured loop near the active site of sortase A as quite variable amongst SrtA enzymes, therefore, we hypothesized that it may affect substrate recognition. The loop is structurally conserved across class A sortases, but varies greatly in sequence, and sometimes in length. Based on this data, several "loop-swapped" sortase variants were designed with the body of $SrtA_{strep}$ and the β-7/β-8 loop from other sortase A enzymes (Figure 4). The recombinant proteins were expressed using standard protein expression protocols (Methods).

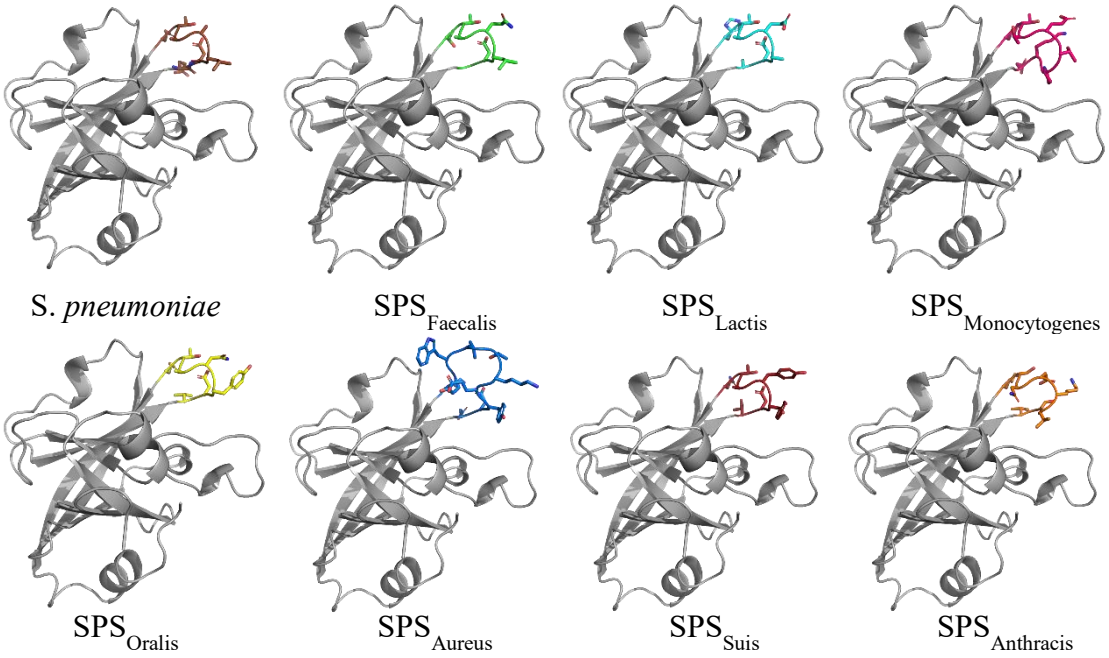



**Figure 4.** "Loop Swapped" Sortase A variants with β-7/β-8 Loop in Stick Configuration. As an example of the naming format, $SPS_{Faecalis}$ stands for '*S. pneumoniae* swap *faecalis*', and indicates that the mutant has the body of *S. pneumoniae* sortase A and the β-7/β-8 loop of *S. faecalis* sortase A. Figure created by Alex Johnson.

Fluorescent peptides were used to mimic the cleavage recognition site of sortase A. Each peptide followed the sequence Abz-LPAT'XGK(DNP), where Abz is aminobenzoic acid, X represents any amino acid, and DNP is the chromophore 2,4-dinitrophenol (Figure 5). Abz fluoresces at 420 nm, and DNP quenches its fluorescence when in close proximity. The peptides

were subjected to excitation at 320 nm while in solution with active sortase A variants. Due to the change in proximity of DNP to Abz during cleavage, fluorescence was quenched in whole peptides but not in cleaved peptides, and fluorescence signal increased with cleavage activity (Figure 2).

Our fluorescence data revealed that the cleavage activity and promiscuity of sortase varied greatly in our different loop-swapped variants, despite all containing the $SrtA_{strep}$ scaffold. The activity of the "loop-swapped" variants appears to follow that of the sortase from which the β-7/β-8 loop came, rather than mimicking the activity of the $SrtA_{strep}$ base. Multiple "loop-swapped" variants showed the high levels of activity at multiple recognition sites, indicating successful design of a more active and promiscuous enzyme (Figure 2). Work in this area is ongoing and these results are preliminary.

## Fluorescent Peptides

Chromophore-containing peptides for use in the fluorescent sortase A activity assay were synthesized via solid phase peptide synthesis (SPPS) (Methods). The Abz-LPATAGK(DNP) peptide was successfully completed and purified (Figure 5A). Due to the presence of the DNP chromophore, this peptide appears faintly yellow in otherwise transparent solution and light yellow in powdered form. The mass of this peptide was calculated to be 943.33 Da. Purification was not performed on the entire crude sample, due to time constraints and the higher priority of the crystallization peptide. The purified peptide and remaining crude peptide were not quantified after lyophilization, but the purified peptide was a notably large amount.

The Abz-LPATGGK(DNP) peptide, which mimics sortase A's natural cleavage motif without any altered residues, was completed (Figure 5B). A large fraction of the resin was lost through the filter during the TFA cleavage step of synthesis, meaning that the purified peptide yield will be decreased. However, it was not purified before the SARS-CoV-2 pandemic halted lab work, and its identity has not been confirmed via MS. Its mass is calculated to be 929.42 Da.

The Abz-LPATSGK(DNP) peptide was completed (Figure 5C). However, it was not purified before the SARS-CoV-2 pandemic halted lab work, and its identity has not been confirmed via MS. Its mass is calculated to be 959.43 Da.

Several times during purification set up, the HPLC showed signs of spiking pressure. This was troubleshooted by reducing the flow rate, and increasing the flow rate more slowly. In cases where bubbles entered the line, the system was purged.

