University of Nevada, Reno

Characterizing Sortase A Mutants with Altered Binding Selectivity

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in
Chemical Engineering

by

Alexander J. Bolt

Dr. Maryam Raeeszadeh-Sarmazdeh/Thesis Advisor

August 2022

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

entitled

be accepted in partial fulfillment of the
requirements for the degree of

Advisor

Committee Member

Graduate School Representative

Markus Kemmelmeier, Ph.D., Dean
Graduate School

ABSTRACT

Post-translational protein fusion has potential to meet the diverse and complex needs of many current problems in biotechnology. Traditional approaches such as genetic fusion are heavily constrained to simple protein fusions that can be expressed within the host cell. Enzymatic ligation mediated by *Staphylococcus aureus s*ortase A (SrtA) broadens the scope of protein fusion by ligating proteins post-translationally, given that the target proteins bear the two substrates recognized by SrtA: short peptides of 3-5 amino acids in length. Nonetheless, one of these tags, polyglycine, is forcibly constrained to the protein N-terminus. Previous efforts have sought to overcome this limitation by engineering SrtA with substrate specificity for a motif found in pilin formation, known as pilin box, using directed evolution. In this work, engineered SrtA mutants are analyzed both on the yeast surface and in soluble form for altered expression levels or activity toward natural polyglycine or the pilin substrate. Expression levels in some mutants increased by over 100-fold. Mutants on yeast showed higher selectivity for pilin box, no longer recognizing polyglycine and maintaining pilin box specificity at a level comparable to eSrtA, an evolved SrtA with higher catalytic activity. A reliable, facile method is used for the quantification of SrtA-mediated ligation product yields using elastin-like polypeptide (ELP) purification and fluorescence assays. SrtA WT showed specificity for both GGG and P0, whereas mutants 1 and 8 were selective toward P0 and GGG, respectively. These findings provide new approaches to characterize transpeptidases with more than one substrate on the yeast surface or in solution, and pave the way for future site-specific protein ligation tools to generate post-translational protein fusions.

ACKNOWLEDGEMENTS

Though graduate school has arguably been the most challenging period in my life thus far, I have also never been blessed with a better support network. Thank you to Dr. Maryam Raeeszadeh-Sarmazdeh and my lab mates for always pushing me to see the best in myself. Thank you to my office coworkers for the laughs and making our work environment hospitable. Mom, dad and my church community: thank you for lifting my spirits when I felt like giving up on myself. Finally, Francesca: I do not know if there could have been a better time for you to enter my life than during the most stressful part of my degree, with the finish line in sight but still so much to get done. Thank you for standing by my side until the end.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 – BACKGROUND & SIGNIFICANCE

The need for more sophisticated protein fusion tools is rapidly growing in the fields of protein and bioprocess engineering. Previous works have used genetic fusion in a range of cellular hosts for the heterologous expression and functionalization of many protein assemblies, including protein nanoparticles (PNPs), metabolons and scaffolds[1]. Introducing short peptide tags or antigens into a DNA sequence for heterologous protein expression does not typically impact protein nanostructure stability or subunit folding efficiency. However, the limitations of genetic fusion manifest when fusing more complex proteins, as these fusions alter the thermodynamic properties of each protein, often resulting in misfolding or aggregation[2,3]. Phage and bacterial expression offer relatively high yields of fusion product at the cost of simple cellular machinery, often leading to inclusion bodies, which require significant time and labor to restore functionality. Conversely, eukaryotic systems like mammalian cells can fold more complex proteins, but the significant reduction in protein yield and difficult maintenance associated with mammalian cell cultures may not be commercially viable[4]. Further complications in genetic fusion arise when proteins originate from different cellular hosts or require sophisticated post-translational modification[5].

Sortase (SrtA)-mediated ligation offers a robust and flexible alternative to traditional genetic fusion through bypassing the need for cells to directly express and fold the protein fusion. SrtA from gram-positive *Staphylococcus aureus* is a calcium (Ca^{2+})-dependent transpeptidase responsible for covalently attaching proteins to the cell wall through a two-step reaction mechanism. The protein of interest bears a C-terminal LPXTG motif, where X is a guest residue (most commonly glutamic acid, or E)[6]. SrtA forms a

thioacyl intermediate with the protein by cleaving between threonine and glycine. Then, an exposed pentaglycine in lipid II on the cell surface performs a nucleophilic attack on the reactive LPXT-SrtA intermediate, thus liberating SrtA and anchoring the protein of interest to the cell wall[7]. Likewise, SrtA-mediated ligation can be performed *in vitro* (Fig. 1). Expressing SrtA in soluble form without its transmembrane-anchoring domain completely retains enzyme activity, and yields of up to 10 mg/L culture of functional SrtA are easily obtained in *E. coli*. Any two proteins "A" and "B" are expressed individually with their respective recombinant tags, LPXTG and polyglycine (though shorter glycine repeats or an aminomethyl group can alternatively participate as the nucleophile with reduced efficiency)[8]. Being only 3-5 amino acids long, these short motifs rarely affect protein stability[5,8]. SrtA-mediated ligation is compatible with physiological conditions but also possible under wide ranges of reaction time, temperature, and pH[5].

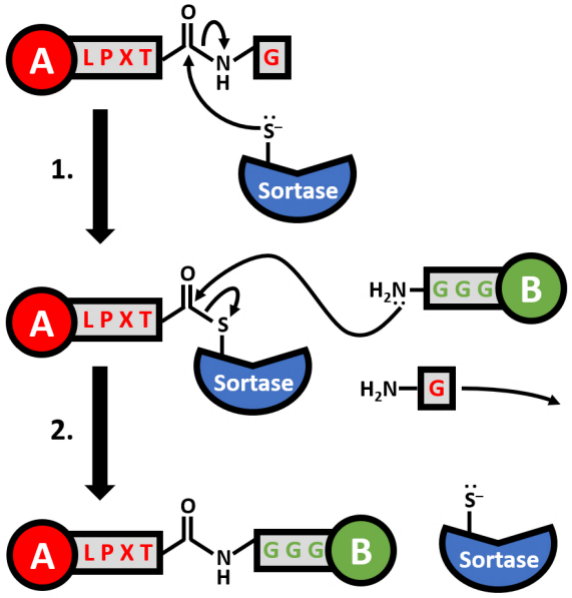


Fig. 1: SrtA-mediated ligation is a two-step reaction mechanism where: 1) SrtA forms a thioacyl intermediate by cleaving between threonine and glycine of LPXTG and 2) polyglycine performs a nucleophilic attack on the LPXT-SrtA intermediate, forming the protein fusion.

Recent SrtA engineering efforts demonstrate the potency of *in vitro* SrtA-mediated ligation as a protein fusion tool with ever-increasing opportunities. Researchers have repeatedly demonstrated the power of directed evolution and rational design approaches to alter or enhance SrtA function, such as improving LPETG-coupling activity[9], removing Ca^{2+} dependency[10], or altering LPXTG substrate specificity[11,12]. Improved SrtA mutants have aided in sophisticated protein-based nanostructure assembly, including: using PNPs to encapsulate therapeutics or decorating the PNP surface with reporter proteins[13-16], linking enzymes for improved substrate channeling or nanoscale bioreactors[17,18], and immobilizing proteins on scaffolds with applications in biosensing or tissue engineering[8,19-21]. One current limitation of the SrtA-mediated ligation toolbox, however, is the constraint of the polyglycine (GGG) tag to the N-terminus of protein "B." The nucleophilic role of GGG is attributed to the N-terminal amine, but certain protein N-termini are buried in the hydrophobic core or required for functional activity, like with tissue inhibitors of metalloproteinases (TIMPs)[22]. Outside of chemo-enzymatic methods[23], if both proteins simultaneously require a free N-terminus or C-terminus, then viable options for SrtA-mediated ligation are limited.

This work seeks to outline the advantages of a different substrate, pilin box, and demonstrate that SrtA can be engineered with altered specificity toward pilin box. Pilin box (P0) is a 13-amino acid tag that mimics the well-characterized SpaA pilin protein recognized by *Corynebacterium diphtheriae* strain NCTC13129 SrtA. This motif obeys the consensus sequence WXXXVXVYPKN distinguished by pilin sortases[7,24,25]. The nucleophile in P0 is the ε-amino group in the side chain of lysine, present in position 10 of the amino acid sequence. K10 in P0 forms an isopeptide bond with the carboxyl group of

threonine in LPETG. Thus, fusing protein "B" to P0 enables N-to-N-terminal or C-to-C-terminal protein fusions (Fig. 2A).

In a previous work, directed evolution was used to generate SrtA mutants with altered P0 specificity[26]. Wild-type SrtA (SrtA WT) and a previously engineered SrtA mutant with improved LPETG-coupling activity (eSrtA) showed little specificity toward P0. These mutants were the end product of three consecutive sorts, each involving: random mutagenesis (of the eSrtA sequence), modified yeast surface display (where the SrtA mutant and LPETG are co-expressed), fluorescent labeling (to detect expression and P0 reaction), and cell sorting (via fluorescence-activated cell sorting or FACS). After the third sort, the DNA sequences of the highest-performing mutants were isolated (Fig. 2B).

The body of this work is divided into two major approaches with one ultimate goal: determine whether SrtA substrate specificity has broadened toward both the GGG and P0 substrates, or shifted from GGG to P0. First, isolated mutants, co-expressed with LPETG on the yeast surface, have been analyzed via flow cytometry to quantify expression and fusion product formation through fluorescent signals. Second, SrtA mutants are expressed in soluble form, and SrtA-ligated products contain a fluorescent protein fused to an elastin-like polypeptide (ELP) repeat bearing the C-terminal LPETG motif. After non-chromatographic ELP purification, yields are detected through sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis and fluorescent signals. SrtA activity is also characterized through by reaction with a fluorogenic LPETG peptide. Notably, this ELP purification method is a useful alternative to conventional SrtA activity assays, which measure the rate of hydrolysis of the LPET-SrtA thioacyl intermediate rather than the rate of fusion product formation. Finally, the sequence-structure-function

relationship of SrtA mutants has been studied through structural analysis on PyMOL, a proprietary open source molecular visualization system, using a well-known crystal structure of LPET-bound wild-type *S. aureus* SrtA[27].

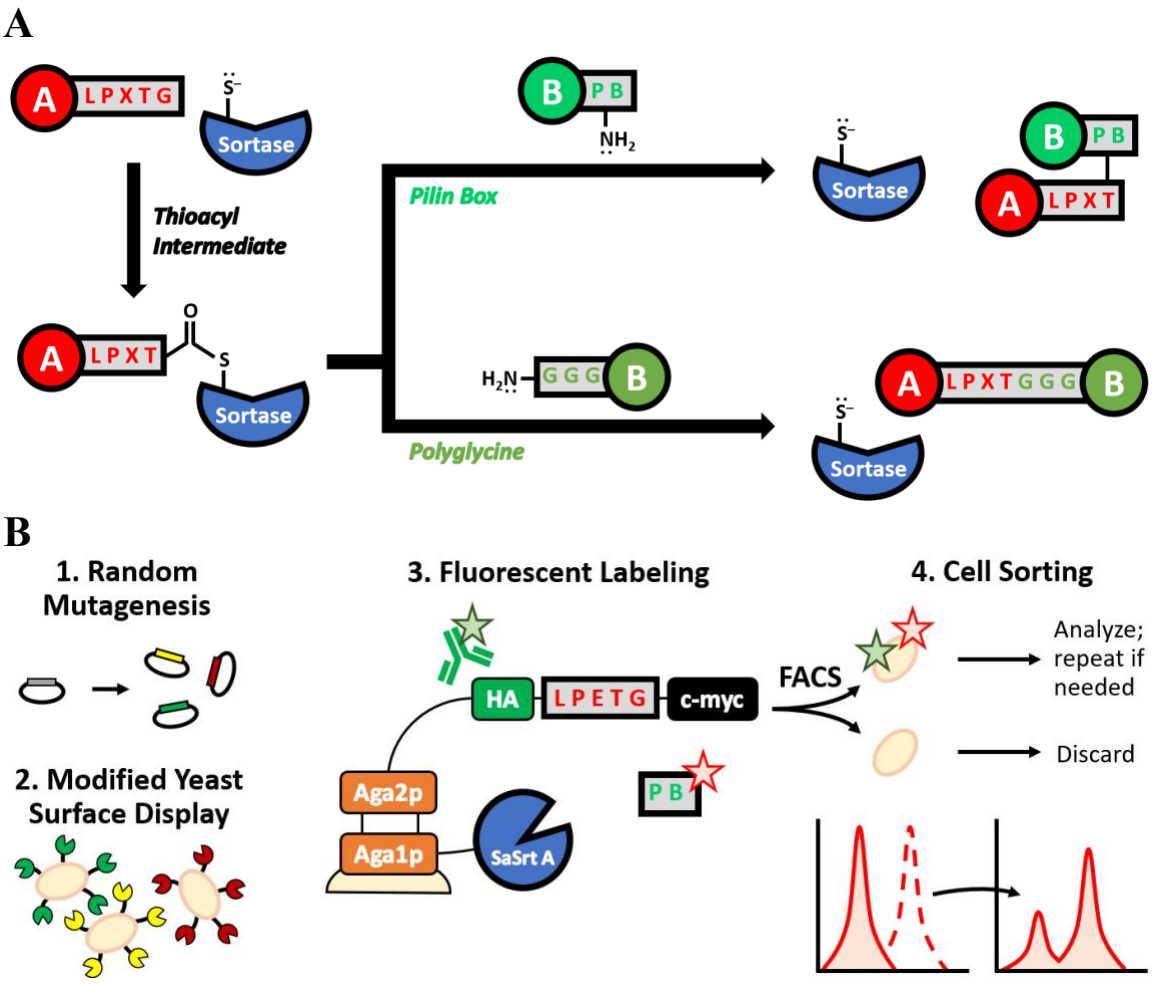


Fig. 2: (A) A second substrate, pilin box (P0), is proposed. P0 reacts via the ε-amino group of a lysine side chain, rather than the N-terminal amine. Thus, as illustrated, the SrtA recognition tag on protein "B" is no longer constrained to the N-terminus. **(B)** In a previous work, SrtA mutants were generated via directed evolution. A mutant library was formed by subjecting the eSrtA DNA sequence to random mutagenesis. The library was screened through modified yeast surface display, fluorescent labeling, and cell sorting. These four steps were repeated for a total of three sorts.

CHAPTER 2

ANALYZING SELECTIVITY OF SORTASE MUTANTS THROUGH YEAST SURFACE DISPLAY

2.1: Introduction

Yeast surface display (YSD) of SrtA provides a robust, powerful platform to analyze SrtA mutant expression and enzymatic activity. YSD involves genetically fusing the proteins of interest to the a-agglutinin surface protein, a two-subunit glycoprotein composed of Aga1p and Aga2p that are linked by disulfide bonds. Proteins can be fused to the N or C-terminus of Aga1p or Aga2p, though Aga1p anchors the protein complex on the yeast cell membrane[28]. Yeast transformants are grown with selective markers to ensure plasmid retention e.g., the pCT302 plasmid carrying the TRP1 gene in tryptophan-deficient media[29]. Different epitopes (e.g. c-myc, HA) or peptides can be easily introduced for labelling with fluorescent antibodies or probes, respectively. Finally, flow cytometry is used to detect fluorescent signals and, when normalized against the proper controls, can provide quantitative analysis of variables like expression and activity.

This work utilizes a modified YSD, previously developed for engineering SrtA specificity toward P0[26], to analyze SrtA mutants. Wild-type SrtA or its mutants are fused to the C-terminus of Aga1p, while LPETG is fused to the C-terminus of Aga2p. Notably, this setup allows the co-expression of SrtA and LPETG on individual plasmids, enabling independent genetic modification of the Aga1p and Aga2p constructs. Furthermore, similar to how previous works have improved the poor reaction kinetics of SrtA by immobilization of SrtA and LPETG on a cobalt resin[18], this system likewise takes advantage of the YSD platform to immobilize both SrtA and LPETG in close proximity. Soluble GGG or P0

nucleophiles are fused to green fluorescent protein (GFP), mCherry, or biotin (bGGG and bP0) and reacted on the yeast surface, where the SrtA transpeptidation reaction can be detected by fluorescent signals through flow cytometry (Fig. 3).

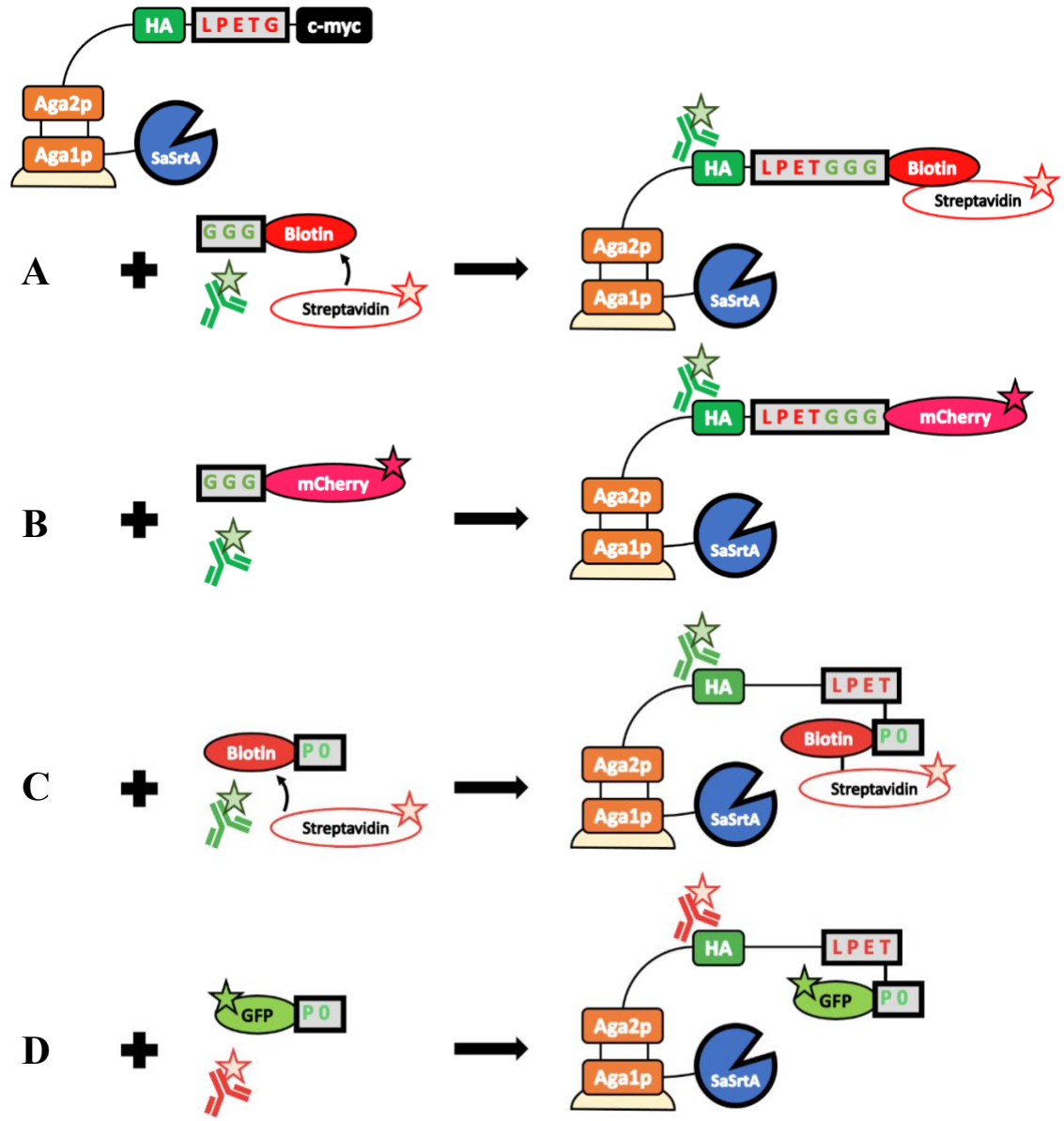


Fig. 3: SrtA and LPETG were displayed on the yeast surface via fusion to the Aga1p-Aga2p complex. SrtA reaction was measured through fluorescent labels and flow cytometry after adding **(A)** bGGG, **(B)** GGG-mCherry, **(C)** bP0, and **(D)** GFP-P0. Biotinylated peptides were labeled with streptavidin-conjugated fluorophores. The fluorescence of GFP and mCherry could be measured directly, requiring no further labeling or reagents.

2.2: Experimental Methods

Strains and Plasmids

pCT plasmids, containing Aga2-HA-LPETG-cmyc, and pCU plasmids, containing Aga1-Srt-AU1, were sequenced using the GAL forward primer. For DNA extraction and sequencing, plasmids were transformed into NEB5α competent *E. coli*. For yeast surface display, plasmids were transformed into BJ5465 (MATa ura3-52 trp1 leu2-delta1 his3-delta200 pep4::HIS3 prb1-delta1.6R can1 GAL) *Saccharomyces cerevisiae* cells.

Soluble Expression and Purification of Fluorescent Proteins

The pET-GGG-mCherry and pET-GFP-P0 plasmids were transformed into BL21 (DE3) cells (Thermo Scientific). Cultures containing 10 mL LB broth supplemented with 100 mg/mL kanamycin were inoculated with transformants on LB kan plates and grown overnight. 500 mL cultures of LB kan broth were inoculated with overnight cultures to a final optical density (OD) between 0.05-0.1. Upon reaching an OD of 0.4-0.6, cells were induced with 0.5 mM IPTG and allowed to grow for 3-4 hours before centrifugation. Cell pellets were resuspended in 20 mL of tris-buffered saline (TBS) (50 mM Tris-HCl pH 7.6, 150 mM NaCl, 50 mM KCl) and 10 µL DNase I. Cells were sonicated, pelleted and passed through a 0.45 µm filter before loading onto a Ni-NTA gravity column. The flowthrough was applied to the column a second time, and proteins were visibly eluted with TBS containing 250 mM imidazole. If needed, proteins were concentrated using 15 mL Amicon centrifugal filter units with 10 kDa molecular weight cutoff according to manufacturer instructions.

Yeast Surface Display of SrtA Mutants

Cells were grown on SD-CAA plates deficient in tryptophan and uracil (-Trp/-Ura). Colonies were used to inoculate 5 mL SD-CAA (-Trp/-Ura) cultures, which were shaken for 16-20 h at 30 °C. Cultures were induced through a galactose-inducible promoter by inoculating 5 mL SG-CAA (-Trp/-Ura) with cells to a final optical density (OD) of 1. Cultures were shaken for another 16-20 h at 30 °C. During induction, each culture expressed Aga2-HA-LPETG-cmyc along with Aga1-Srt, where SrtA is SrtA WT or any SrtA mutant. Following induction, cells were pelleted to an OD of 0.6 and washed three times with phosphate-buffered saline (PBS) containing either 0.1% or 4% BSA (PBSA).

SrtA Reaction on the Yeast Surface

Polyglycine and P0 peptides with C-terminal biotin modification were received from Genscript. The lyophilized pellets were resuspended in TBS containing 10 mM $CaCl_2$. 100 µL of the peptide-buffer mixture was used to resuspend the yeast cell pellets, which were incubated in a thermocycler at 37 °C. GGG reactions were incubated at 100-500 µM for 3 h, and P0 reactions were incubated at 100-500 µM for 16 h. Upon reaction completion, cells were pelleted and washed 3-5 times with 0.1% or 4% PBSA. Cells were labelled with fluorescein isothiocyanate (FITC)-conjugated mouse anti-HA (human influenza hemagglutinin) to measure expression, mouse anti-cmyc with Alexa Fluor 488 to measure LPETG cleavage, and Alexa Fluor 647-conjugated streptavidin to measure SrtA reaction. Cells were incubated on ice for 1 h while protected from light, and then washed 3-5 times with either 0.1% or 4% PBSA.

Flow Cytometry

After a final resuspension in 750 µL 0.1% or 4% PBSA, samples were run on the BD Accuri C6 Plus flow cytometer. 10000 events were recorded per sample. FlowJo, a flow cytometry data analysis software, was used to gate the FITC-negative population in each sample. Overall and FITC-negative median fluorescence levels (MFL) were then extracted via FlowJo for both expression (FL-1; FITC) and reaction (FL-3; PerCP or perFL-4; APC). Final MFLs were calculated using the following formula:

$$MFL_{exp} = MFL_{exp}^{overall} - MFL_{exp}^{FITC-negative}$$

$$MFL_{rxn} = MFL_{rxn}^{overall} - MFL_{rxn}^{FITC-negative}$$

The subscripts "exp" and "rxn" stand for expression and reaction, respectively. The negative control (un-induced cells) MFL was subtracted from sample MFLs, and sample MFLs were normalized against eSrtA as shown:

$$nMFL_{exp}^{sample} = \frac{MFL_{exp}^{sample} - MFL_{exp}^{NC}}{MFL_{exp}^{eSrt} - MFL_{exp}^{NC}}$$

$$nMFL_{rxn}^{sample} = \frac{MFL_{rxn}^{sample} - MFL_{rxn}^{NC}}{MFL_{rxn}^{eSrt} - MFL_{rxn}^{NC}}$$

2.3: Results & Discussion

The yeast surface display platform was used to analyze SrtA mutants based on expression, LPETG cleavage, and reaction with either the GGG or P0 substrates. Select SrtA mutants were sequenced, and the DNA sequences were aligned to identify notable mutations (Table 1). Mutations were more likely to occur in flexible loops and regions known for either substrate recognition or enzymatic activity.

Table 1: Key mutations in SrtA mutants expressed on the yeast surface.

Region	Mut 4	Mut 7	Mut 11
H2 Helix			**N98K**
ß3 Sheet			**S102G**
ß3-ß4 Loop		**D111N**	**N107S**
ß4-H3 Loop	**F122S / N127S / Q129R**		**N127D**
H3 Helix			**N132D / A135T**
ß5 Sheet			**V142A**
ß6 Sheet	**R151H**		
ß6-ß7 Loop	**K177R**	**V166A / V168A**	
ß7 Sheet	**L181F / D186G / N188D**	**Y187H**	
ß7-ß8 Loop	**K190T**	**E189A**	
ß8 Sheet		**W194R**	

Three mutants, shown in Table 1, were selected for further analysis due to superior expression levels. Some amino acid positions showed multiple mutations, such as N127 in mutants 4 and 11. Other mutants (10 and 16) failed to show improvement in expression when displayed on yeast (Fig. B-1). As previously described, SrtA mutant MFLs were first normalized to SrtA WT or eSrtA when drawing conclusions about expression or reaction.

To best analyze SrtA activity with the GGG and P0 substrates, numerous conditions were tested. First, mutant activity was measured using two fluorescent proteins fused to SrtA recognition motifs: GGG-mCherry and GFP-P0. Relative to eSrtA, SrtA mutant activity remained unchanged with GGG-mCherry and increased with GFP-P0 by up to 32% with mutant 4 (Fig. B-2). Given that SrtA activity stayed low at concentrations up to 548 μM, alternatives were explored. GGG and P0 peptides with biotin modification at the C-terminus (bGGG and bP0) were used. Substrate concentrations at 300-500 μM yielded higher signal in Q2 (FITC-positive, APC-positive) for both substrates. With bGGG, 100 μM was enough to saturate cells expressing SrtA WT and eSrtA, while activity was mostly abolished for the mutants (Fig. 4).

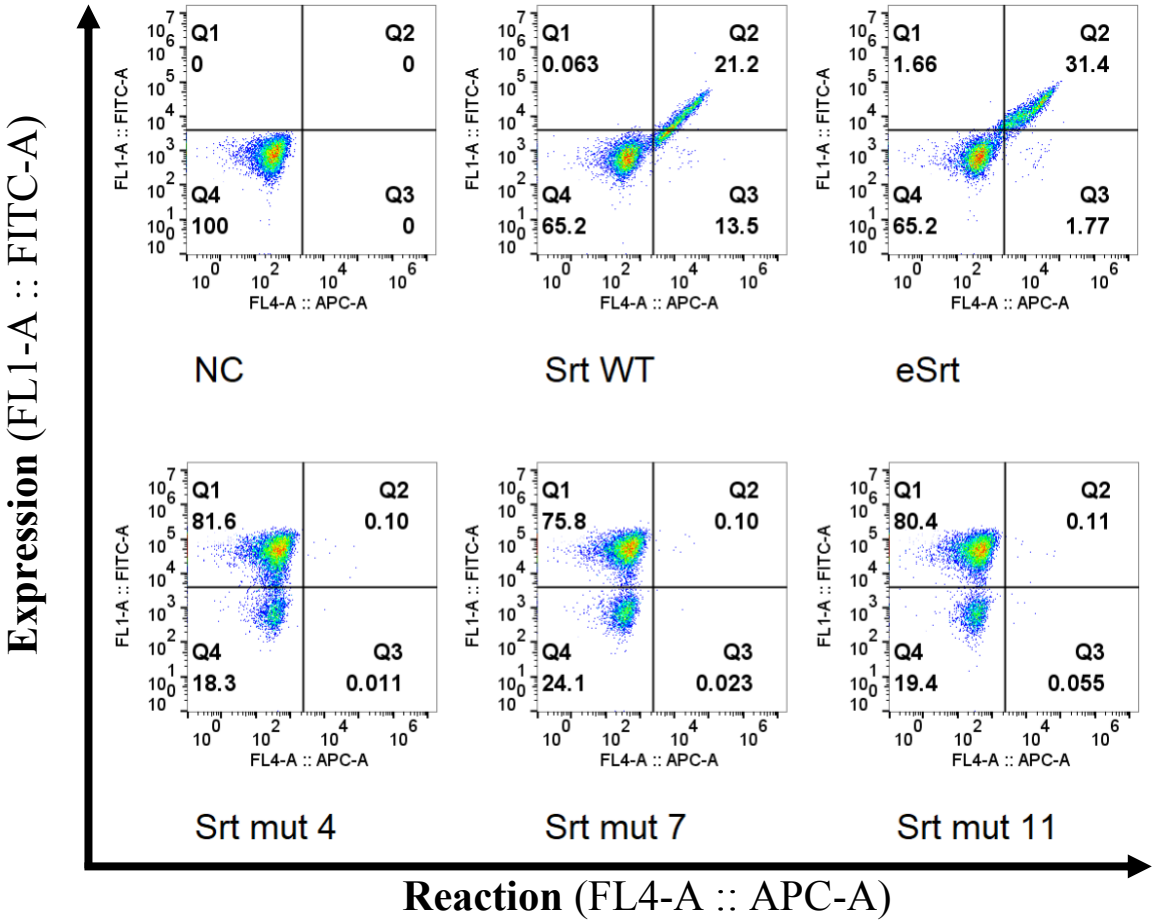


Fig. 4: SrtA WT and mutants on the yeast surface were incubated in 100 µM bGGG at 3h and 37 °C. Under these conditions, SrtA WT and eSrtA reached reaction completion, whereas substrate specificity for bGGG was abolished in mutants 4, 7 and 11.

Unlike bGGG, incubating at higher concentrations of bP0 simultaneously displayed signal in Q2 and Q3 (FITC-negative, APC-positive). The Q3 population was hypothesized to be non-specific binding between the P0 peptide and the yeast cell surface. When reaction temperature was lowered to 25 °C (Fig. 5A) and the mixture incubated for 24 h with 500 µM bP0, Q2 and Q3 signal were highest. Higher Q3 signal was observed in un-induced cells or SrtA variants with lower expression levels, like eSrtA. Decreasing bP0 concentration to 100 µM did not remove this Q3 signal (Fig. 5B). To further test this hypothesis and attempt to abolish non-specific binding, the BSA concentration in PBSA

was increased to 4% and yeast pellets were washed an additional two times for a total of five rounds (Fig. 6A). Increasing to 450 µM bP0 along with extensive washing led to higher signal in Q2 and Q3, indicating a concentration-dependent tendency of the bP0 peptide to non-specifically interact with the yeast surface (Fig. 6B). Though higher concentrations of bP0 increased Q2 signal across all SrtA variants, inactive SrtA (Inact SrtA) carrying the C184G mutation had higher signal than mutants 4, 7 and 11.

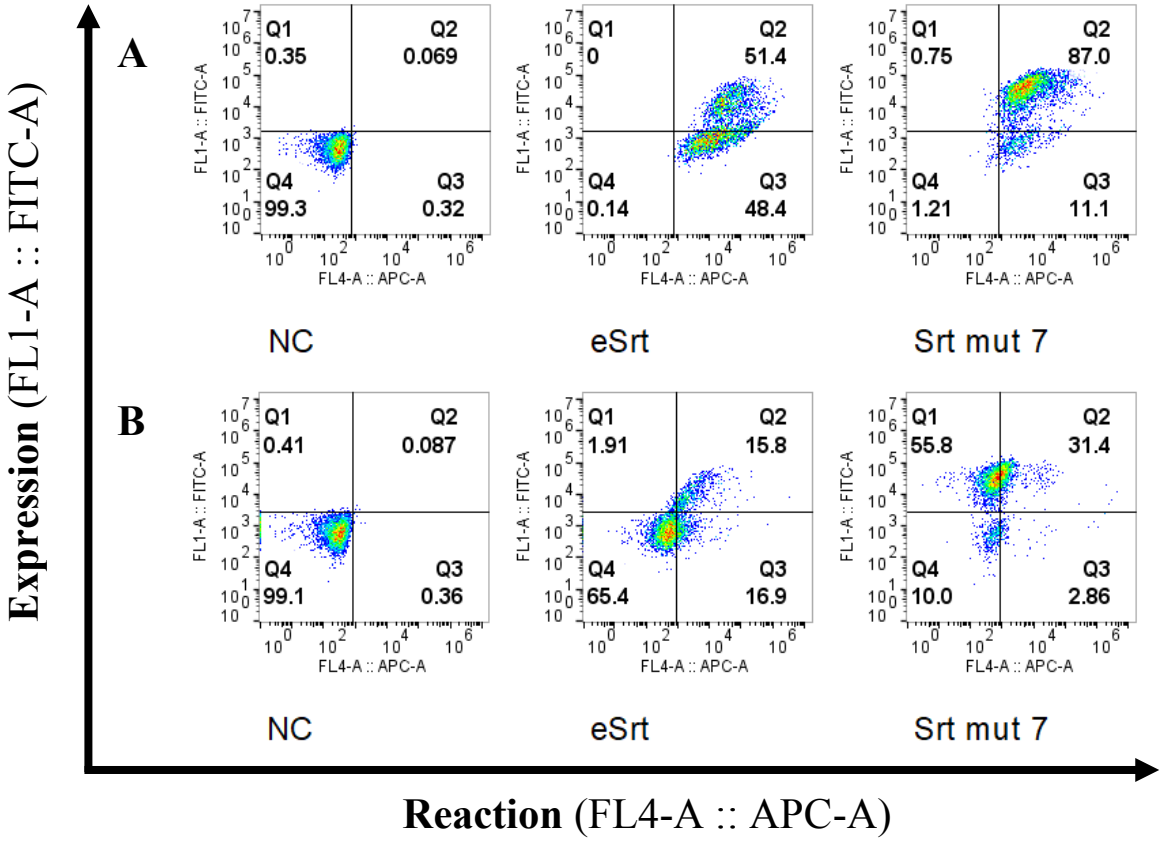


Fig. 5: The effects of different sets of reaction conditions for bP0 incubation are compared. Samples were incubated with **(A)** 500 µM bP0 for 24 h at 25 °C or **(B)** 100 µM bP0 for 16 h at 37 °C.