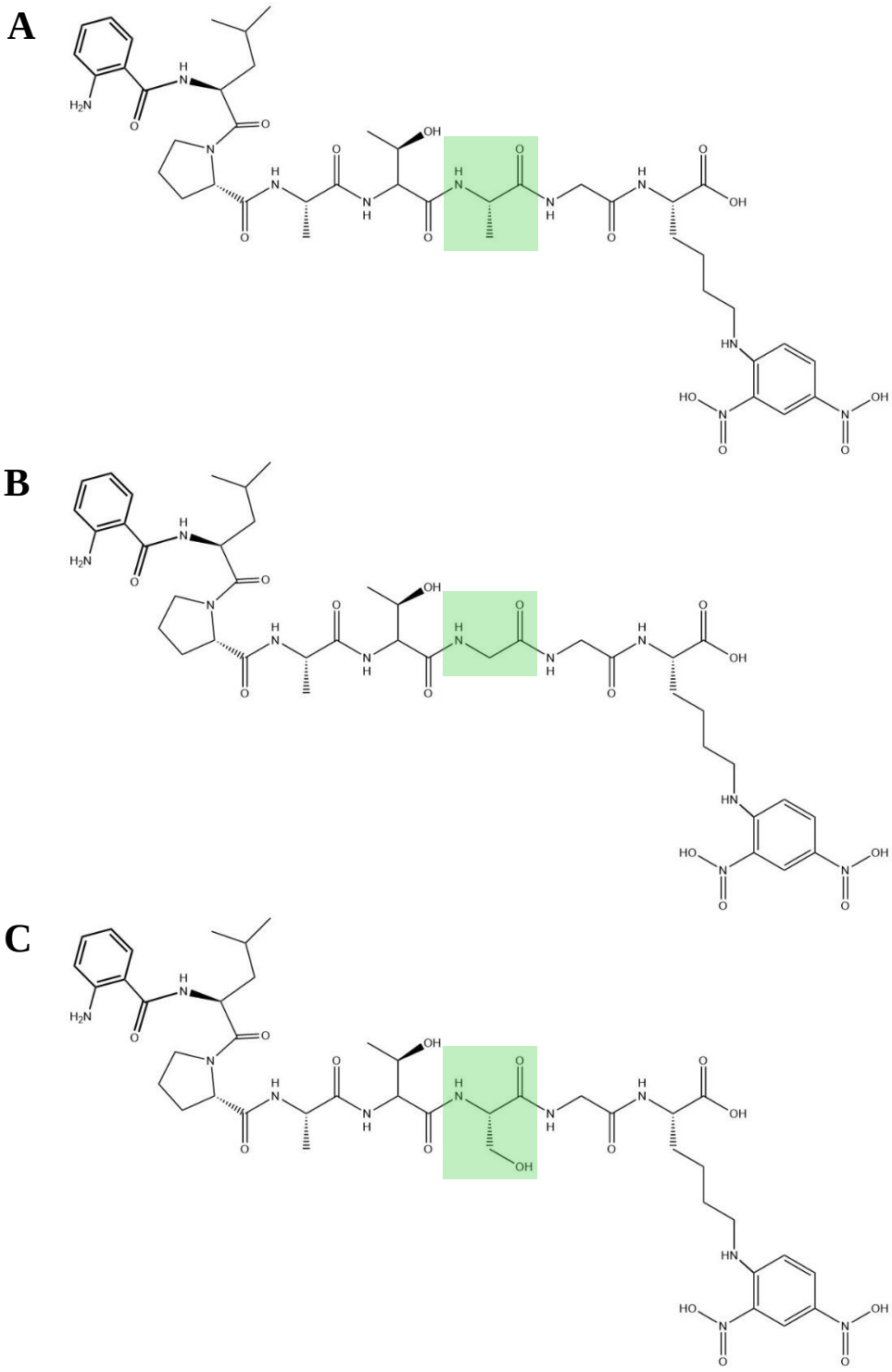


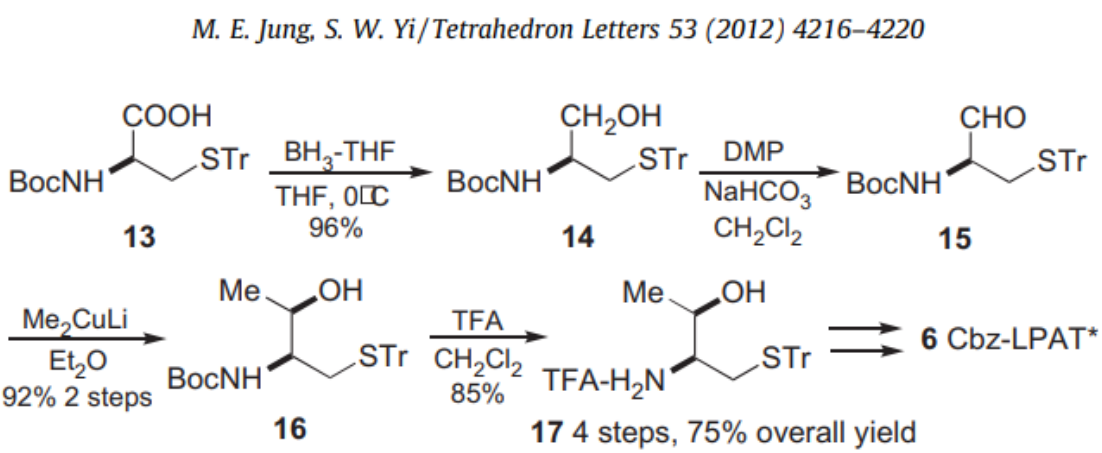

**Figure 5.** Fluorescent peptides used to mimic the sortase A cleavage motif. Differences between the peptides are highlighted in green. **A.** The structure of Abz-LPATAGK(DNP). **B.** The structure of Abz-LPATGGK(DNP). **C.** The structure of Abz-LPATSGK(DNP).

## Synthesis of a Trapping Peptide: Two Routes

### Unnatural Peptide

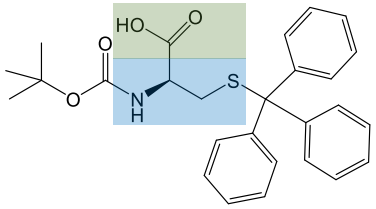
Following the successful results of the "loop-swapped" variants, bound-state structural data became important to further understand the interaction of the β-7/β-8 loop with the peptide. Therefore, synthesis of the unnatural trapping peptide LPAT* was attempted following the improved 2012 synthesis scheme developed by Jung and Yi (Figure 6A).[6]

**A**




**Scheme 1.** Improved method for synthesis of **6**.

**B**




**C**

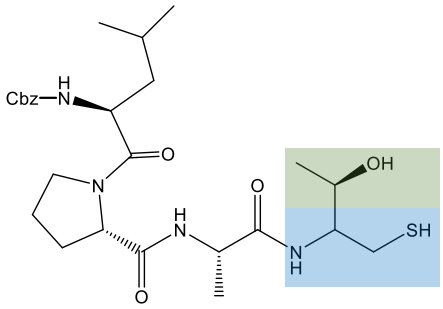



**Figure 6. A.** Synthetic route for unnatural peptide LPAT*. Figure obtained from Jung and Yi, (2012). **B.** Structure of the bis-protected cysteine starting material (compound 13). **C.** Structure of the unnatural peptide product (compound 6). The blue rectangle highlights the part of the starting material conserved, and the green rectangle highlights the reduction of the starting material's carboxylic acid.

The first synthetic step, reduction of the cysteine's C-terminal carboxylic acid, was attempted three times. In the first attempt, borane tetrahydrofuran complex (referred to as 'borane') was allowed to react with the bis-protected cysteine for 80 minutes total, as described in Methods.

This was shorter than the four total hours (1 hour at 0 °C, 3 hours at room temperature) suggested by Jung and Yi in their 2012 paper.[6] The resulting crude product was purified via flash column chromatography. During purification, thin layer chromatography (TLC) monitoring showed that the cysteine eluted off the silica gel column in fractions 4 through 7. However, the TLC did not show a difference between the starting material control sample and the crude product control sample, and no product was able to be isolated when fractions 4 through 7 were concentrated. It was determined that the product had been lost.