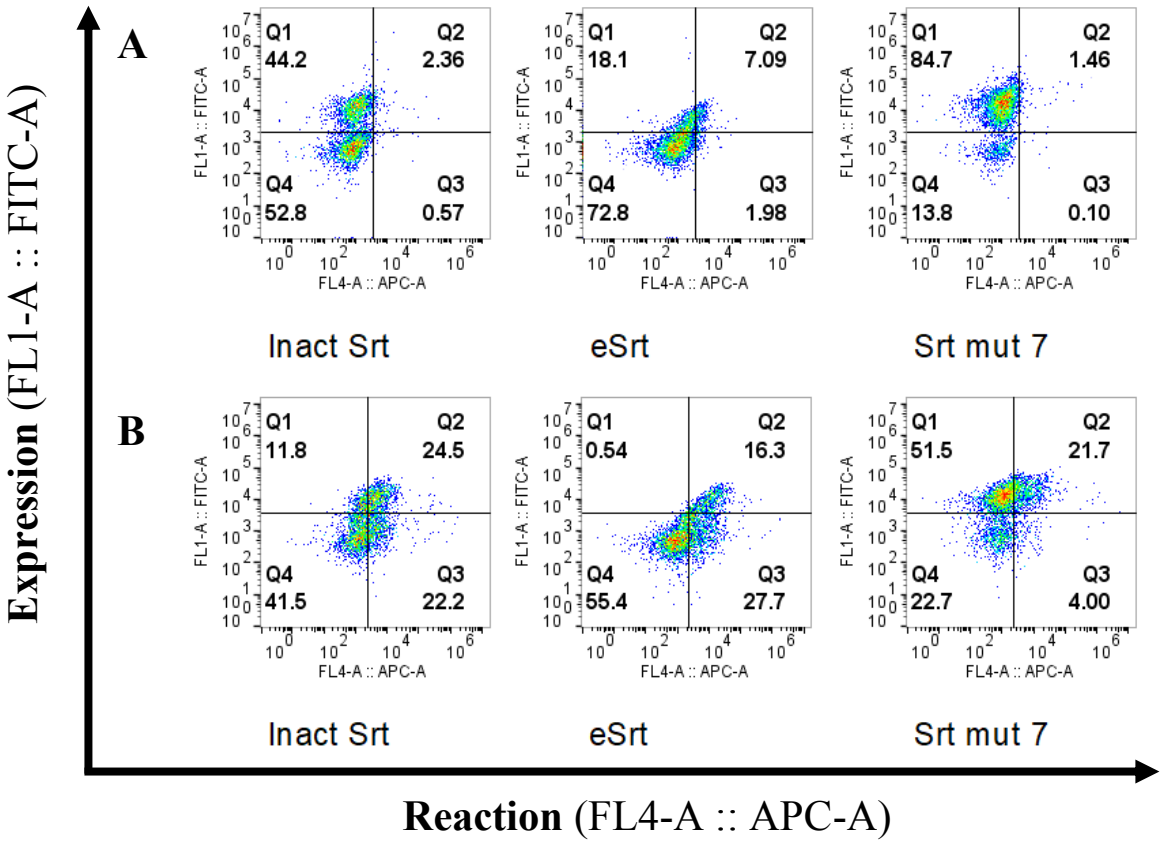


Fig. 6: The effects of longer washing cycles on bP0 reaction were analyzed with (A) 100 µM bP0 for 16h at 37 °C and **(D)** 450 µM bP0 for 16 h at 37 °C.

SrtA mutants demonstrated markedly different expression and reaction levels when compared to eSrtA. As described in the methods, median fluorescence levels were extracted from flow cytometry data. Expression levels for mutants 4, 7, and 11 increased by over 100-fold (Fig. 7A). Mutants demonstrated lower reaction kinetics than eSrtA for both bGGG (Fig. 7B) and bP0 (Fig. 7C) substrates, though this reduction in yield was more apparent with bGGG. Consequently, the ratio of bP0 reaction to bGGG reaction increased across mutants by at least 4-fold (Fig. 7D).

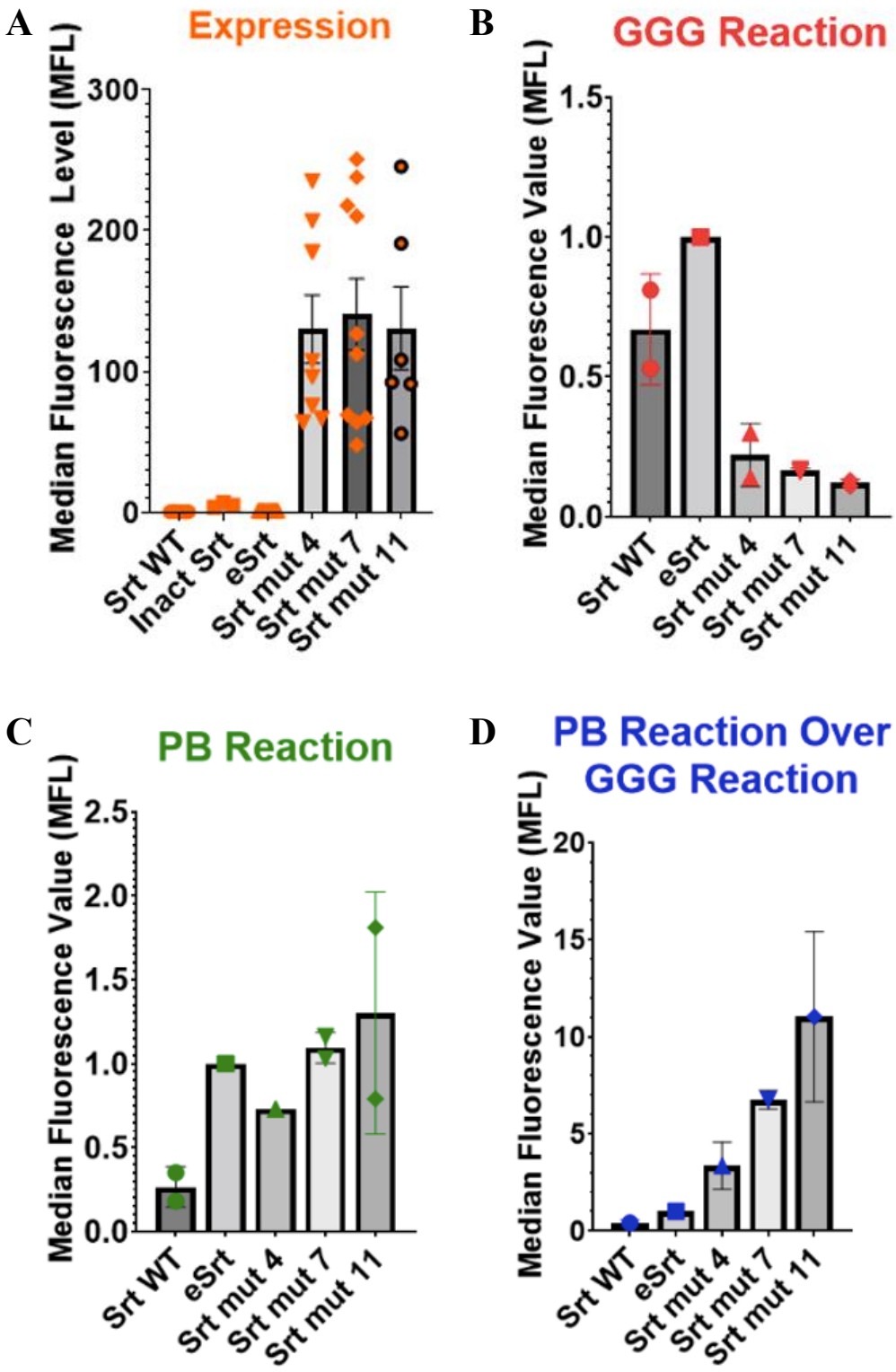


Fig. 7: The median fluorescence level data of SrtA mutants and controls are compared regarding (A) expression, (B) 100 μM bGGG reaction, (C) 100 μM bP0 reaction, and (D) ratio of bP0 to bGGG reaction.

In order to rationalize the sequence-structure-function relationship of SrtA mutants, the well-known crystal structure of wild-type *S. aureus* SrtA bound to LPETG was analyzed in PyMOL (Fig. 8). Of 22 mutations discovered across the three highest-performing mutants displayed on the yeast surface, 16 were found on flexible loops or oriented toward the substrate cleft.

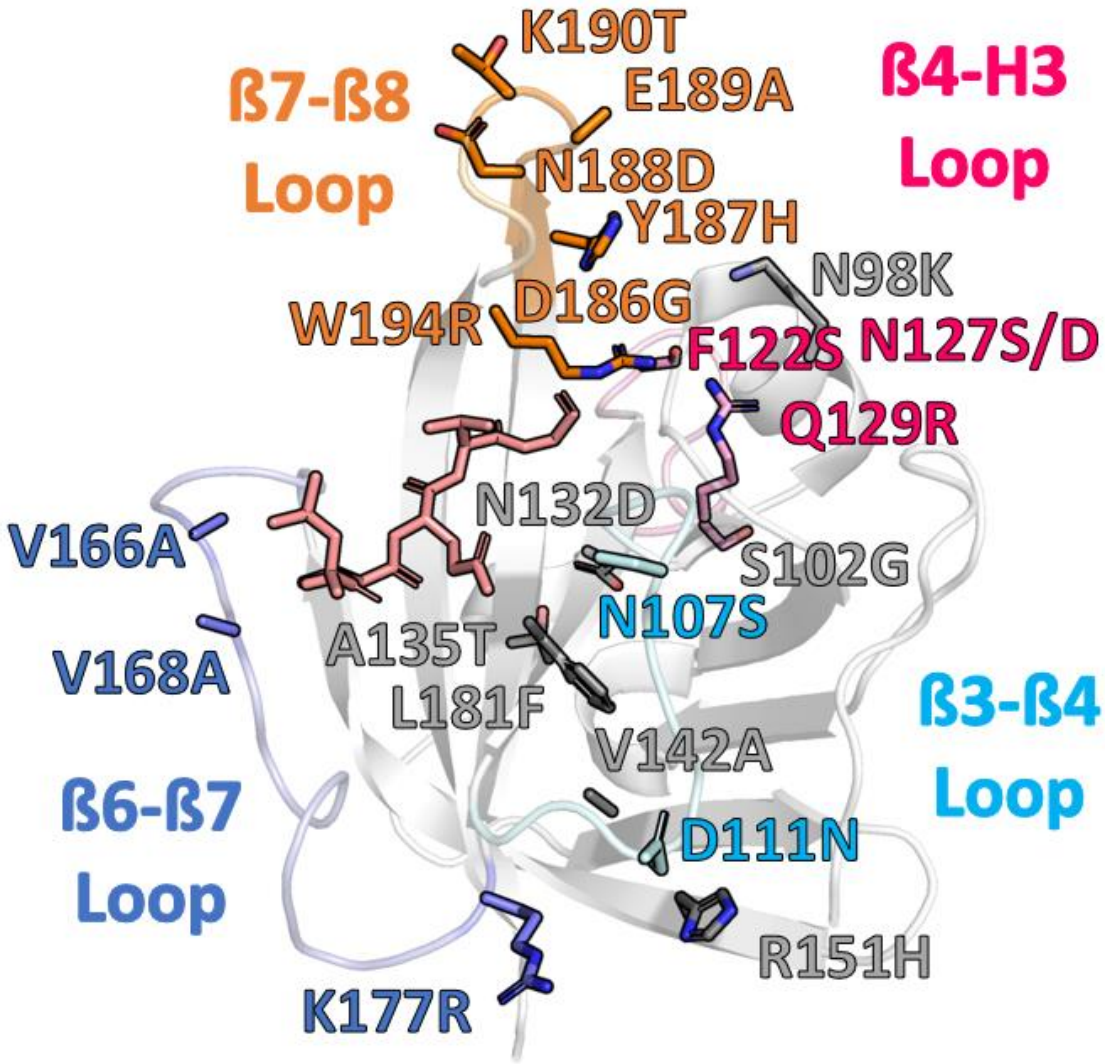


Fig. 8: The *S. aureus* SrtA-LPETG complex is depicted in white. Mutations from three high-performing mutants analyzed on the yeast surface were heavily localized to the ß3-ß4 loop (cyan), ß4-H3 loop (pink), ß6-ß7 loop (blue), and ß7-ß8 loop (orange). Other mutations not found in random loop regions are also shown (gray).

2.4: Conclusion

SrtA mutants, generated after three rounds of random mutagenesis and sorting, were analyzed. Though the mutants expressed at significantly higher levels than eSrtA, reaction with both the bGGG and bP0 peptide decreased. Substrate specificity for bGGG was almost completely abolished, but for bP0 was reduced by less than 50% relative to eSrtA. It remains possible that these mutations increased substrate selectivity for bP0 over bGGG across the three mutants. However, high Q2 signal in Inact SrtA suggests that bP0 may non-specifically bind the displayed protein or the yeast surface, since Inact SrtA is unable to form a thioacyl intermediate with bP0 or bGGG. As the mutants failed to show higher signal than Inact SrtA, the hypothesis that bP0 forms a thioacyl intermediate with SrtA mutants could not be validated.

To generate the alignment of P0 with LPET-bound SrtA in PyMOL, the orientation of GGG in the substrate cleft was first considered. Previous works had successfully shown GGG bound to *S. aureus* sortase B[30], which became the rationale for the orientation of GGG with LPET-SrtA shown here (Fig. D-1). Notably, several mutations were found in the ß4-H3 loop, leading to the hypothesis that interaction with this loop is key to P0 reaction. In order to simultaneously form the same interactions with other loops implicated in GGG binding, the only possible orientation for P0 was to be tightly packed between the ß7-ß8 and ß2-H1 loops. The direction of P0 was also considered: the C-terminus was oriented toward LPETG and the substrate cleft so that Y8 in the P0 sequence was buried in the hydrophobic core, newly formed by the S102G mutation (Fig. 9A). The P0 residue thought to participate in formation of the thioacyl intermediate, K10, was successfully positioned 3.5 Å away from the catalytic C184 residue in SrtA (Fig. D-2). This orientation

also explained the preference of most mutations toward smaller or less hydrophobic amino acids, such as F122S (Fig. 9B), V166A (Fig. 9C), and W194R (Fig. 9D). F122 is additionally known to play a role in calcium binding[31], whereas Y187H and W194R may play a role in protonation to resolve the thioacyl intermediate. Mutations in the core or away from the catalytic site may be responsible for improved stability and expression, such as L181F.

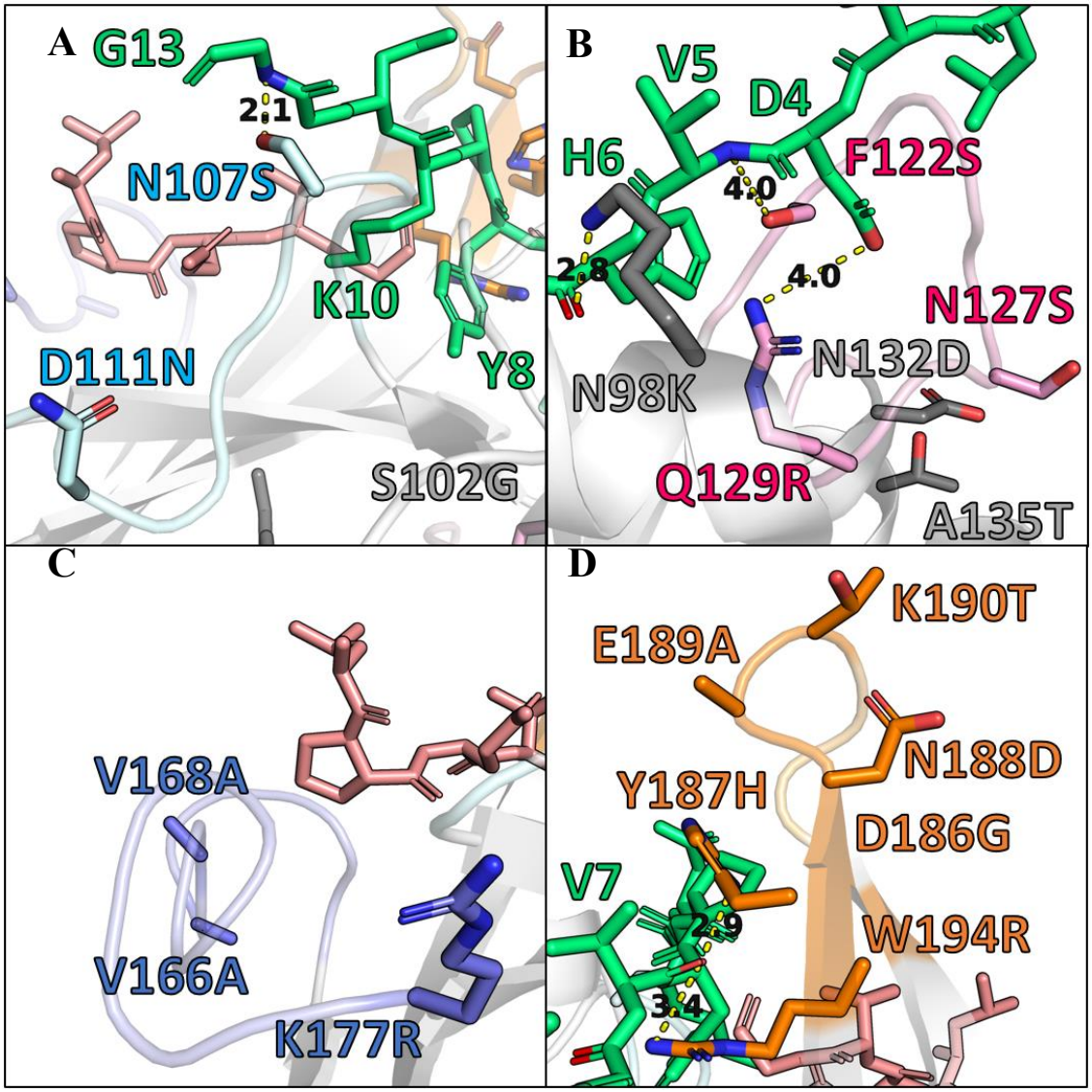


Fig. 9: Most SrtA mutations were found close to the substrate cleft or in key loop regions, such as **(A)** the ß3-ß4 loop in cyan, **(B)** the ß4-H3 loop in pink, **(C)** the ß6-ß7 loop in blue, or the **(D)** ß7-ß8 loop in orange. LPETG and P0 are depicted as red and green, respectively.

The major objective of this work was to engineer *S. aureus* SrtA to accept P0, derived from a common substrate accepted by *C. diphtheria* SrtA, as the nucleophile which resolved the LPET-SrtA thioacyl intermediate. Previously, it was discovered that *C. diphtheria* SrtA substrate recognition was largely influenced by the ß7-ß8 loop. Perhaps unsurprisingly, the ß7-ß8 loops in mutants 4 and 7 were significantly mutated. Several mutations matched residues in the ß7-ß8 loop of *C. diphtheria* SrtA, including Y186H, E189A, and K190T (Fig. 10A). Though the order of these residues was not anticipated, it remained likely that these mutations were preferential for P0 recognition.

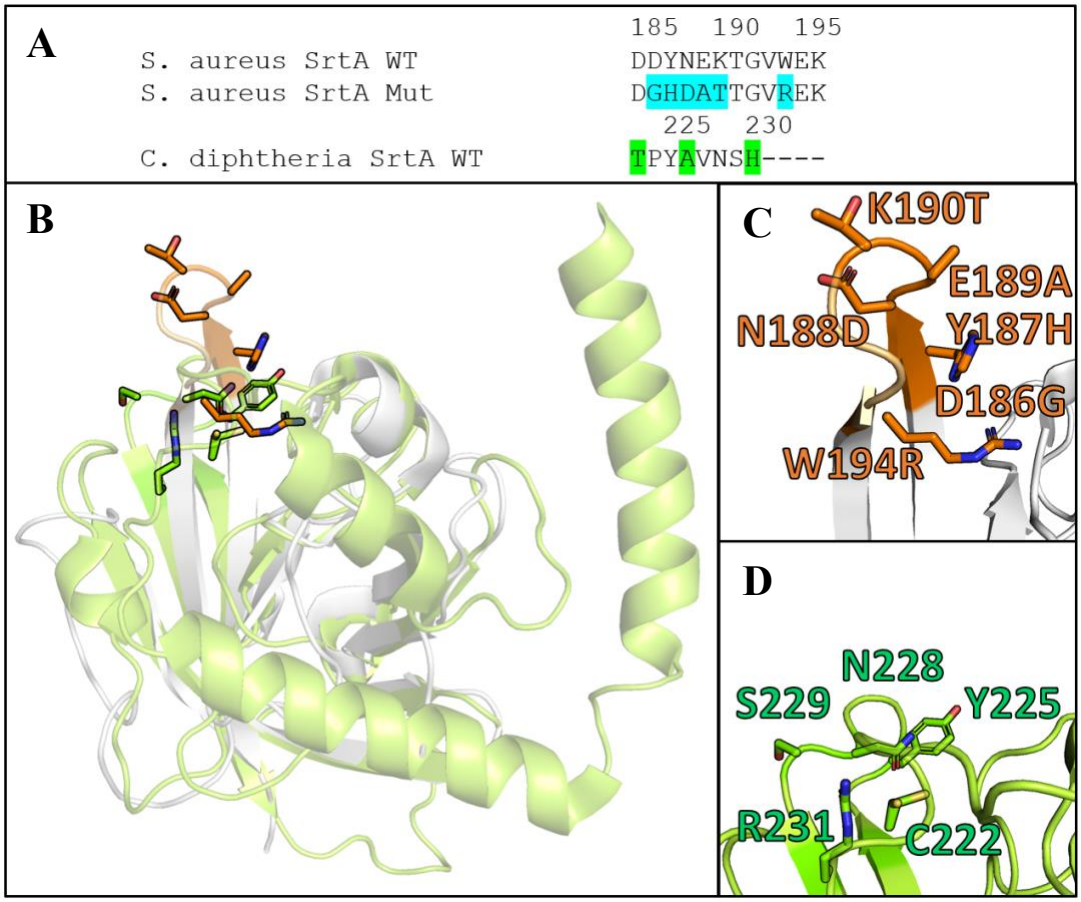


Fig. 10: The ß7-ß8 loops of *S. aureus* (orange) and *C. diphtheria* SrtA (green) are compared. **(A)** *S. aureus* and *C. diphtheria* SrtA ß7-ß8 sequences are overlapped, with emphasis on mutations to *S. aureus* SrtA (blue) and residues matching those mutations in *C. diphtheria* SrtA (green). SrtA structures were **(B)** superimposed using PyMOL, focusing on the ß7-ß8 loops for **(C)** *S. aureus* SrtA and **(D)** *C. diphtheria* SrtA.

CHAPTER 3

SOLUBLE SORTASE-MEDIATED LIGATION AND PURIFICATION USING ELPS

3.1: Introduction

Processes for the soluble expression and purification of fusion proteins often contain several bottlenecks at laboratory and industry-scale. Many fusion proteins for purification purposes are expressed with relative ease and utilize affinity chromatography for one-step purification. However, chromatography resins are expensive, making scale-up complex and costly. Overcoming the limitations of genetic fusion and affinity chromatography with a simpler and cheaper purification process could enable high-potential fusion proteins, with applications in drug delivery, biosensing, and protein assemblies, to see commercial use[32].

One promising purification alternative involves fusion to elastin-like polypeptides (ELPs) and subsequent non-chromatographic purification of the fusion product. ELPs are oligomeric repeats of the pentapeptide VPGXG, where X is any residue excluding proline. ELPs have an inverse transition temperature (T_t) which determines their solubility: ELPs are soluble below their T_t but otherwise aggregate and form precipitates when T_t is exceeded. This reversible aggregation temperature can be altered through mild perturbations to solution pH or ionic strength, thus enabling temperatures to be shifted toward physiological conditions. Guest residue composition and number of repeats affect T_t through altering the hydrophobicity of the protein. This unique property enables both free ELPs and ELP fusion proteins to undergo several cycles of reversible precipitation and resolubilization known as inverse transition cycling (ITC), where subsequent rounds of

ITC reduce impurities at the expense of lower protein yields. Only basic laboratory equipment is required, omitting the need for purchasing columns or resins. In previous works, ELP fusion and ITC were used to purify cowpea chlorotic mottle virus (CCMV) and E2 nanoparticles assembled through SrtA-mediated ligation[13,14].

Where genetic fusion and expression in a cellular host fail to yield ELP-tagged target proteins for ITC, SrtA-mediated ligation offers a flexible alternative. In this work, ELP 60-$(EV_4A_2G_2E)_6$ is expressed with a C-terminal LPETG tag for SrtA-mediated ligation[33]. ELP-LPETG is fused to GGG-mCherry or GFP-P0 and products are purified via ITC (Fig. 11). This method more reliably quantifies rate of SrtA product formation rather than rate of hydrolysis, which is often measured by conventional SrtA assays and kits. In most cases, ELP fusion to the C-terminus results in higher folding and activity of the fusion protein[34]. As ELP-LPETG is fused to the N-terminus of GGG-mCherry, GFP-P0 is also tested. SDS-PAGE is used to evaluate yields, where coomassie blue staining detects all proteins with weak binding to free ELP-LPETG, and copper staining is used to more reliably detect free ELP-LPETG and ELP protein fusions[35,36]. Both stains are compared to best visualize free ELP-LPETG and ELP fusion products. The apparent molecular weight of ELPs corresponding to SDS-PAGE can be 20% larger than the theoretical molecular weight[37]. Lastly, kinetic parameters of SrtA mutants are determined through cleavage of a fluorogenic peptide. Nucleophile is added to favor the rate of transpeptidation over hydrolysis, and better evaluate substrate specificity for GGG and P0[38]. Reactions are also performed in the presence of a well-characterized cone snail venom inhibitor of wild-type *S. aureus* SrtA, M2-conotoxin[39]. Biotinylated M2-conotoxin (bMC) competitively inhibits

SrtA, and was used to assess whether the P0 transpeptidation mechanism mimics product formation with GGG.

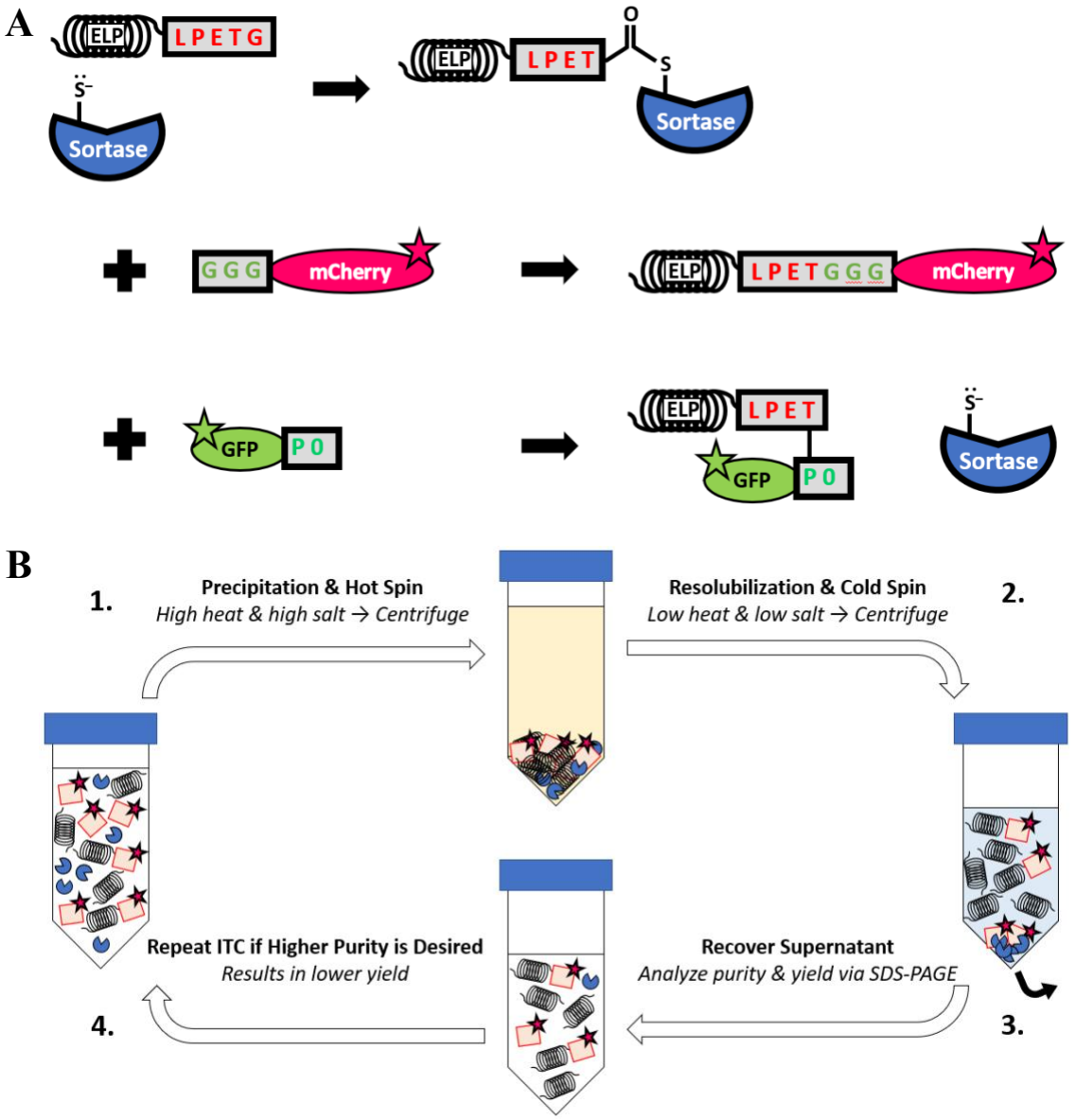


Fig. 11: (A) Soluble SrtA forms a thioacyl intermediate with ELP-LPETG, which can undergo transpeptidation with either GGG-mCherry or GFP-P0. (B) Free ELP-LPETG and ELP fusion products can be purified through ITC. 1) Proteins are precipitated through high salt concentration at high temperature. 2) ELPs and ELP fusion products are resolubilized in cold low-salt buffer, and the pellet containing the impurities is discarded. 3) Purity and yield are analyzed via SDS-PAGE. ITC can be repeated as desired for higher purity, but a lower proportion of ELP-containing products will be recovered each time.

3.2: Experimental Methods

Strains and Plasmids

All constructs used the pET plasmid for bacterial expression and were sequenced using the T7 forward primer. For DNA extraction and sequencing, plasmids were transformed into NEB5α competent *E. coli*. For soluble expression in bacteria, plasmids were transformed into BL21(DE3) chemically competent *E. coli* (Thermo Scientific). The pET28a-ELP-LPETG plasmid was received through Addgene.

Soluble Protein Expression and Purification

Cultures containing 10 mL LB broth supplemented with 100 mg/mL kanamycin were inoculated with transformants on LB kan plates and grown overnight at 37 °C. 500 mL cultures of LB kan broth were inoculated with overnight cultures to a final optical density (OD) between 0.05-0.1. When expressing wild-type SrtA, cultures were shaken for 24 h at 37 °C and 250 rpm before centrifuging. When expressing SrtA mutants or fluorescent proteins, cells were induced with 0.5 mM IPTG upon reaching OD 0.4-0.6. Cultures induced with IPTG were grown for 3-4 hours before centrifugation. All cell pellets were resuspended in 20 mL of TBS and 10 µL of DNase I. Cells were sonicated, pelleted at 13,000 x g and passed through a 0.45 µm filter. SrtA WT was loaded onto a Ni-NTA gravity column. The flowthrough was applied to the column a second time, and proteins were visibly eluted with TBS containing 250 mM imidazole. If needed, proteins were concentrated using 15 mL Amicon centrifugal filter units with 10 kDa molecular weight cutoff. SrtA mutants were not purified. For SrtA mutants expressed as insoluble inclusion bodies, lysates were dissolved in TBS containing 6 M urea and dialyzed in TBS overnight.

Proteins not purified were quantified through SDS-PAGE, using ImageJ to estimate protein concentrations, given a known concentration of SrtA WT.

Elastin-like Polypeptide (ELP) Purification

Sonicated cell lysates containing ELP-LPETG proceeded to inverse transition cycling (ITC). 3 M ammonium sulfate stock was added to cell lysates to a final concentration of 0.5 M, and samples were incubated at 37 °C for 30 min. For the "hot" spin, samples were pelleted at 16,000 x g for 10 min at room temperature. After discarding the supernatant, 1.5 mL of ice-cold 50 mM Tris-HCl (pH 7.6) and 10 mM tris(2-carboxyethyl)phosphine (TCEP) was added to resolubilize the pellet. Upon turning translucent, the pellet was dissolved by pipetting up and down. Finally, for the "cold spin," samples were centrifuged at 16,000 x g at 4 °C to remove any insoluble contaminants. The supernatant was either stored below -80 °C or kept on ice for future analysis.

SrtA Reaction with GFP-LPETG and Fluorescent Nucleophilic Substrates

Reaction mixtures of 50 µL containing 220 µM SrtA WT, 10-100 µM GFP-LPETG, 10-100 µM GGG-mCherry, and TBS (pH 8.0) with 10 mM $CaCl_2$ were incubated for 3 h at 37 °C. Aliquots were taken for SDS-PAGE.

SrtA Reaction with ELP-LPETG and Fluorescent Nucleophilic Substrates

Reaction mixtures of 150 µL containing 20 µM SrtA WT or mutants, 20 µM ELP-LPETG, 80 µM GGG-mCherry or GFP-P0, and TBS (pH 8.0) with 10 mM $CaCl_2$ were incubated for 4 h at 42 °C. An aliquot was taken for SDS-PAGE. Samples were purified

via ITC, as previously described, into a new volume of 100 µL. Samples were then analyzed via the fluorescence microplate assay and SDS-PAGE.

SDS-PAGE Gel Analysis and Staining

ELP and SrtA reaction samples were visualized via SDS-PAGE using two different stains. Coomassie Brilliant Blue G-250 was used to visualize most proteins. Gels were first washed in deionized water before incubation in staining solution and destaining solution for 1 h each. As ELPs lacked a sufficient number of amino acids to be stained with Coomassie Brilliant Blue, ELPs and SrtA reaction mixtures were also visualized using copper (0.5 M $CuCl_2$) stain. After washing in DI water and incubating in copper stain for 10 min, gels were ready to be imaged.

Fluorescence Microplate Assay

ITC-purified ELP and SrtA reaction samples (90 µL) were added to black 96-well plates and fluorescence was detected on the Synergy HTX multi-mode plate reader with gain set to 100 V. GFP-P0 and GGG-mCherry were detected with excitation and emission at 485/20 and 528/20, respectively.

SrtA Kinetic Assays Using a Quenched Fluorescent Peptide

Kinetic parameters for SrtA mutants were determined through cleavage of the Dabcyl-QALPETGEE-Edans peptide (Anaspec). 15-250 µM of fluorogenic substrate was incubated in the presence of 50 µM SrtA WT or mutants. Tests were performed in the presence and absence of 500 µM bGGG peptide, 5 µM bP0 peptide, or 5 µM bMC.

3.3: Results & Discussion

SrtA WT and mutants were sequenced and expressed in soluble form in *E. coli* (Table 2). Six mutants were analyzed, including four new mutants and mutants 4 and 7, previously analyzed on the yeast surface. Soluble expression levels of mutants 2 and 7 were lower compared to other mutants, possibly since these mutants were not as stable when expressed using the limited cellular machinery of *E. coli* (Fig. C-1) Sequencing revealed mutations were mostly found in secondary structure elements like ß-sheets, while other mutations favored random loops and regions known for substrate interaction (Table 2).

Table 2: Key mutations in SrtA mutants expressed in soluble form in *E. coli*.

Region	Mut 1	Mut 2	Mut 8	Mut 9
N-Terminus				**K67R**
ß2 Sheet		**E85K**	**I83T**	
H2 Helix	**E95K / R99K**			
ß3-ß4 Loop	**E106K**			
ß4-H3 Loop	**N127S**	**F122L / N127D**		
ß5 Sheet		**K145E**		
ß6 Sheet				**M155V / D160S**
ß6-ß7 Loop	**D170G / E171G / K177R**			
ß7 Sheet				**Q178R**
ß8 Sheet		**T203A**		**K198R / F200S**

Before testing the activity of SrtA mutants, a standardized ELP purification protocol was developed to circumvent the need to separately purify SrtA mutants and fusion products with expensive chromatography resins (Fig. C-2). SDS-PAGE gels were copper-stained to best visualize free ELP-LPETG, though ELP fusion products could be visualized with coomassie blue (Fig. C-3). Secondly, substrate and SrtA WT concentrations were optimized for the ligation of GGG-mCherry to GFP-LPETG and ELP-LPETG. Though GFP-LPETG tests yielded GFP-LPETGGG-mCherry fusion product, large amounts of impurities made ELP-LPETG the more feasible substrate for reliable

quantification (Fig. C-4). The ELP SrtA reaction was then optimized by varying concentrations of GGG-mCherry (Fig. C-5) and SrtA WT (Fig. C-6).

To determine the impact of different mutations on transpeptidation activity, SrtA WT and mutants were used to ligate ELP-LPETG to either GGG-mCherry or GFP-P0. After one round of ITC, yield was determined via fluorescent detection and SDS-PAGE. Relative to SrtA WT, fluorescence increased for both substrates across mutants 1, 4 and 8, though mutants preferred GFP-P0 with mutant 8 showing the highest selectivity (Fig. 12).