In the second attempt, the procedure used for synthesis in the first attempt was repeated with a smaller molar excess of borane, and the crude product was not purified. To determine whether the reduction reaction had been successful, NMR spectroscopy was performed on both the crude product and a sample of the starting material. No peak for a carboxylic acid (10 to 12 ppm) was observed in either spectra, which was unexpected (Figure 7A, B). NMR analysis of the crude product showed no presence of primary alcohol, indicating that the reduction of the carboxylic acid failed (Figure 7B). It was then concluded that the first attempt had also failed, and that the product had not been lost as it was nonexistent. In consultation with mentors, it was determined that in both cases, the borane had not been allowed to react with the bis-protected cysteine for long enough.

Following this conclusion, the procedure was repeated and the reaction was allowed to proceed for one hour at 0 °C and overnight (approximately 17 hours) at room temperature. NMR spectroscopy was performed on both the crude product and on a new sample of the starting material, and again no peak for a carboxylic acid (10 to 12 ppm) was observed in either the starting material or the crude product NMR spectra (Figure 8A, B). The continued absence of the carboxylic acid peak is unusual and requires further investigation. The NMR spectra of the crude

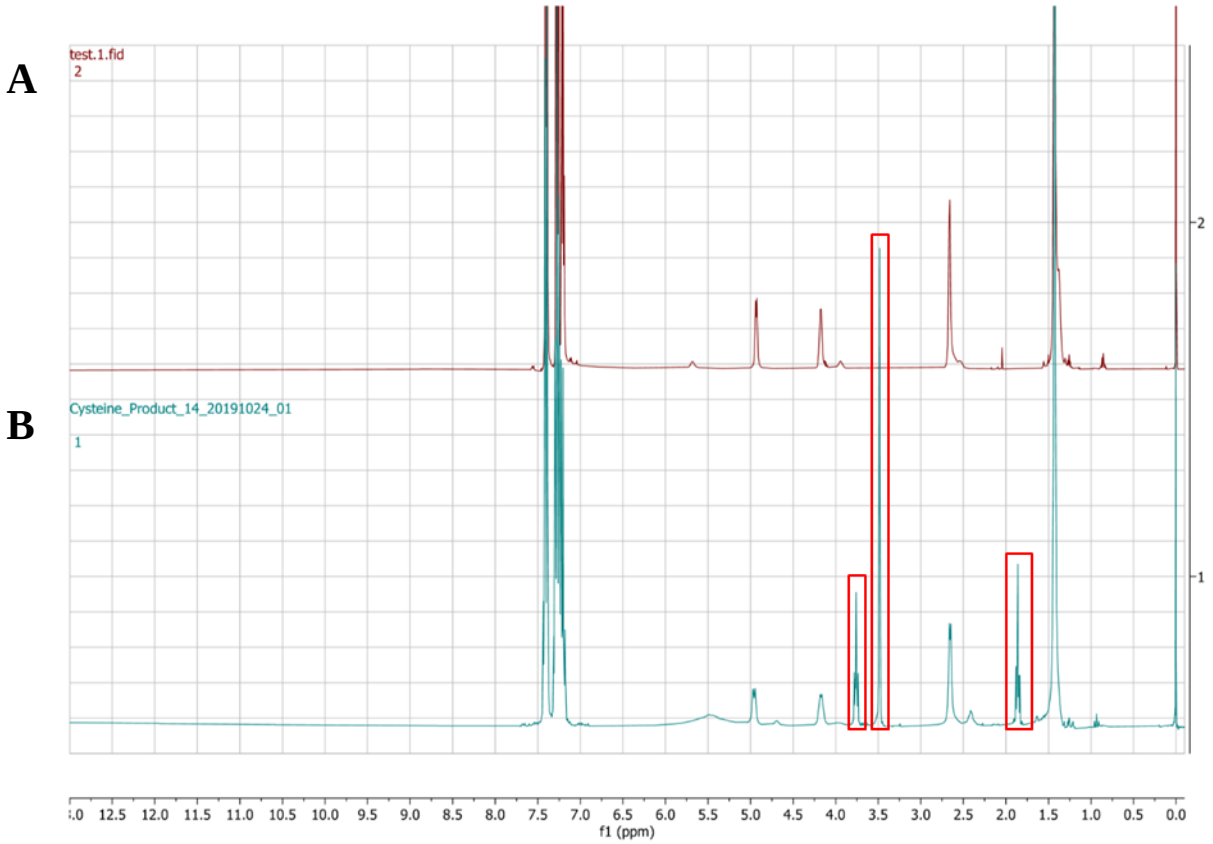



**Figure 7. A.** NMR spectrum of bis-protected cysteine starting material, taken at 500 MHz. **B.** NMR spectrum of the crude product from synthetic step 1, attempt 2. This spectrum was taken at 300 MHz on the same day as Figure 7A. Peaks highlighted in red are not present on the starting material spectrum and are consistent with known chemical shifts of solvent peaks for MeOH (3.49) and THF (1.85, 3.76).[11]

product showed some presence of primary alcohol at 3.41-3.44 ppm and 3.62-3.69 ppm, indicating that the reduction was at least partially successful (Figure 8B). However, the starting material was clearly still present in the crude product, indicating that conditions for the reduction need to be improved in order to increase the reaction efficiency.

It is clear that the borane reduction reaction in synthetic step 1 is not optimized for maximum product yield. In order to understand and improve on these limitations, further investigation is necessary. Increased length of reaction time did allow the reaction to be successful; however, 17 hours is a very long reaction time and other steps, such as altering the ratio of reaction time at 0 °C versus at room temperature, altering the "room temperature", altering the molar ratio of bis-

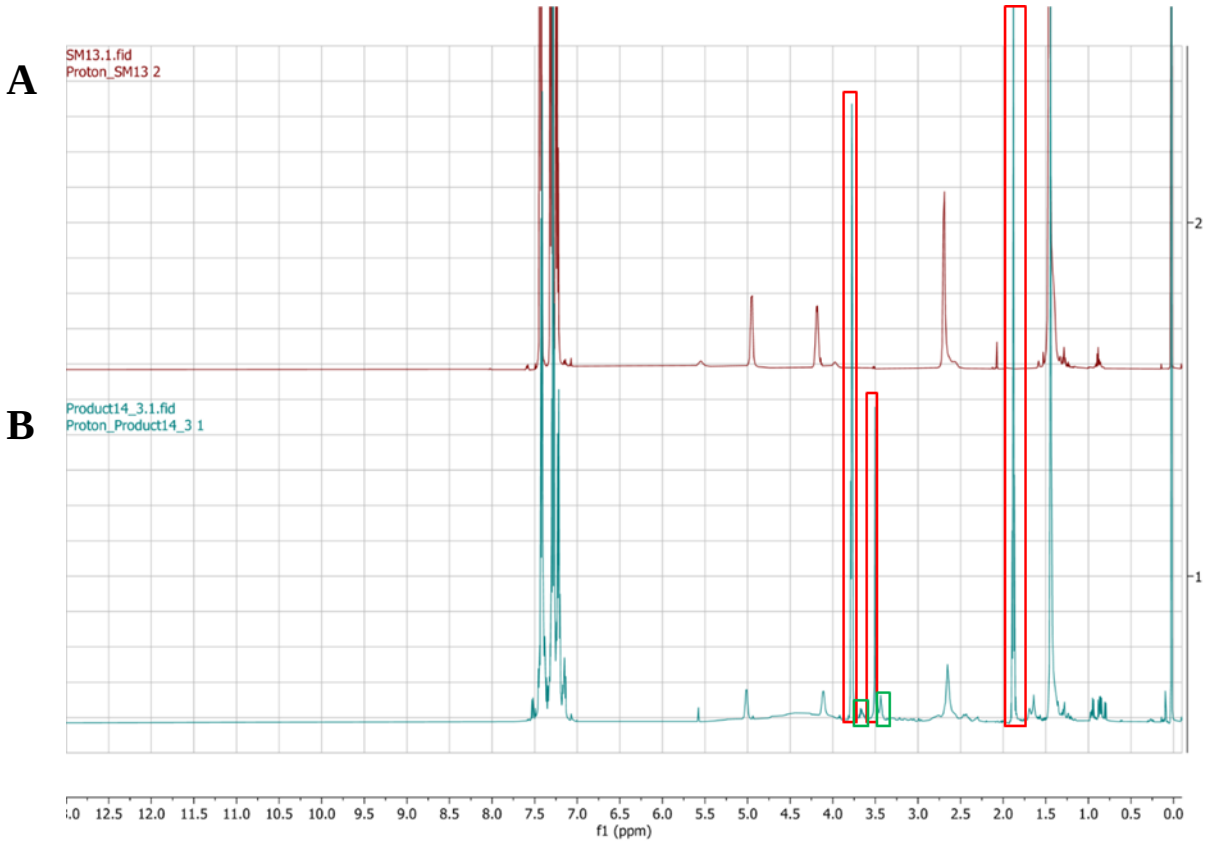



**Figure 8. A.** NMR spectrum of bis-protected cysteine starting material, taken at 500 MHz. **B.** NMR spectrum of the crude product from synthetic step 1, attempt 3. This spectrum was taken at 500 MHz on the same day as Figure 8A. Peaks highlighted in red are not present on the starting material spectrum and are consistent with known chemical shifts of solvent peaks for MeOH (3.49) and THF (1.85, 3.76).[11] Peaks highlighted in green are not present on the starting material spectrum and are not consistent with solvent peaks. These green peaks are at chemical shifts (3.41-3.44 and 3.62-3.69), which are consistent with expected peaks from a primary alcohol on a bis-protected cysteine.

protected cysteine to borane complex, and altering reaction vessel volume, all may lead to better yields. Further investigation is required to determine how changing these factors affects the reaction. Additionally, investigation is required into the lack of signal for the carboxylic acid across all spectra taken. One possible route is confirmation of the bis-protected cysteine's formula via mass spectrometry. The synthesis of an unnatural cysteine-containing trapping peptide was discontinued after the third attempt at synthetic step number one, due to time and budgetary constraints.

## Natural Peptide

Following termination of synthesis of the unnatural peptide, synthesis of a natural sulfur-containing peptide was pursued. The natural peptide Bz-GLPACGG-$NH_2$, where Bz is benzoic acid which does not fluoresce, was synthesized in the same way as the fluorescent peptides, using SPPS (Methods) (Figure 9A). This "crystallization peptide" was successfully completed and purified. As it does not contain any chromophores, this peptide appeared clear in solution, and white in powered form, which is unlike the fluorescent peptides described earlier in this paper. The mass of this peptide was calculated to be 676.30 Da. Because the peptide contains a cysteine residue, it has the potential to form self-dimers via disulfide bond (Figure 9B). The mass of this dimer was calculated to be 1350.58 Da. The monomer and dimer were both determined, via mass spectrometry, to elute at the same time, and were collected in the same screw-cap vial. Immediately before the peptide-containing peak, eluted another high-intensity peak, which was determined to not contain any monomer or dimer of the desired peptide. It was nonetheless curious due to its high intensity signal and similar elution time to the peptide. The entire crude peptide solution was purified and lyophilized, but the resulting powdered peptide was not quantified.

The purification of this peptide encountered the same difficulties involving pressure spikes that were described earlier in this paper, and approached troubleshooting in the same way.