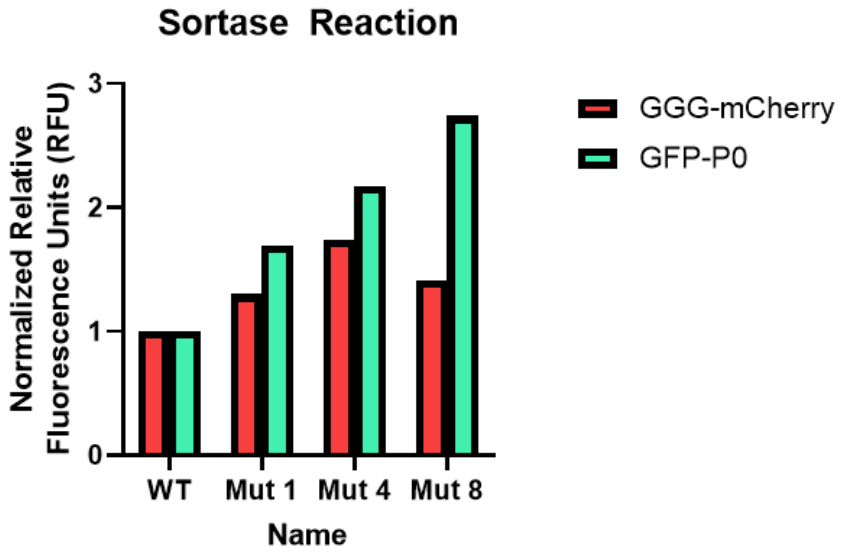


Fig. 12: SrtA reaction is compared across WT and three mutants for GGG-mCherry and GFP-P0 by analyzing the fluorescent signal of ITC-purified fusion products.

Fusion product yield was simultaneously measured via SDS-PAGE before and after one round of ITC (Fig. 13). If either substrate or enzyme was absent, no fusion product yield was observed. SrtA WT and mutant 8 retained GGG specificity, whereas GGG transpeptidation activity was lost in mutants 1 and 4. The band corresponding to ELP-LPETGGG-mCherry fusion product is expected around 53 kDa.

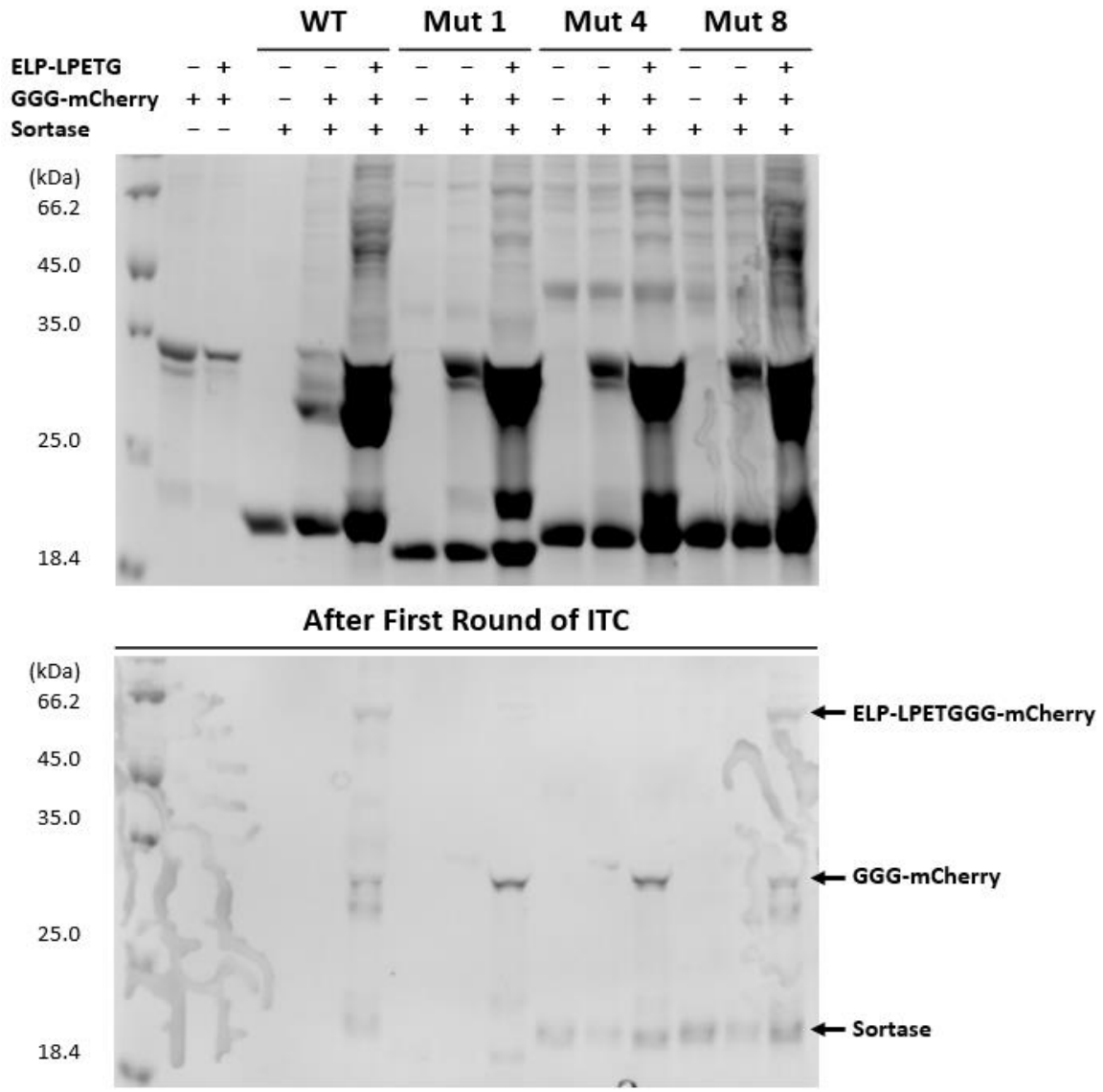


Fig. 13: SrtA reaction mixtures were analyzed via SDS-PAGE before and after one round of ITC. SrtA WT and mutants were incubated in the presence of ELP-LPETG and GGG-mCherry with an expected fusion product of ELP-LPETGGG-mCherry at approximately 53 kDa.

Similarly, GFP-P0 fusion products were analyzed via SDS-PAGE before and after one round of ITC (Fig. 14), with similar controls to ensure bands representing impurities are not mistaken for ELP-LPETG-P0-GFP fusion product. The fusion product band is expected around 55 kDa. Faint bands can be observed for SrtA WT and mutant 1, and

possibly for mutants 4 and 8, though this remains inconclusive due to improper staining of the gel containing ITC-purified samples.

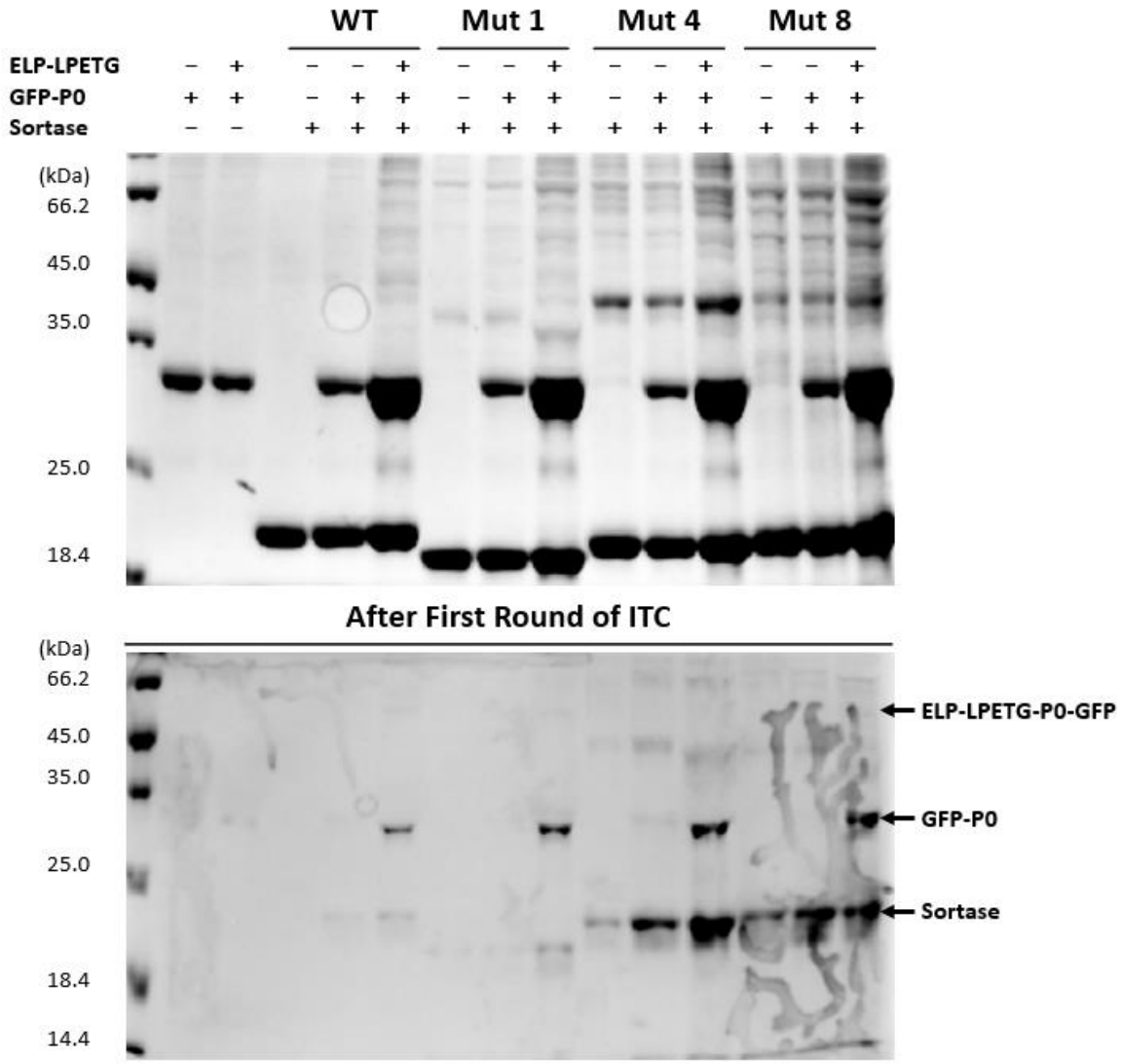


Fig. 14: SrtA reaction mixtures were analyzed via SDS-PAGE before and after one round of ITC. SrtA WT and mutants were incubated in the presence of ELP-LPETG and GFP-P0 with an expected fusion product of ELP-LPET-P0-GFP at approximately 55 kDa.

As with mutants characterized on the yeast surface, mutations found across mutants 1, 2, 8, and 9 were analyzed in PyMOL (Fig. 15). Notably, mutations were mostly found in secondary structure elements, like ß-sheets, or constrained to the ß6-ß7 loop.

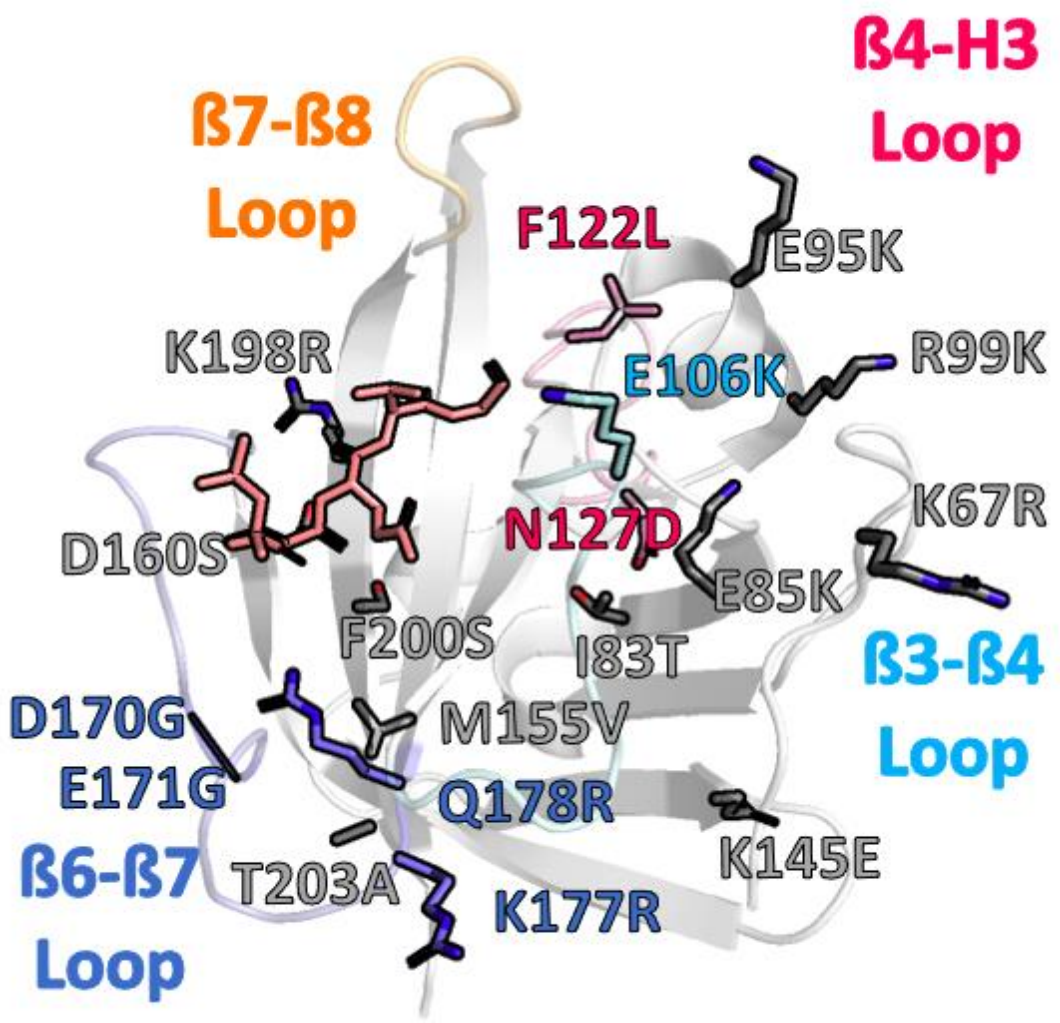


Fig. 15: The *S. aureus* SrtA-LPETG complex is depicted in white. Mutations from mutants 1, 2, 8, and 9 expressed in *E. coli* were mostly constrained to ß-sheets or the ß6-ß7 loop (blue). The ß3-ß4 loop (cyan), ß4-H3 loop (pink), and ß7-ß8 loop (orange) are also shown. Other mutations not found in random loop regions are depicted in gray.

3.4: Conclusion

SrtA mutants, with truncated N-termini omitting the transmembrane anchoring domain, were expressed in soluble form in *E. coli*. Using fluorescence detection and SDS-PAGE, the effect of these mutations was analyzed pertaining to substrate specificity for GGG-mCherry and P0-GFP. Though SrtA WT maintained specificity for both GGG and P0, mutants 1 and 8 showed higher selectivity toward P0 and GGG, respectively.

Further mutational is required to gain insight into effects on substrate specificity. The only mutation present in mutant 8, I83T, had negligible effect on GGG reaction, though it is unclear whether P0 specificity was conserved. I83T is buried in the hydrophobic protein core, and might favor protein stability. Notably, of the four mutants, mutant 1 shows the most mutations in the ß3-ß4 loop and ß4-H3 loops. D170G, E171G (Fig. 16A) and E106K (Fig. 16B) all represent shifts to smaller or more linear amino acids, possibly removing steric hindrance that would otherwise block the bulky P0 substrate. F122L (Fig. 16C) and Q178R found on mutants 2 and 9, respectively, might provide less steric hindrance or ionic interactions that also facilitate P0 binding. E171 and F122 are known to participate in calcium binding, though these mutations might not be enough to completely disrupt these interactions.

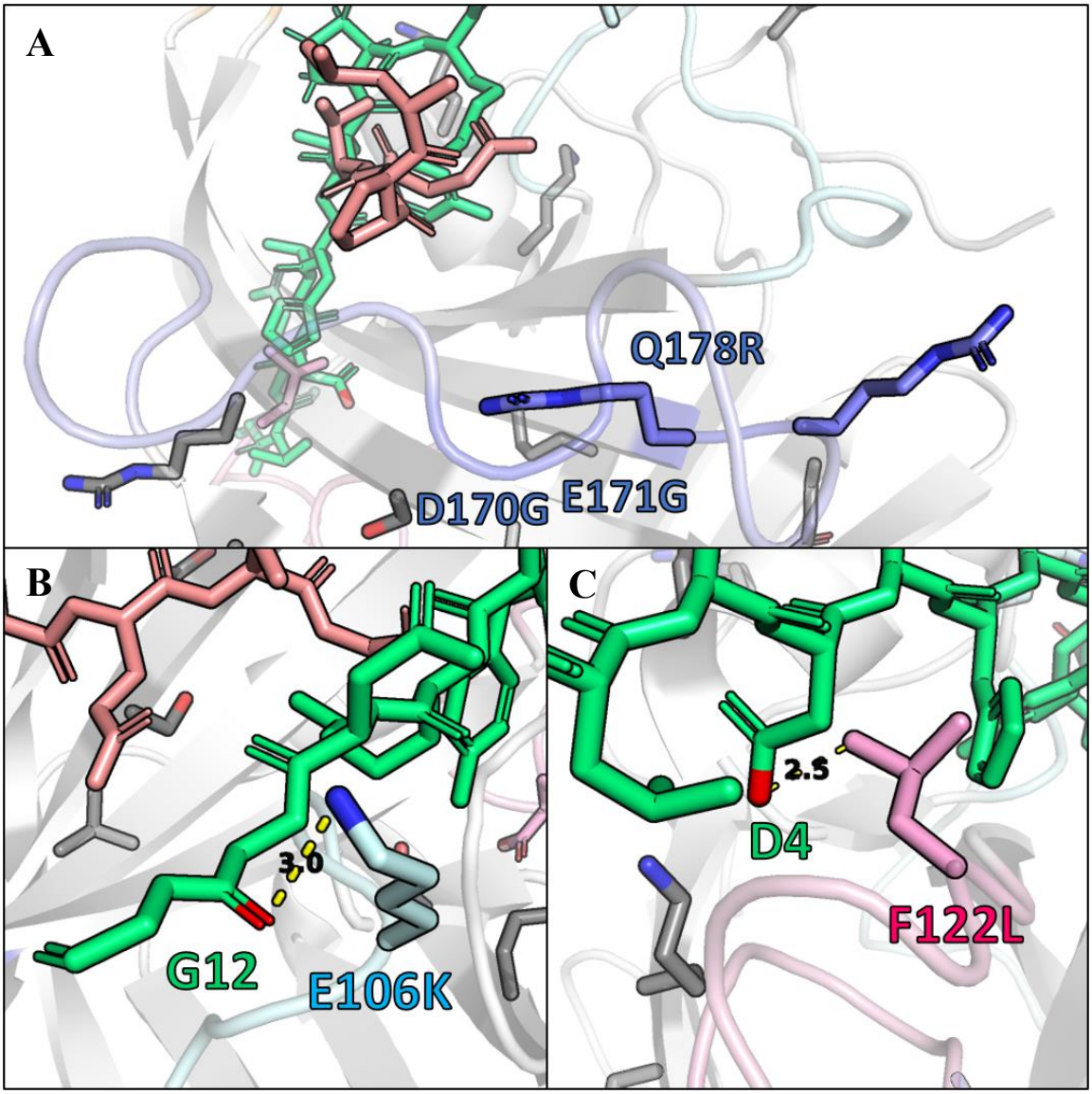


Fig. 16: **(A)** Mutations to smaller or basic amino acids may provide a more advantageous binding pocket for LPETG or P0. **(B)** K106, a more linear amino acid, could accommodate the bulky P0 substrate, or form desirable ionic interactions. **(C)** L122, a smaller and aliphatic amino acid, might form hydrogen bonds or hydrophobic interactions with P0.

CHAPTER 4

RECOMMENDATIONS & FUTURE WORK

Previous SrtA engineering efforts have arisen from either directed evolution, rational design, or a combination of both approaches. In this study, analysis of SrtA mutants generated via directed evolution revealed many key insights into the SrtA sequence-structure-function relationship. However, to further increase SrtA P0 activity or specificity, several rounds of random mutagenesis and sorting may prove too difficult or impossible: certain mutations, or combinations thereof, that would theoretically maximize P0 activity may not be thermodynamically stable. Therefore, to build on this work, a combination of directed evolution and targeted rational design approaches may be advantageous. Previous works have highlighted the importance of the ß7-ß8 loop in *C. diphtheriae* SrtA for recognition of the P0 motif, and the ability to alter *S. aureus* SrtA substrate specificity by directly mutating the ß7-ß8 loop. By keeping mutations that improve stability and expression while targeting the ß7-ß8 loop for P0 activity, highly stable and selective SrtA mutants could be discovered.

This work generally showcases how directed evolution can be used to improve SrtA stability and broaden or shift substrate specificity. Overall, mutants with mutations that favor regions of flexibility or substrate interaction had more extreme effects on activity and specificity. To better elucidate the sequence-structure-function relationship of SrtA, all mutants should be expressed in both soluble and on the yeast surface. Secondly, specific point mutations should be studied, particularly to identify whether mutations have an individual or combination of effects on stability, activity, or substrate specificity.

The P0 peptide shows promise to expand the SrtA-mediated ligation toolbox, though more evidence is needed to verify that engineered SrtA mutants are catalyzing the transpeptidation of LPETG and P0 instead of non-specifically binding to P0. On the yeast surface, increasing P0 concentration leads to a higher fluorescent signal attributed to P0. This higher signal appears to be independent of whether HA, fused to Aga2p, is present, suggesting that un-induced yeast cells are binding to P0. Furthermore, Inact SrtA, carrying the C184G mutation, showed similar spread on dual scatter plots to SrtA mutants. More analysis and optimization of the P0 reaction is needed to confirm whether transpeptidation, non-specific binding, or both interactions are present. If non-specific binding is found to be an issue, the original directed evolution protocol can be re-visited and correct to select against mutants that promote undesirable side reactions.

Soluble SrtA-mediated ligation with fluorescent proteins provides a quantitative and facile approach to characterizing sortase mutants, though additional tests would make this method more comprehensive. Currently, tests that incorporate excess nucleophile or inhibitor are needed to ensure that the transpeptidation reaction is being measured rather than rate of hydrolysis, and to report the inhibition constant, both of which are critical measures in characterizing SrtA mutants. Additionally, studies of the kinetic parameters such as the Michaelis constant have not been included, which could be determined through activity assays using a fluorogenic cleavable peptide with a sortase recognition motif.

REFERENCES

1 Raeeszadeh-Sarmazdeh, M., Hartzell, E., Price, J. V. & Chen, W. Protein nanoparticles as multifunctional biocatalysts and health assessment sensors. *Current opinion in chemical engineering* **13**, 109-118 (2016).

2 Kim, H. *et al.* Bioengineering strategies to generate artificial protein complexes. *Biotechnology and bioengineering* **112**, 1495-1505 (2015).

3 Diaz, D., Care, A. & Sunna, A. Bioengineering strategies for protein-based nanoparticles. *Genes* **9**, 370 (2018).

4 Bolt, A. J., Do, L. D. & Raeeszadeh-Sarmazdeh, M. Bacterial Expression and Purification of Human Matrix Metalloproteinase-3 using Affinity Chromatography. *JoVE*, e63263, doi:doi:10.3791/63263 (2022).

5 Levary, D. A., Parthasarathy, R., Boder, E. T. & Ackerman, M. E. Protein-protein fusion catalyzed by sortase A. *PLoS One* **6**, e18342, doi:10.1371/journal.pone.0018342 (2011).

6 Kruger, R. G. *et al.* Analysis of the substrate specificity of the Staphylococcus aureus sortase transpeptidase SrtA. *Biochemistry* **43**, 1541-1551 (2004).

7 Hendrickx, A. P., Budzik, J. M., Oh, S. Y. & Schneewind, O. Architects at the bacterial surface - sortases and the assembly of pili with isopeptide bonds. *Nat Rev Microbiol* **9**, 166-176, doi:10.1038/nrmicro2520 (2011).

8 Parthasarathy, R., Subramanian, S. & Boder, E. T. Sortase A as a novel molecular "stapler" for sequence-specific protein conjugation. *Bioconjugate chemistry* **18**, 469-476 (2007).

9 Chen, I., Dorr, B. M. & Liu, D. R. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc Natl Acad Sci U S A* **108**, 11399-11404, doi:10.1073/pnas.1101046108 (2011).

10 Jeong, H. J., Abhiraman, G. C., Story, C. M., Ingram, J. R. & Dougan, S. K. Generation of Ca2+-independent sortase A mutants with enhanced activity for protein and cell surface labeling. *PLoS One* **12**, e0189068, doi:10.1371/journal.pone.0189068 (2017).

11 Dorr, B. M., Ham, H. O., An, C., Chaikof, E. L. & Liu, D. R. Reprogramming the specificity of sortase enzymes. *Proc Natl Acad Sci U S A* **111**, 13343-13348, doi:10.1073/pnas.1411179111 (2014).

12 Bentley, M. L., Gaweska, H., Kielec, J. M. & McCafferty, D. G. Engineering the substrate specificity of Staphylococcus aureus sortase a: the β6/β7 loop from SrtB confers npqtn recognition to SrtA. *Journal of biological chemistry* **282**, 6571-6581 (2007).

13 Chen, Q. *et al.* Sortase A-mediated multi-functionalization of protein nanoparticles. *Chemical Communications* **51**, 12107-12110, doi:10.1039/C5CC03769G (2015).

14 Schoonen, L., Nolte, R. J. & van Hest, J. C. Highly efficient enzyme encapsulation in a protein nanocage: towards enzyme catalysis in a cellular nanocompartment mimic. *Nanoscale* **8**, 14467-14472 (2016).

15 Tang, S., Xuan, B., Ye, X., Huang, Z. & Qian, Z. A modular vaccine development platform based on sortase-mediated site-specific tagging of antigens onto virus-like particles. *Scientific reports* **6**, 1-9 (2016).
16 Gau, E. *et al.* Sortase-mediated surface functionalization of stimuli-responsive microgels. *Biomacromolecules* **18**, 2789-2798 (2017).
17 Patterson, D. P., Prevelige, P. E. & Douglas, T. Nanoreactors by programmed enzyme encapsulation inside the capsid of the bacteriophage P22. *ACS nano* **6**, 5000-5009 (2012).
18 Wang, H. H., Altun, B., Nwe, K. & Tsourkas, A. Proximity-Based Sortase-Mediated Ligation. *Angewandte Chemie* **129**, 5433-5436 (2017).
19 Le, R. K., Raeeszadeh-Sarmazdeh, M., Boder, E. T. & Frymier, P. D. Sortase-mediated ligation of PsaE-modified photosystem I from Synechocystis sp. PCC 6803 to a conductive surface for enhanced photocurrent production on a gold electrode. *Langmuir* **31**, 1180-1188 (2015).
20 Raeeszadeh-Sarmazdeh, M., Parthasarathy, R. & Boder, E. T. Site-specific immobilization of protein layers on gold surfaces via orthogonal sortases. *Colloids and Surfaces B: Biointerfaces* **128**, 457-463 (2015).
21 Raeeszadeh-Sarmazdeh, M., Parthasarathy, R. & Boder, E. T. Fine-tuning sortase-mediated immobilization of protein layers on surfaces using sequential deprotection and coupling. *Biotechnology Progress* **33**, 824-831 (2017).
22 Van Doren, S. R. *et al.* Inactivation of N-TIMP-1 by N-terminal acetylation when expressed in bacteria. *Biopolymers: Original Research on Biomolecules* **89**, 960-968 (2008).
23 Witte, M. D. *et al.* Preparation of unnatural N-to-N and C-to-C protein fusions. *Proc Natl Acad Sci U S A* **109**, 11993-11998, doi:10.1073/pnas.1205427109 (2012).
24 Dasgupta, S., Samantaray, S., Sahal, D. & Roy, R. P. Isopeptide ligation catalyzed by quintessential sortase A: mechanistic cues from cyclic and branched oligomers of indolicidin. *J Biol Chem* **286**, 23996-24006, doi:10.1074/jbc.M111.247650 (2011).
25 Scott, J. R. & Zahner, D. Pili with strong attachments: Gram-positive bacteria do it differently. *Mol Microbiol* **62**, 320-330, doi:10.1111/j.1365-2958.2006.05279.x (2006).
26 Raeeszadeh-Sarmazdeh, M. Engineering Sortase A for generating site-specific protein 3D assemblies. (2014).
27 Zong, Y., Bice, T. W., Ton-That, H., Schneewind, O. & Narayana, S. V. Crystal structures of Staphylococcus aureus sortase A and its substrate complex. *J Biol Chem* **279**, 31383-31389, doi:10.1074/jbc.M401374200 (2004).
28 Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nature biotechnology* **15**, 553-557 (1997).
29 Boder, E. T. & Wittrup, K. D. in *Methods in enzymology* Vol. 328 430-444 (Elsevier, 2000).
30 Zong, Y., Mazmanian, S. K., Schneewind, O. & Narayana, S. V. L. The Structure of Sortase B, a Cysteine Transpeptidase that Tethers Surface Protein to the

Staphylococcus aureus Cell Wall. *Structure* **12**, 105-112, doi:https://doi.org/10.1016/j.str.2003.11.021 (2004).

31 Naik, M. T. *et al.* Staphylococcus aureus sortase A transpeptidase: calcium promotes sorting signal binding by altering the mobility and structure of an active site loop. *Journal of Biological Chemistry* **281**, 1817-1826 (2006).

32 Meyer, D. E. & Chilkoti, A. Purification of recombinant proteins by fusion with thermally-responsive polypeptides. *Nature biotechnology* **17**, 1112-1115 (1999).

33 Ott, W. *et al.* Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy. *ACS Nano* **11**, 6346-6354, doi:10.1021/acsnano.7b02694 (2017).

34 Christensen, T. *et al.* Fusion order controls expression level and activity of elastin-like polypeptide fusion proteins. *Protein Science* **18**, 1377-1387 (2009).

35 MacEwan, S. R., Hassouneh, W. & Chilkoti, A. Non-chromatographic Purification of Recombinant Elastin-like Polypeptides and their Fusions with Peptides and Proteins from Escherichia coli. *JoVE*, e51583, doi:doi:10.3791/51583 (2014).

36 Hassouneh, W., Christensen, T. & Chilkoti, A. Elastin-like polypeptides as a purification tag for recombinant proteins. *Current protocols in protein science* **61**, 6.11. 11-16.11. 16 (2010).

37 Meyer, D. E. & Chilkoti, A. Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system. *Biomacromolecules* **3**, 357-367 (2002).

38 Ilangovan, U., Ton-That, H., Iwahara, J., Schneewind, O. & Clubb, R. T. Structure of sortase, the transpeptidase that anchors proteins to the cell wall of Staphylococcus aureus. *Proc Natl Acad Sci U S A* **98**, 6056-6061, doi:10.1073/pnas.101064198 (2001).

39 Younis, S., Taj, S. & Rashid, S. Structural studies of Staphylococcus aureus Sortase inhibiton via Conus venom peptides. *Archives of biochemistry and biophysics* **671**, 87-102 (2019).

APPENDIX A: DNA Sequences, Plasmid Maps and Protein Sequence Alignment

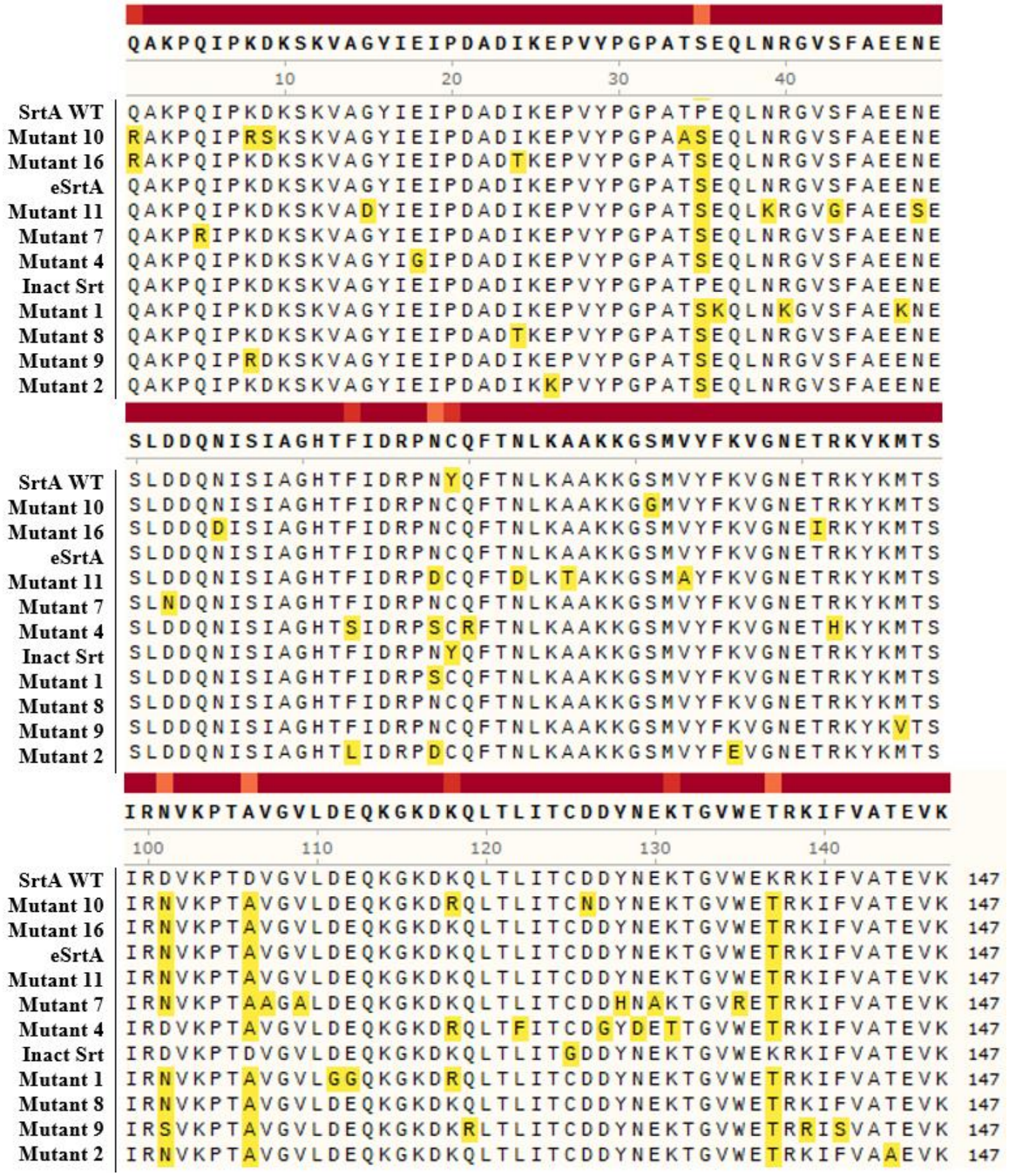


Fig. A-1: Alignment of SrtA mutant sequences with respect to SrtA WT. Mutations are highlighted in yellow.

Table A-1: Protein sequences of substrates used in SrtA reactions.