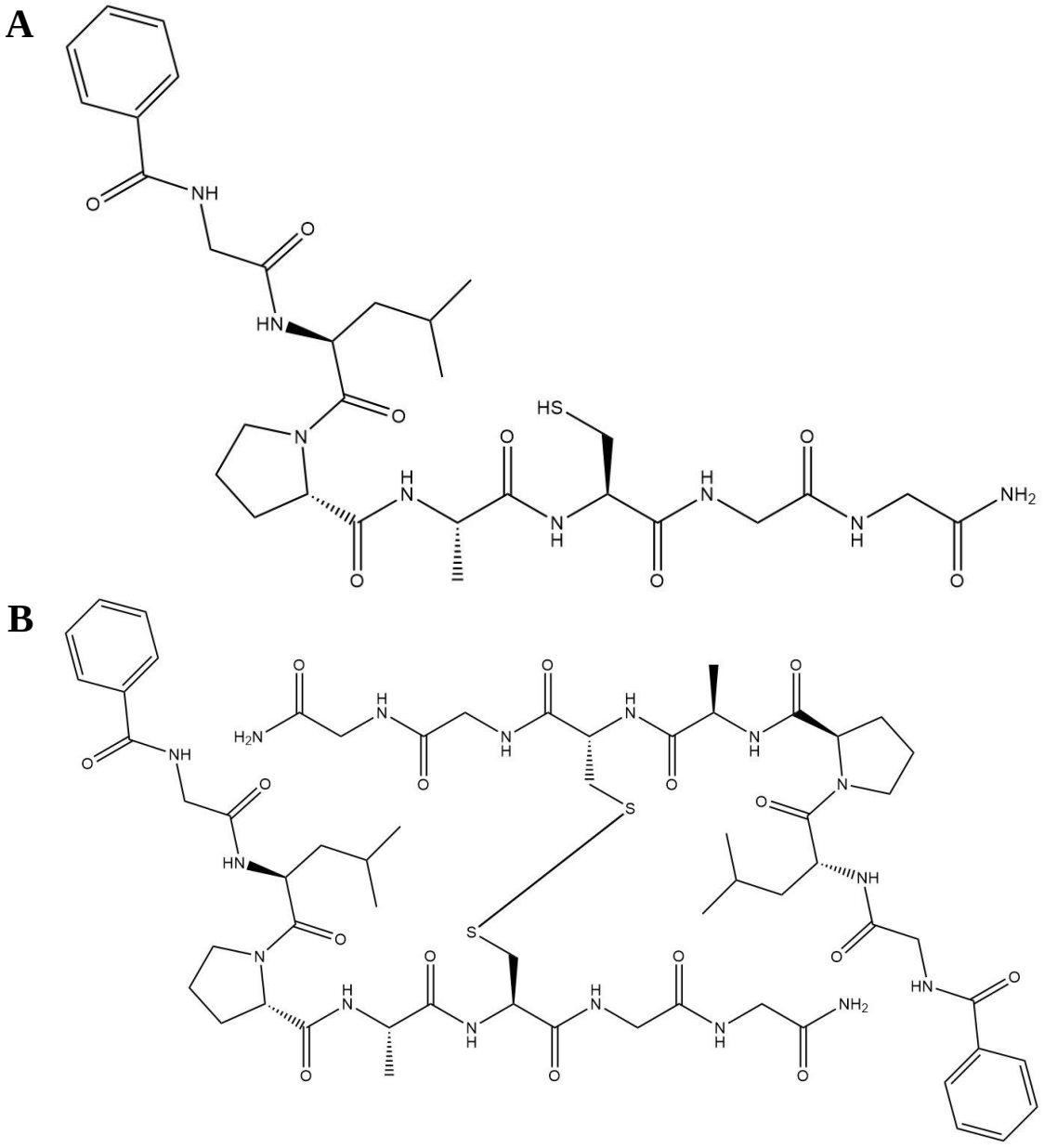



**Figure 9. A.** The structure of monomeric Bz-GLPACGG-$NH_2$. **B.** The structure of dimeric Bz-GLPACGG-$NH_2$, showing expected atom connectivity with an exaggerated disulfide bond.

## Future directions

Co-crystallization of the Bz-GLPACGG-$NH_2$ peptide with sortase A variants will be pursued in future work. Co-crystallization is not anticipated to be attempted with $SrtA_{strep}$-based "loop-swapped" variants, due to the difficulty of crystallizing $SrtA_{strep}$ in a biological monomer.[7]

Instead, co-crystallization will be attempted with wild type and potentially also with "loop-swap" variants of sortase A homologues that have already been shown to crystallize well. The sortase A homologues of *S. agalactiae* and of *S. mutans* have both been shown to crystallize in simple conditions and are candidates for this future work.[8–10]

## Methods

*Protein expression and purification.* Kanamycin-resistant plasmids (Genscript) containing the hexahistidine-tagged sortase variant genes were expressed in *E. coli* cells. Cells were grown in 10 ml of LB or 2XYT media with Kanamycin (Kan) ("overnight" culture) for 16 to 18 hours at 37 °C, with shaking, then grown in 1 L of the same media, also with Kan and also at 37 °C, with shaking, until $OD_{600}$ readings were between 0.6 and 0.8. Cultures were then induced for overexpression with isopropylthio-β-D-galactoside (IPTG) and the temperature was lowered to 18 °C. After overexpressing for 18-20 hours, the growth was pelleted via centrifugation, and the pellet was frozen in liquid nitrogen and stored in -80 °C. The cell pellet was thawed in a room-temperature water bath, then resuspended in 30 mL of lysis buffer via vortexing. The resulting cell suspension was lysed via sonication for 30 seconds, two times. This cell lysate was pelleted via centrifugation. A simplified workflow of this process is presented in Figure 10A.

The supernatant was then run through a nickel column and washed with low imidazole buffer, then eluted with high imidazole buffer. In some cases, before further purification, the hexahistidine tag was cleaved using tobacco etch virus protease (TEV protease), and the cleaved-sortase containing eluent was run through a second nickel column. Then, the cleaved-sortase containing wash was further purified via size exclusion chromatography on an ÄKTA prime plus FPLC, and the monomer-containing fractions were assessed for impurities and preliminary presence of desired protein via tricine gel electrophoresis. A simplified workflow of this process is presented in Figure 10B.

*Solid State Peptide Synthesis (SPPS).* All natural peptides were synthesized using solid state peptide synthesis (Figure 11). Each peptide was synthesized in a batch size of 0.1 mmol, using ~0.1724g Rink Amide Resin (resin) as the solid support, with each peptide in its own plastic synthesis vessel with a 20 μm filter (Resprep) to prevent the evacuation of resin during vacuum-draining. The resin was swollen in excess 1-methyl-2-pyrrolidinone (NMP) before addition of the first amino acid.

Each amino acid, as well as the resin, contained an N-terminal Fluorenylmethyloxycarbonyl (Fmoc) protecting group. Prior to each conjugation, the Fmoc group was cleaved off of the N-terminus of the resin or the most recent amino acid addition (Figure 11). The resin and attached peptide were washed with an excess of 20% piperidine in NMP, with rocking at ~65 rpm for 10-20 minutes, two times. The solution was then washed with NMP with rocking at ~65 rpm for 5 minutes, three times, to remove the free Fmoc group from solution (Figure 12A). Between each wash, the vessel was vacuum-drained, in order to remove unwanted chemical debris such as the free Fmoc group and unreacted piperidine.

In a separate vial, the C-terminal amino acid, which was added in 3x molar excess, was primed using 3x molar excess HBTU and Diisopropylethylamine (DIPEA) (Figure 11). The amino acid mixture was vortexed to mix, then decanted into the synthesis vessel (Figure 12B, C). The amino acid vial was washed with NMP to ensure the transfer of all reagents. The resulting conjugation solution was rocked at ~65 rpm for a minimum of 2 hours. After the conjugation time was complete, the solution was vacuum-evacuated to remove the byproducts of the conjugation reaction the solution, then washed with NMP with rocking at ~65 rpm for 10 minutes, three times. Vacuum-draining was also performed between these washes. Abz and Bz were conjugated using the same method.

After each addition, a Kaiser assay was performed, described below. In the event of a successful result, the conjugation reaction was considered complete, and the process of conjugating the next amino acid, using the same procedure, was started.

Once all residues and Abz or Bz in a peptide had been conjugated, the completed peptide was washed with NMP with rocking at ~65 rpm for 10 minutes, three times, to remove chemical debris from conjugation. Then, the peptide solution was washed with excess DCM with rocking at ~65 rpm for 10 minutes, three times. Next, the peptide solution was washed in trifluoracetic acid (TFA), for 30 minutes, two times, in order to cleave the peptide off of the resin. The TFA washes were performed without rocking, and without a synthesis vessel lid to prevent pressure buildup. The TFA was then removed from solution via rotovap, and the peptide was added dropwise to ice cold ether. The ether-peptide solution was centrifuges at 4500 rpm for 5 minutes at 4 ℃, and the ether supernatant was decanted off. The crude peptide-containing pellet was vacuum-dried, then stored at 4 ℃ (Figure 12D).

*Kaiser Assay.* The Kaiser assay is a colorimetric assay that detects the presence of free primary amines in solution. In the presence of these amines, the solution will turn a distinctive musky blue color.

The Kaiser test kit consists of three reagents. In order to test each conjugation, 10 μL of each reagent were pipetted into a clean, labelled microcentrifuge tube, and additionally one tube for a positive control and one tube for a negative control. These tubes were centrifuged for a few seconds to ensure that no reagents were on the sides of the tubes. Then, vacuum-dried resin from the conjugation solution, positive ninhydrin powder, and negative ninhydrin powder, were placed

into their appropriate tubes, via pipette tip. These tubes were capped, then incubated at 95 ℃ for 5 minutes. A successful result was indicated by no color change in the conjugated resin solution, which was either clear, or faintly yellow for peptides containing the chromophore DNP. An unsuccessful result was indicated by a musky blue color.

*Mass Spectrometry.* Electron Spray Ionization Mass Spectrometry (ESI-MS) was performed on an Advion Expression Mass Spectrometer. ESI-MS was used to positively identify the presence of the desired peptide in the initial spectrum collection peaks. The mass found in each collected peak was compared to the calculated mass of the peptide (ChemDraw).

*High Performance Liquid Chromatography Peptide Purification.* Peptide purification was performed on a Dionex UltiMate 3000 UHPLC with Diode Array detector. The crude peptide was dissolved in just enough 1:1 acetonitrile: sterile $H_2O$, and 10 μL was added to either a microcentrifuge tube or an HPLC vial. Each type of vessel required a base liquid volume, to ensure that the HPLC injection needle could remove the full volume of injection without taking up any bubbles. If a microcentrifuge tube was used, then 30 μL of sterile H20 was used. If an HPLC vial was used, 25 μL of sterile H20 was used.

First, 5 μL of sample was injected onto a semi-prep column (Phenomenex Luna 5 μM C18(2) 100Å LC 250 x 10 mm column) and eluted using a gradient (beginning at 20% buffer B (0.1% formic acid in acetonitrile), 80% buffer A (0.1% formic acid in 95% nanopure $H_2O$, 5% acetonitrile)). A UV-Vis spectrum of the elution profile was collected (“initial spectrum”). Data was collected for every wavelength between 200 nm – 700 nm, and was visualized at whichever

wavelength provided clearest visualization of peaks. Then, another 5 μL was injected, and eluted using the same method, and the peak(s) most likely to be the peptide of interest were collected in clean, labelled screw-cap vials. Peaks were judged based on their size, and to a lesser extent, their elution time.

A 5 μL aliquot of each peak collection was added to a new microcentrifuge tube or HPLC vial with the proper base of sterile H2O, and 5 μL of each diluted sample was injected onto the analytical column (Phenomenex Aeris 3.6 μM WIDEPORE C4 200Å LC 150 x 21 mm column), which was eluted onto the mass spectrometer to confirm presence of the desired peptide in that peak.

Then, enough crude peptide solution for several injections was added to the original microcentrifuge tube or HPLC vial. Depending on the estimated concentration of the crude peptide solution, several injections of volume 5 μL -100 μL were run in series, and the appropriate peaks were collected in the appropriate screw-cap vials. This was repeated until no crude peptide solution remained. Once the sample in the microcentrifuge tube or HPLC vial was at the base liquid volume, additional 1:1 acetonitrile: sterile $H_20$ was added to the sample to dilute the peptide and allow for further injections. This was repeated until the UV-Vis signal was at sufficiently low intensity, indicating that most of the peptide had been collected (Figure 12D).

*Peptide Storage.* Following purification, or in the case that purification could not be completed, the acetonitrile was removed from each purified or crude peptide solution via vacuum evaporation (Figure 12D). Then, the solution was frozen in dry ice at an angle, in order to maximize exposed surface area within the labelled screw-cap vial. The cap of the vial was then replaced with a porous Kim Wipe, which was secured with a rubber band, and the water was

removed via lyophilization. After lyophilization, the caps were replaced on the screw-cap vials and powdered peptides were stored at 4 ℃.

*Synthesis of the unnatural amino acid T**. N-Boc-S-Trityl-(D)-Cysteine (Bis-protected Cysteine) was used as a starting material (Figure 6B). The planned synthetic scheme was four steps: First, reduction of the carboxylic acid C-terminus to a primary alcohol. Second, the oxidation of that primary alcohol to an aldehyde. Third, the nucleophilic alkylation of the aldehyde, to mimic the side chain of threonine. Fourth, deprotection of the N-Boc group. The synthesis would be followed by deprotection of the S-Trityl group, and finally, the addition of the resultant unnatural amino acid (T*) to the C-terminus of an LPA peptide, using solid phase peptide synthesis. The structure of the LPAT* peptide is shown in Figure 6C.

*Synthetic step one for T**. Bis-protected Cysteine (1 g, or 0.002157 mol) was dissolved in 5 mL of anhydrous THF in anhydrous conditions under argon gas. Molar excess of $BH_3$-THF complex was added, slowly, via syringe, at 0 °C. The mixture was allowed to react, with stirring, for 20 minutes at 0 °C. It was then allowed to react for 1 hour, with stirring, at room temperature (approximately 25 °C). Then, the reaction was quenched with molar excess methanol and concentrated. The quenching step was performed twice to ensure full deactivation of the $BH_3$.

*Purification of T* intermediate.* A small portion of the crude product was saved as a control for thin layer chromatography (TLC) monitoring of the purification. A flash-column chromatography column was used to purify the crude product. The column was packed with

Silia*Flash* P60, 40-63 μm irregular silica gel (SiliCycle) as the stationary media, and run in a glass column with a 1 L bulb. Laboratory-grade sand (Macron Fine Chemicals) was used as a filter at the top and bottom of the silica layer, and cotton was used as a porous plug at the bottom of the column. A 1:1 Methane: Hexanes solution was used as the mobile media to elute the reduced cysteine product off the column. Fractions of approximate volume 36 mL were collected in borosilicate glass disposable culture tubes.

*Nuclear Magnetic Resonance (NMR) Spectroscopy.* NMR spectra were collected on either a Varian MercuryPlus 300 MHz FT-NMR instrument or a Bruker Avance III 500 MHz FT-NMR instrument, at Western Washington University. All NMR samples were in Chloroform-D solvent.

**Acknowledgements**

The author is grateful to Dr. Jeanine Amacher, Dr. John Antos, and Dr. Jim Vyvyan for their mentorship and guidance throughout this project. Dr. Amacher helped to plan and direct funding for the project, provided lab space and supplies, and served as an advisor for all parts of the project. Dr. Antos also contributed lab space and supplies, and advised on many parts of the project especially on natural peptide synthesis. Dr. Vyvyan also contributed lab space and supplies, and advised on the unnatural amino acid synthesis portion of the project, including analysis of NMR data. The author would also like to thank Dr. Hla Win-Piazza for their assistance in collecting and analyzing NMR data.

The author thanks Sarah Struyvenberg for her training in protein purification and her leadership on the "loop-swapped" project, as well as her contributions to developing the fluorescence assay. The author thanks Izzi Piper for her training in SPPS, as well as her contributions to developing the fluorescence assay. The author thanks Nicholas M. Horvath for his previous work with the Bz-GLPACGG-$NH_2$ peptide in Dr. Antos' lab. The author thanks Jordan Valgardson for his computational work with the β-7/β-8 loop. The author would also like to recognize Savanna Takasaki for her work creating cyclic structures using sortase A, which are mentioned in the introduction.

The author thanks Alex Johnson for the use of his figure of the "loop-swapped" sortase variants (Figure 4). The author additionally thanks Izzi piper for the use of her figures of the heat maps of the activity of the "loop-swapped" sortase variants (Figure 2).