Name	MW (kDa)	Sequence
ELP-LPETG	24.89	MCVPGEGVPGVGVPGVGVPGVGVPGVGVPGAGVPGAGVPG GGVPGGGVPGEGVPGEGVPGVGVPGVGVPGVGVPGVGVPG AGVPGAGVPGGGVPGGGVPGEGVPGEGVPGVGVPGVGVPG VGVPGVGVPGAGVPGAGVPGGGVPGGGVPGEGVPGEGVPG VGVPGVGVPGVGVPGVGVPGAGVPGAGVPGGGVPGGGVPG EGVPGEGVPGVGVPGVGVPGVGVPGVGVPGAGVPGAGVPG GGVPGGGVPGEGVPGEGVPGVGVPGVGVPGVGVPGVGVPG AGVPGAGVPGGGVPGGGVPGEGLPETGG
GFP-LPETG	30.81	MALNDIFEAQKIEWHEMVSKGEELFTGVVPILVELDGDVN GHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTT LTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDD GNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEY NYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQ QNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEF VTAAGITLGMDELYKDDDDKLPETGHHHHHH
GGG-mCherry	28.08	MGGGVSKKGEEDNMAIIKEFMRFKVHMEGSVDGHEFEIEG EGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKA YVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDS SLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMY PEDGALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPG AYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYK LEHHHHHH
GFP-P0	29.86	MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATY GKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMK QHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKN GIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH YLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKD DDDKWLQDVHVYPKLEHHHHHH

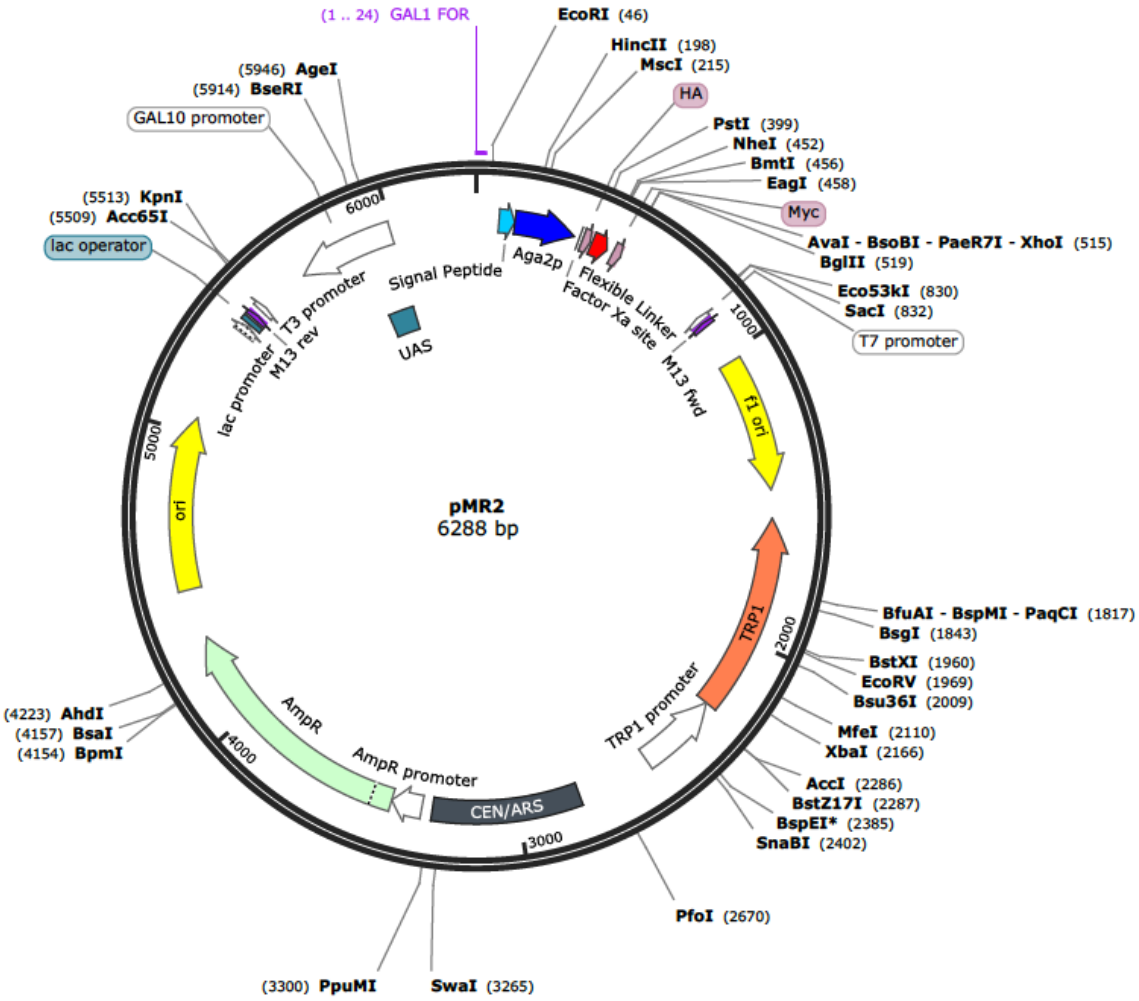


Fig. A-2: The plasmid map of pMR2 is shown, which contains the LPETG peptide fused to the C-terminus of Aga2p on the yeast surface.

APPENDIX B: Flow Cytometry Data

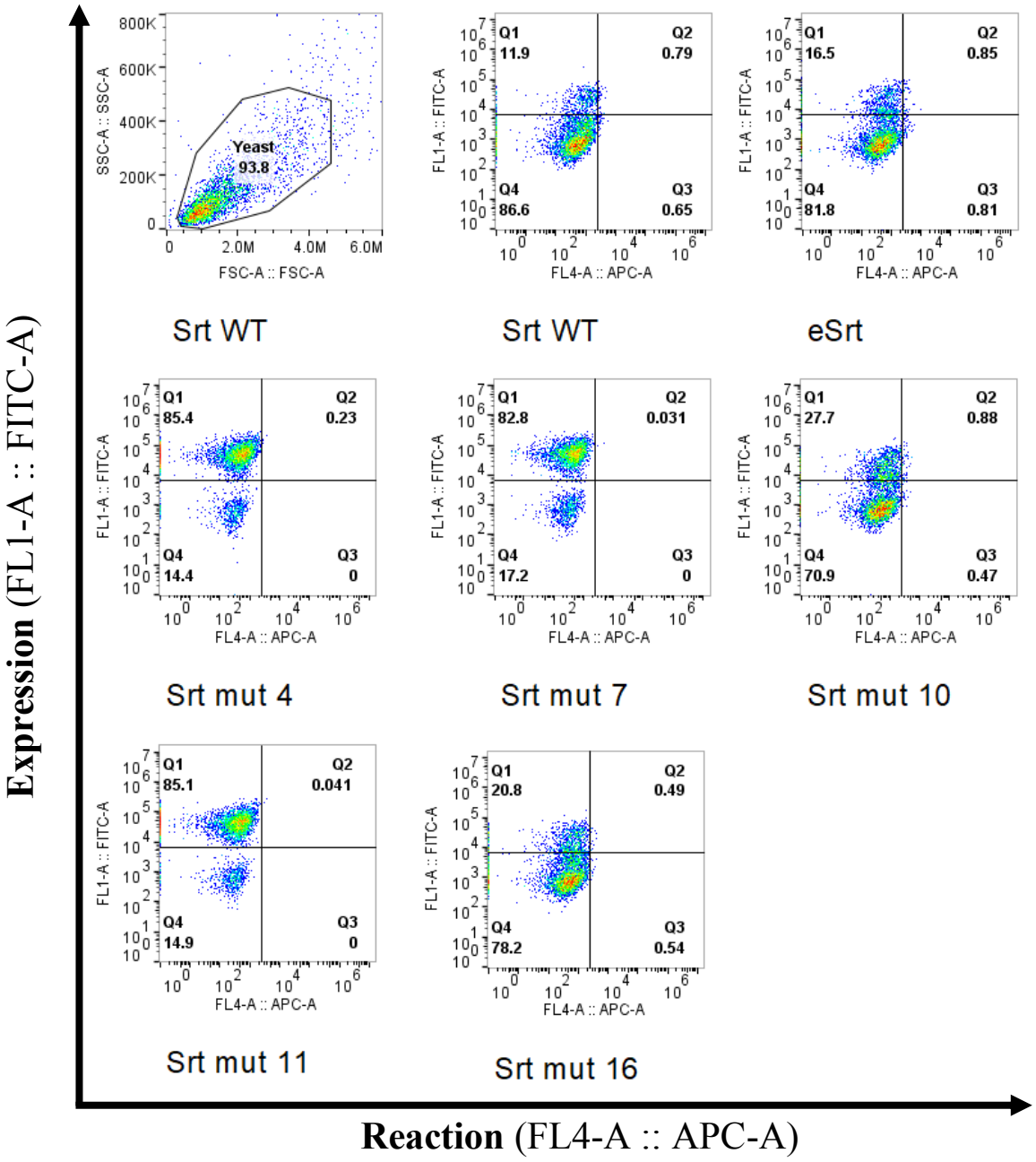


Fig. B-1: GGG-mCherry reaction on the yeast surface with SrtA WT and mutants. The reaction was performed with 584 µM GGG-mCherry substrate at 3 h and 37 °C.

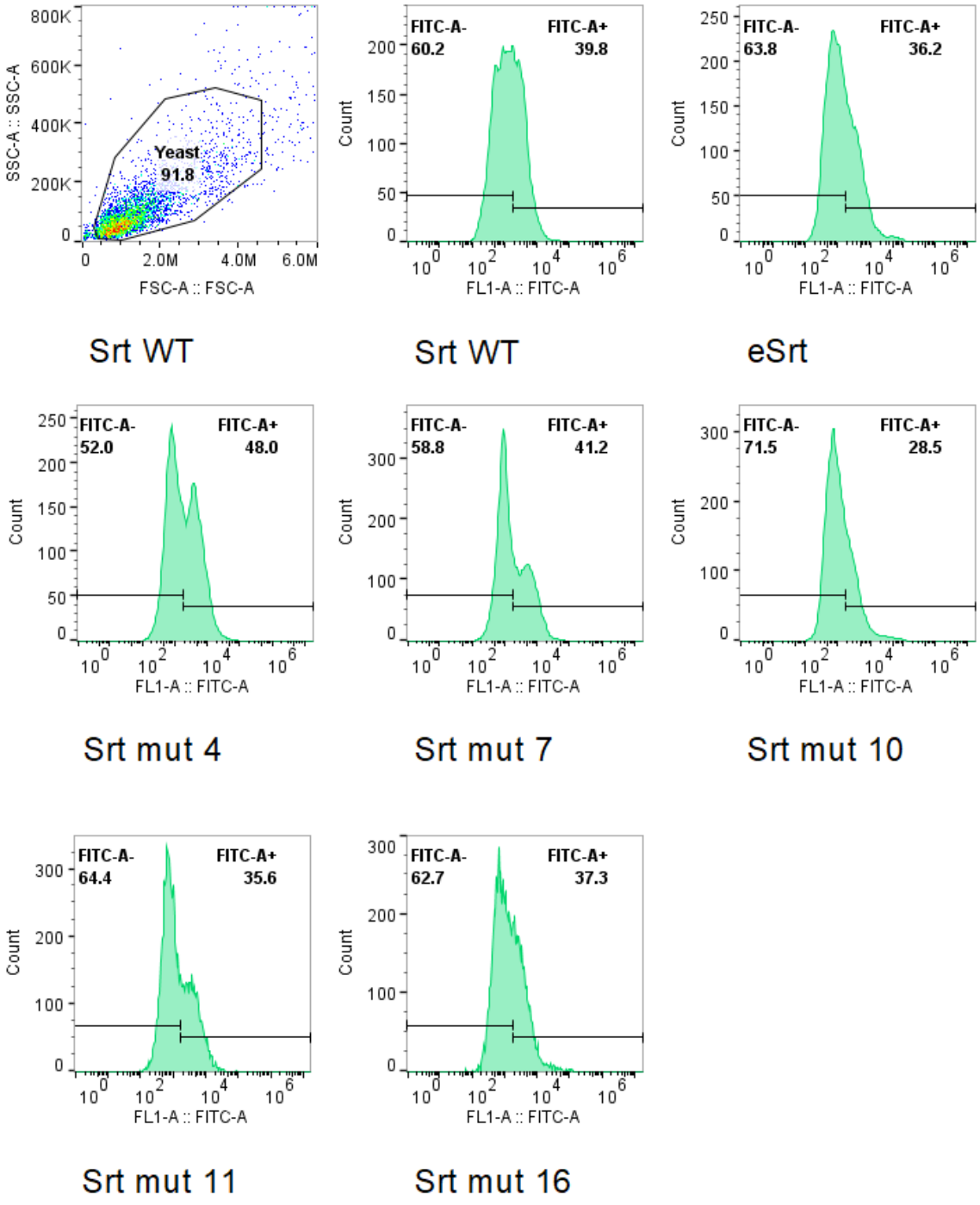


Fig. B-2: GFP-P0 reaction on the yeast surface with SrtA WT and mutants. The reaction was performed with 485 µM GFP-P0 substrate at 3 h and 37 °C.

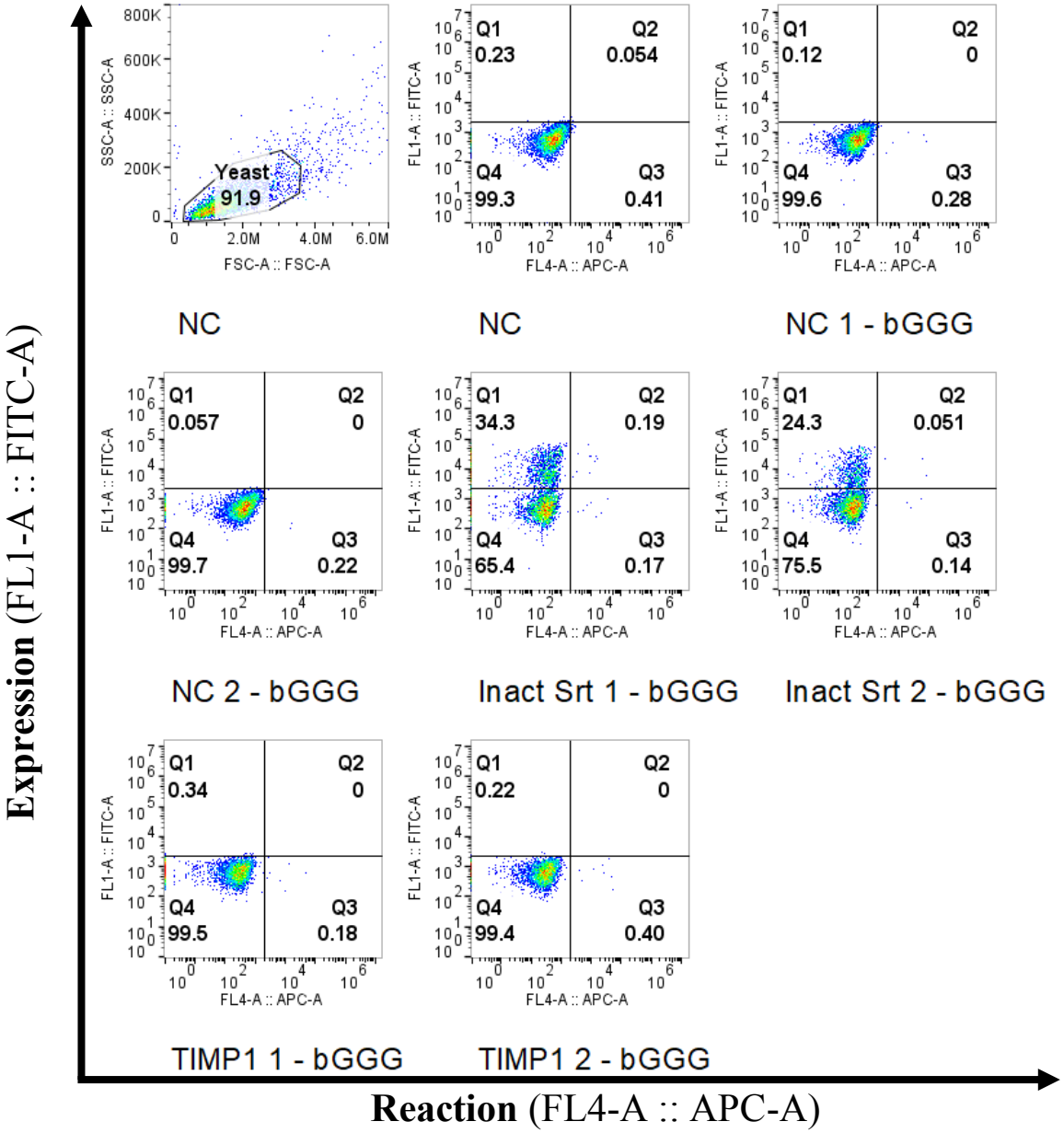


Fig. B-3: Control samples were incubated in 100 µM bGGG for 3 h at 37 °C and labelled for flow cytometry as described in the methods. TIMP1 replicates were only labelled with bP0 and streptavidin-647 due to lack of an HA epitope for measuring expression.