The author would also like to thank the other members of the Amacher lab, especially Melody Gao, Iain Mackley, and Haley Bamonte, for their unwavering support and enthusiastic investment in this project. The author would like to thank Chris Swanson for his patience in

answering many questions about biochemistry lab work. The author is also grateful to Mariah Kollasch and Michael Lee for their friendship and interest in this project. The author would like to thank Eileen Lee, Nancy Brill, and Aaron Johnston, for their friendship, support, and invaluable feedback on the communication of this project to people without a background in biochemistry.

## References


(1) Jung, M. E.; Clemens, J. J.; Suree, N.; Liew, C. K.; Pilpa, R.; Campbell, D. O.; Clubb, R. T. Synthesis of (2R,3S) 3-Amino-4-Mercapto-2-Butanol, a Threonine Analogue for Covalent Inhibition of Sortases. *Bioorganic & Medicinal Chemistry Letters* **2005**, *15* (22), 5076–5079. https://doi.org/10.1016/j.bmcl.2005.07.073.

(2) Spirig, T.; Weiner, E. M.; Clubb, R. T. Sortase Enzymes in Gram-Positive Bacteria: Sortase Enzymes in Gram-Positive Bacteria. *Molecular Microbiology* **2011**, *82* (5), 1044–1059. https://doi.org/10.1111/j.1365-2958.2011.07887.x.

(3) Antos, J. M.; Truttmann, M. C.; Ploegh, H. L. Recent Advances in Sortase-Catalyzed Ligation Methodology. *Current Opinion in Structural Biology* **2016**, *38*, 111–118. https://doi.org/10.1016/j.sbi.2016.05.021.

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

(5) Zong, Y.; Bice, T. W.; Ton-That, H.; Schneewind, O.; Narayana, S. V. L. Crystal Structures of Staphylococcus Aureus Sortase A and Its Substrate Complex. *J. Biol. Chem.* **2004**, *279* (30), 31383–31389. https://doi.org/10.1074/jbc.M401374200.

(6) Michael E. Jung; Sung Wook Yi. Synthesis of Threo-b-Aminoalcohols from Aminoaldehydes via Chelation-Controlled Additions. Total Synthesis of L-Threo Sphingosine and Safingol. *Tetrahedron Letters* **2012**, *53* (32), 4216–4220. https://doi.org/10.1016/j.tetlet.2012.05.153.

(7) Misra, A.; Biswas, T.; Das, S.; Marathe, U.; Roy, R. P.; Ramakumar, S. RCSB PDB - 4O8L: Structure of sortase A from Streptococcus pneumoniae https://www.rcsb.org/structure/4o8l (accessed Jun 16, 2020).

(8) Khare, B.; Samal, A.; Vengadesan, K.; Rajashankar, K. R.; Ma, X.; Huang, I.-H.; Ton-That, H.; Narayana, S. V. L. Preliminary Crystallographic Study of the Streptococcus Agalactiae Sortases, Sortase A and Sortase C1. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **2010**, *66* (Pt 9), 1096–1100. https://doi.org/10.1107/S1744309110031106.

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

(10) Wallock-Richards, D. J.; Marles-Wright, J.; Clarke, D. J.; Maitra, A.; Dodds, M.; Hanley, B.; Campopiano, D. J. Molecular Basis of Streptococcus Mutans Sortase A Inhibition by the Flavonoid Natural Product Trans-Chalcone. *Chem. Commun.* **2015**, *51* (52), 10483–10485. https://doi.org/10.1039/C5CC01816A.

(11) Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I. NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics* **2010**, *29* (9), 2176–2179. https://doi.org/10.1021/om100106e.

**Supporting Information**

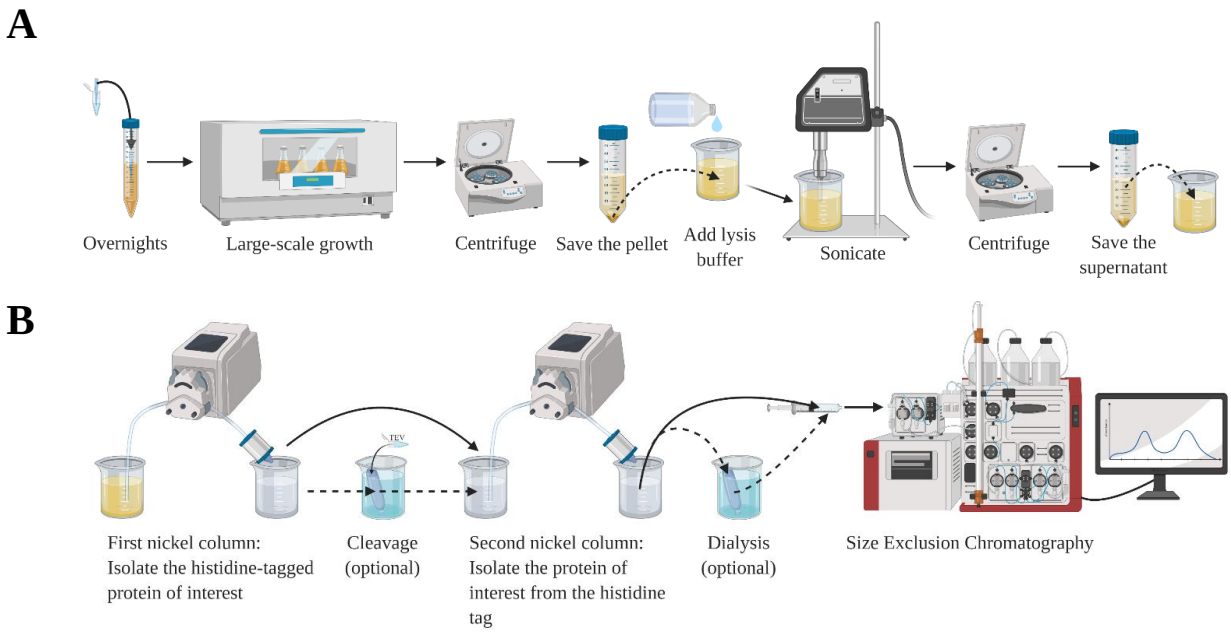



**Figure S1. A.** Workflow of protein expression and retrieval from bacterial cells **B.** Workflow of protein purification.

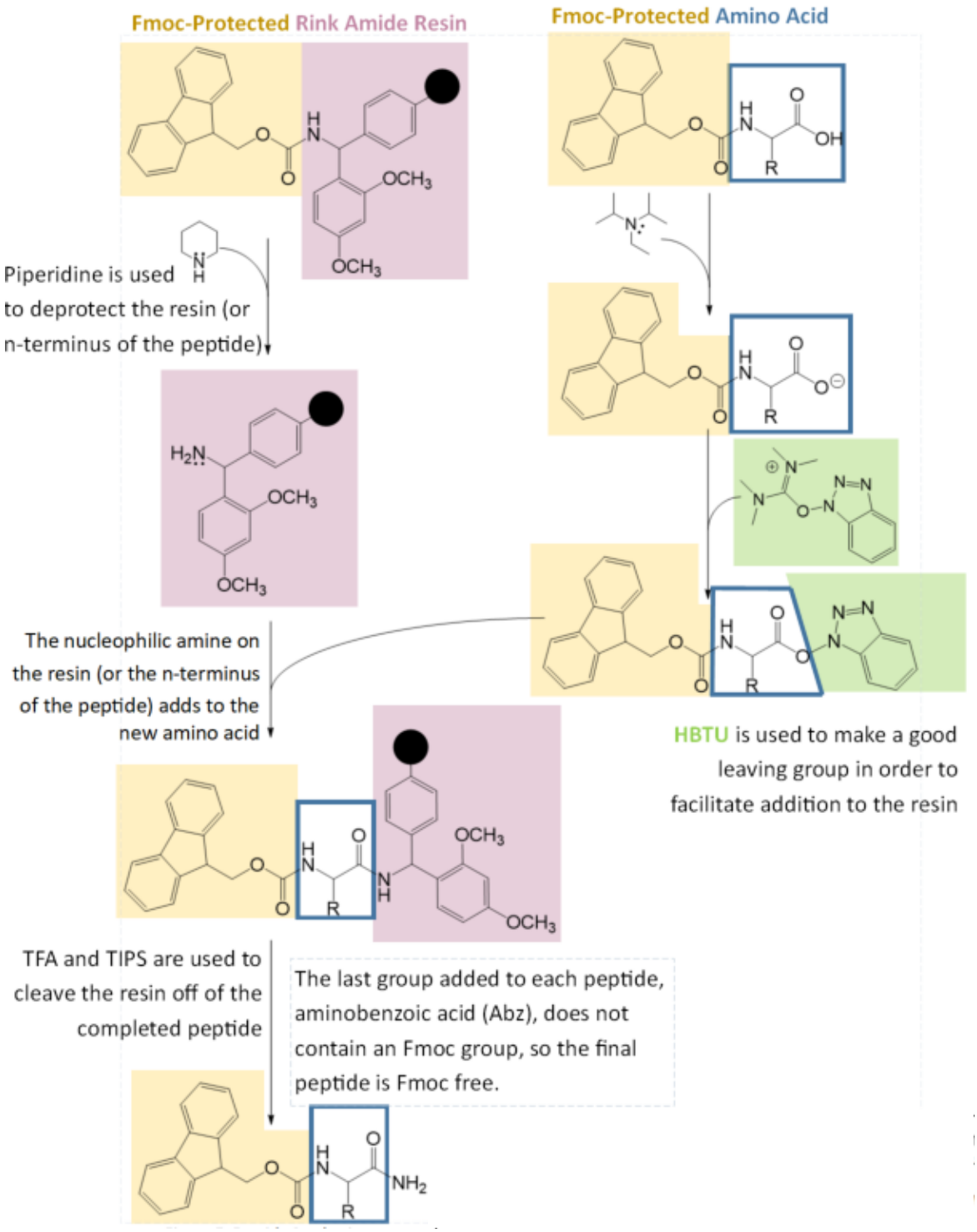



**Figure S2.** Simplified scheme of one round of the amino acid addition in solid phase peptide synthesis.

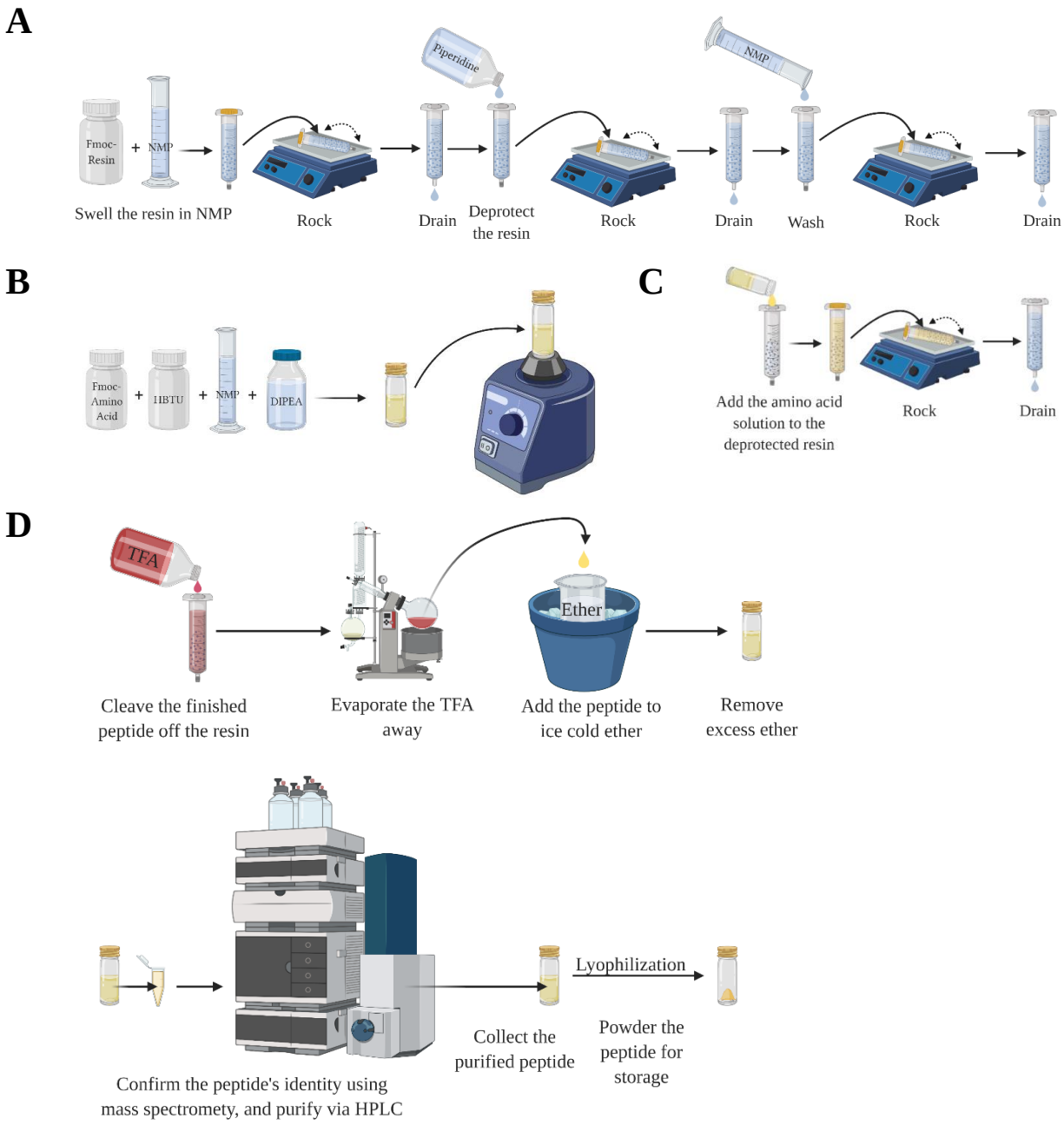



**Figure S3.** Workflow of laboratory procedure for solid phase peptide synthesis **A.** Workflow of Fmoc deprotection. **B.** Workflow of amino acid preparation for conjugation. **C.** Workflow of resin/peptide and amino acid conjugation. **D.** Workflow of peptide cleavage from resin, and purification.