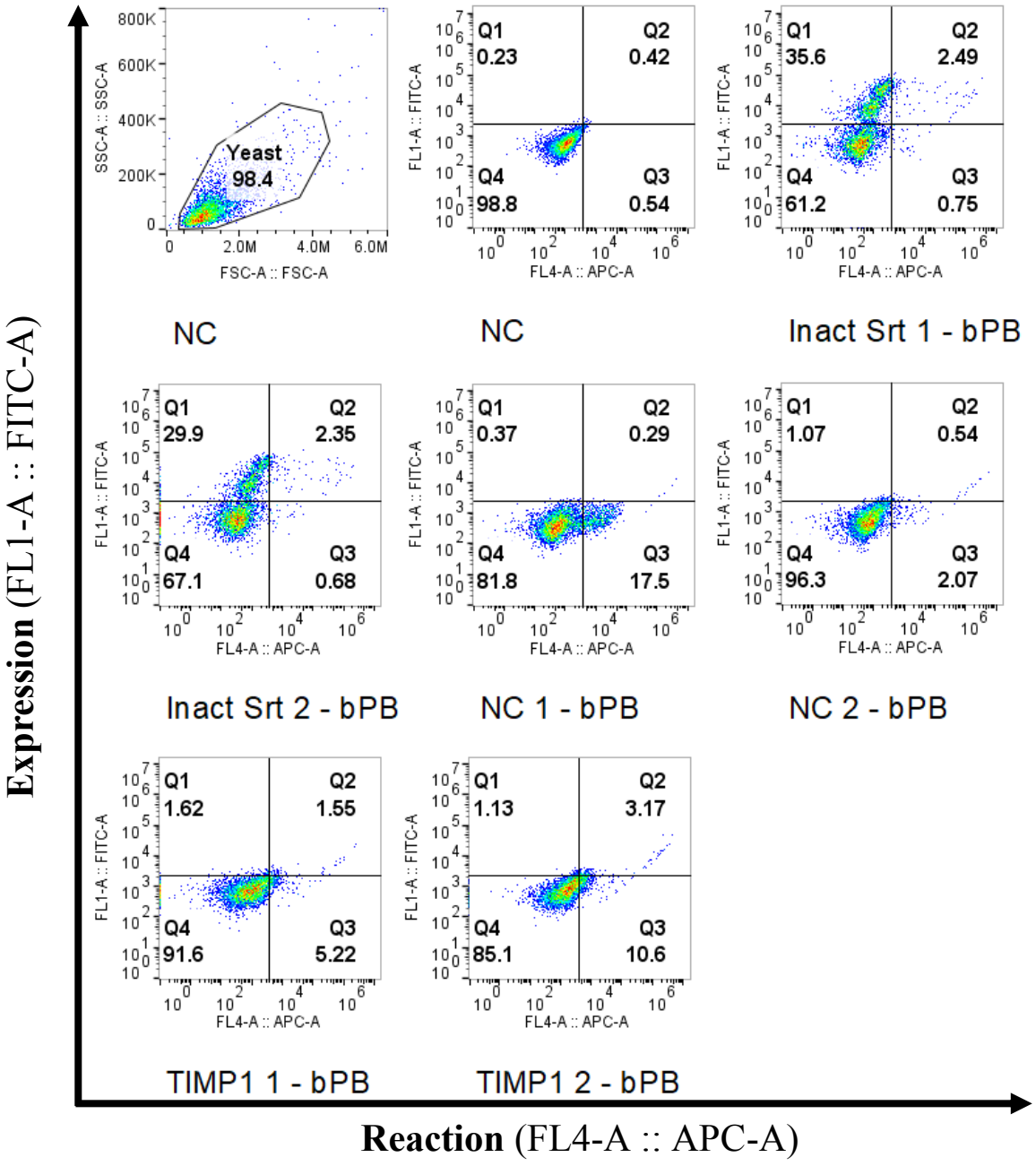
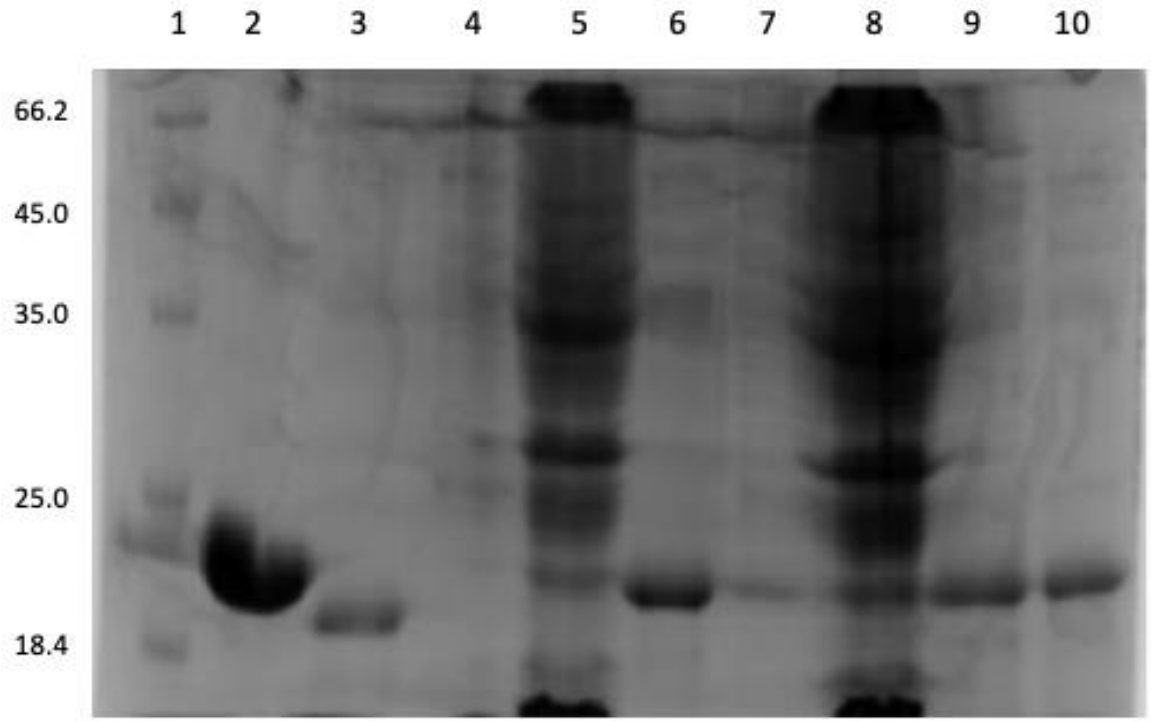


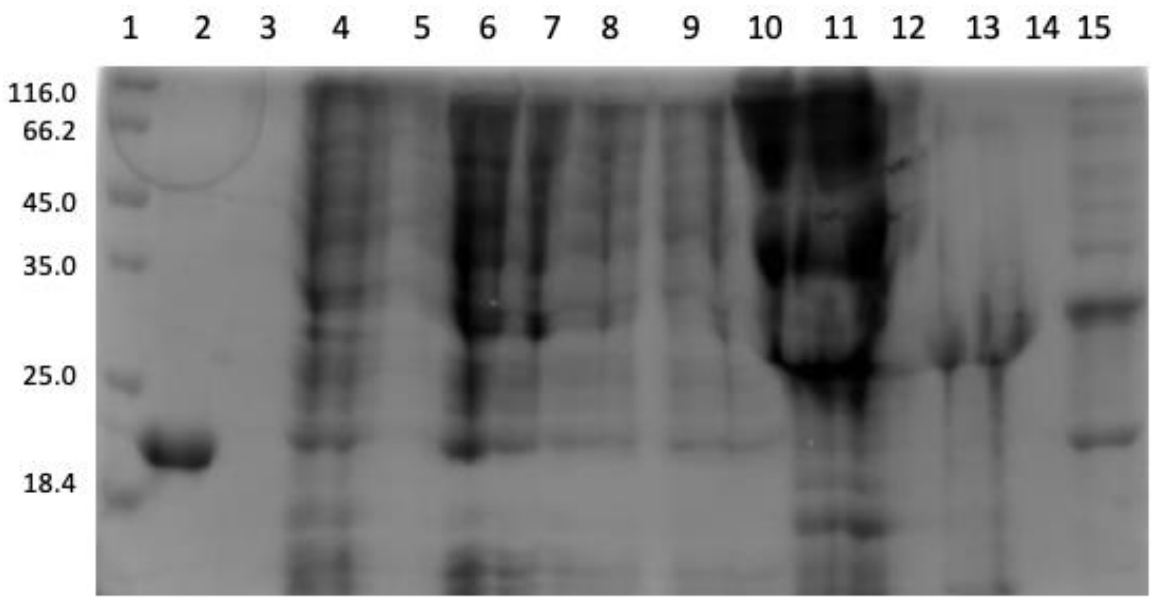
Fig. B-4: Control samples were incubated in 100 µM bP0 for 16 h at 37 °C and labelled for flow cytometry as described in the methods. TIMP1 replicates were only labelled with bP0 and streptavidin-647 due to lack of an HA epitope for measuring expression.

APPENDIX C: SDS-PAGE GELS



Lane	Name
2	SrtA WT (74 μM)
3	Mutant 1
4	Mutant 2
5	Mutant 2: denatured lysate
6	Mutant 4
7	Mutant 7
8	Mutant 7: denatured lysate
9	Mutant 8
10	Mutant 9

Fig. C-1: SrtA WT and mutants were run on SDS-PAGE. Aside from mutants 2 and 7, SrtA mutants were successfully expressed in soluble form in *E. coli*.



Lane	Name
2	SrtA WT (74 μM)
3	ELP-LPETG: un-induced
4	ELP-LPETG: induced (1st replicate)
5	ELP-LPETG: induced (2nd replicate)
6	ELP-LPETG: lysate (1st replicate)
7	ELP-LPETG: lysate (2nd replicate)
8	ELP-LPETG: 37 °C precipitate
9	ELP-LPETG: 37 °C hot spin
10	ELP-LPETG: 4 °C resolubilized pellet
11	ELP-LPETG: 4 °C cold spin
12	
13	ELP-LPETG: 4 °C cold spin after TCEP incubation
14	
15	GGG-mCherry (20 μM)

Fig. C-2: ELP-LPETG is induced, lysed, purified using one cycle of ITC, and run on SDS-PAGE with coomassie blue stain. ELP-LPETG is difficult to visualize using coomassie blue stain, but this stain could be used to detect impurities. Notably, impurities are reduced significantly after incubation in TCEP (lane 11 to lane 13). Furthermore, a faint outline can be seen around the ELP-LPETG band (lane 13; see Fig. C-2).

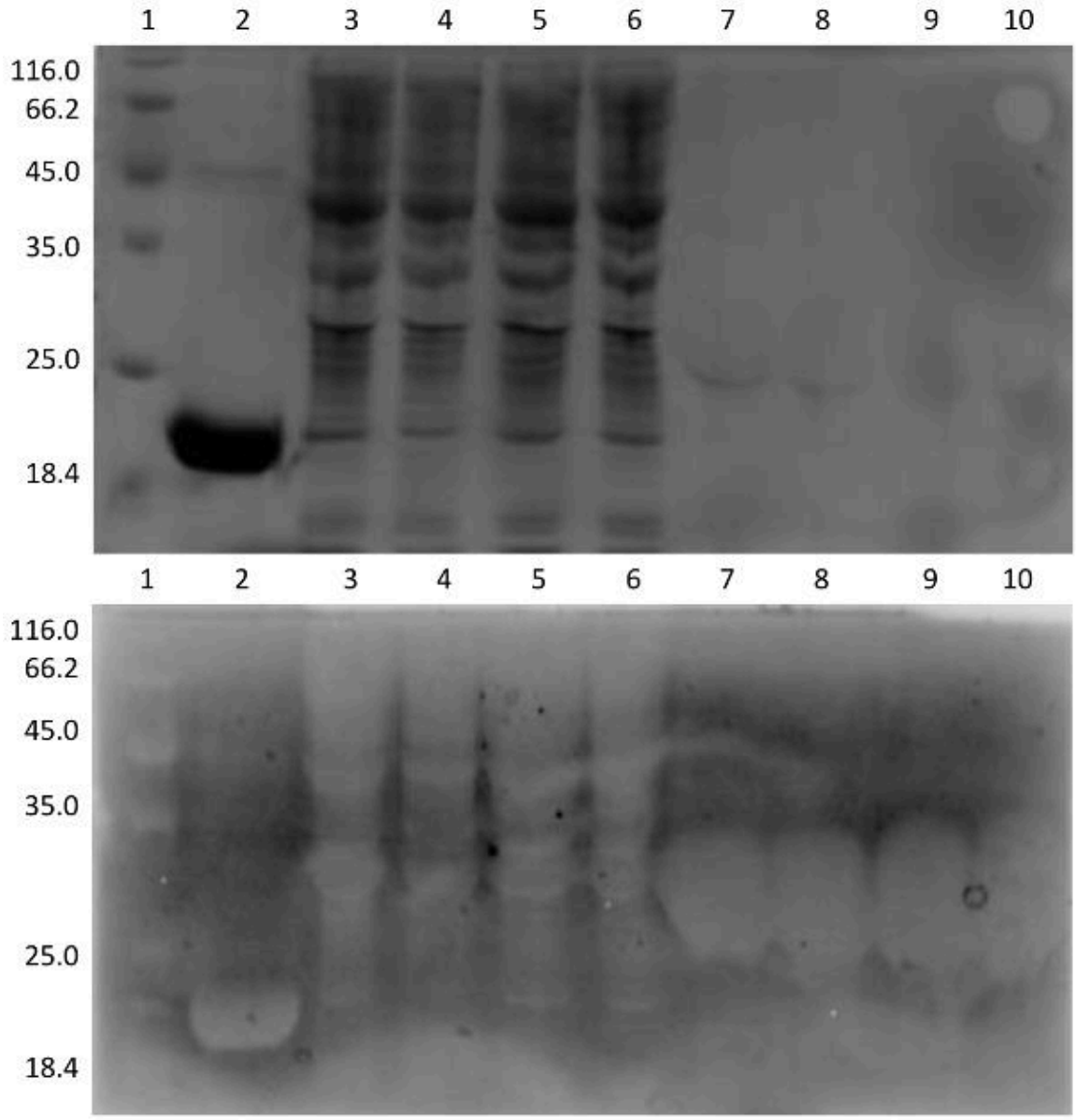


Fig. C-3: Two stains are compared for detecting ELP-LPETG. Gels contain 74 μM SrtA WT (lane 2), induced ELP-LPETG (lanes 3-6), and ELP-LPETG products after one cycle of ITC (lanes 7-10). The coomassie blue gel (top) shows a faint outline for the ELP-LPETG product, but the copper-stained gel (bottom) reveals large bands of pure ELP-LPETG. coomassie blue (top) and copper-stained (bottom) gels are shown for comparison.

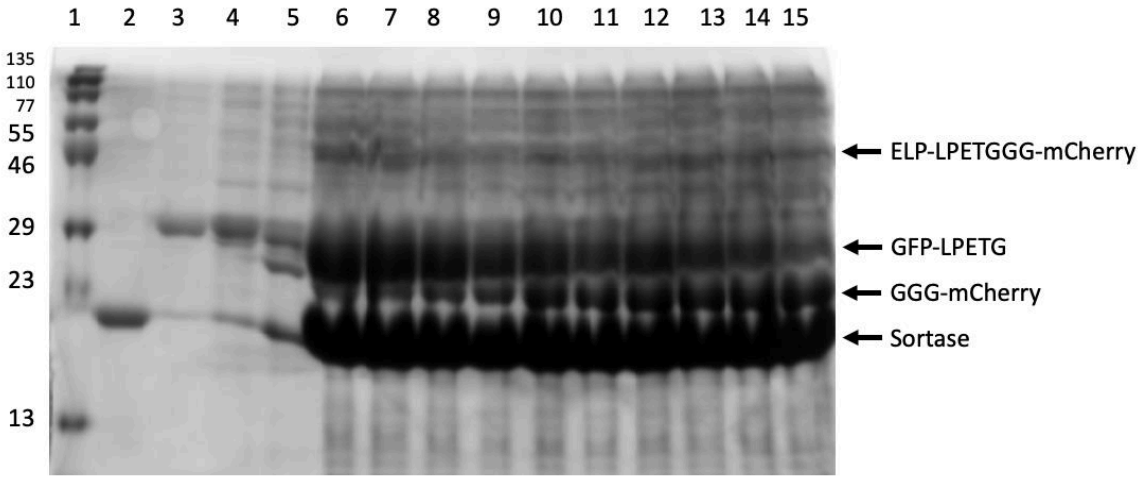


Fig. C-4: The SrtA-mediated ligation of GFP-LPETG and GGG-mCherry is analyzed across a range of concentrations (lanes 6-15). The GFP-LPETG to GGG-mCherry ratio for lanes 6 and 15 are 10:1 and 1:10, respectively.

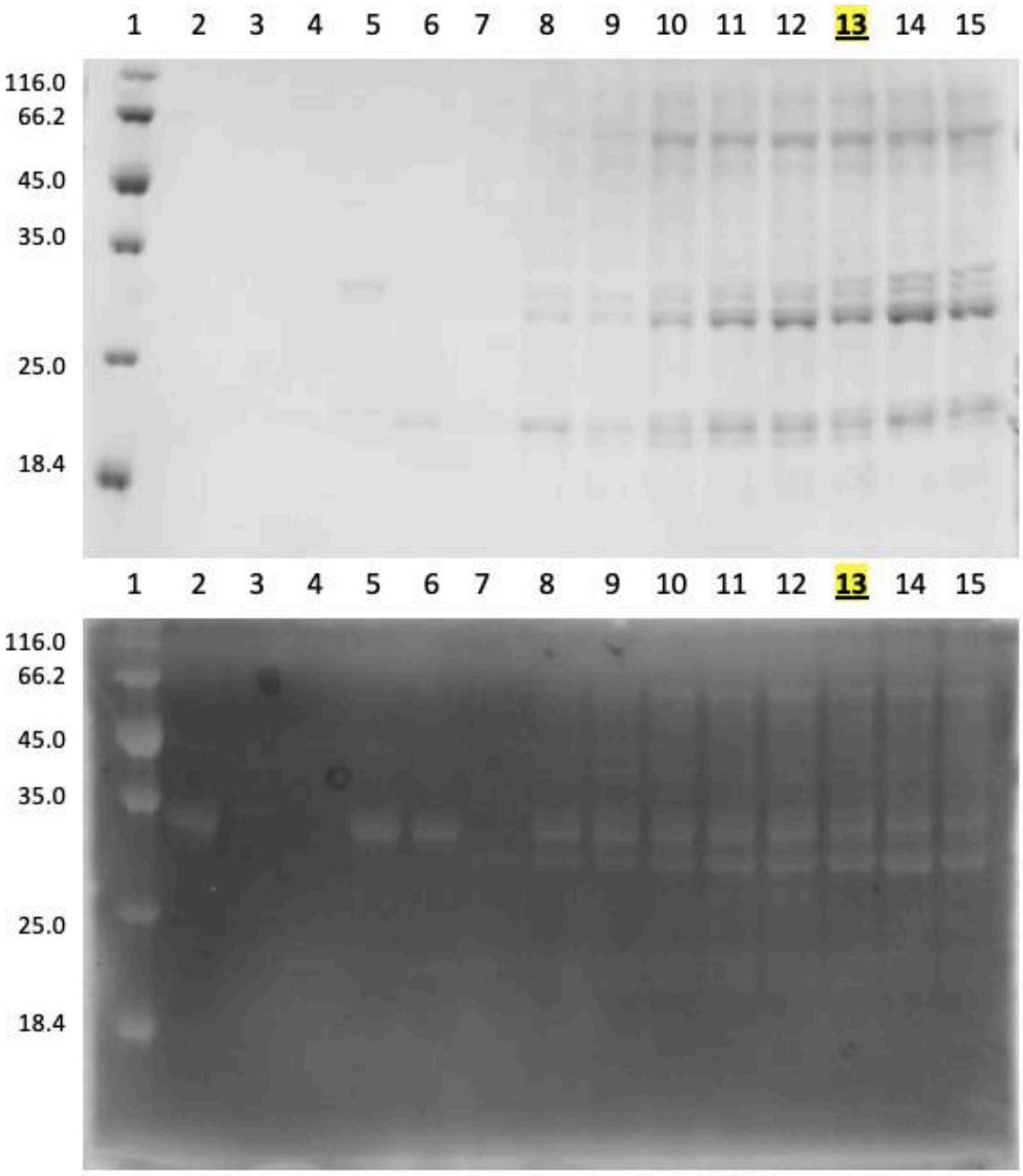


Fig. C-5: The GGG-mCherry concentration was optimized for the ligation of ELP-LPETG and GGG-mCherry, mediated by SrtA WT. Concentrations from 10-100 μM were tested. ImageJ was used to determine that lane 13, corresponding to 80 μM GGG-mCherry, had the highest yield. ELP-LPETG and SrtA WT were maintained at 20 μM. coomassie blue (top) and copper-stained (bottom) gels after one round of ITC are shown for comparison.

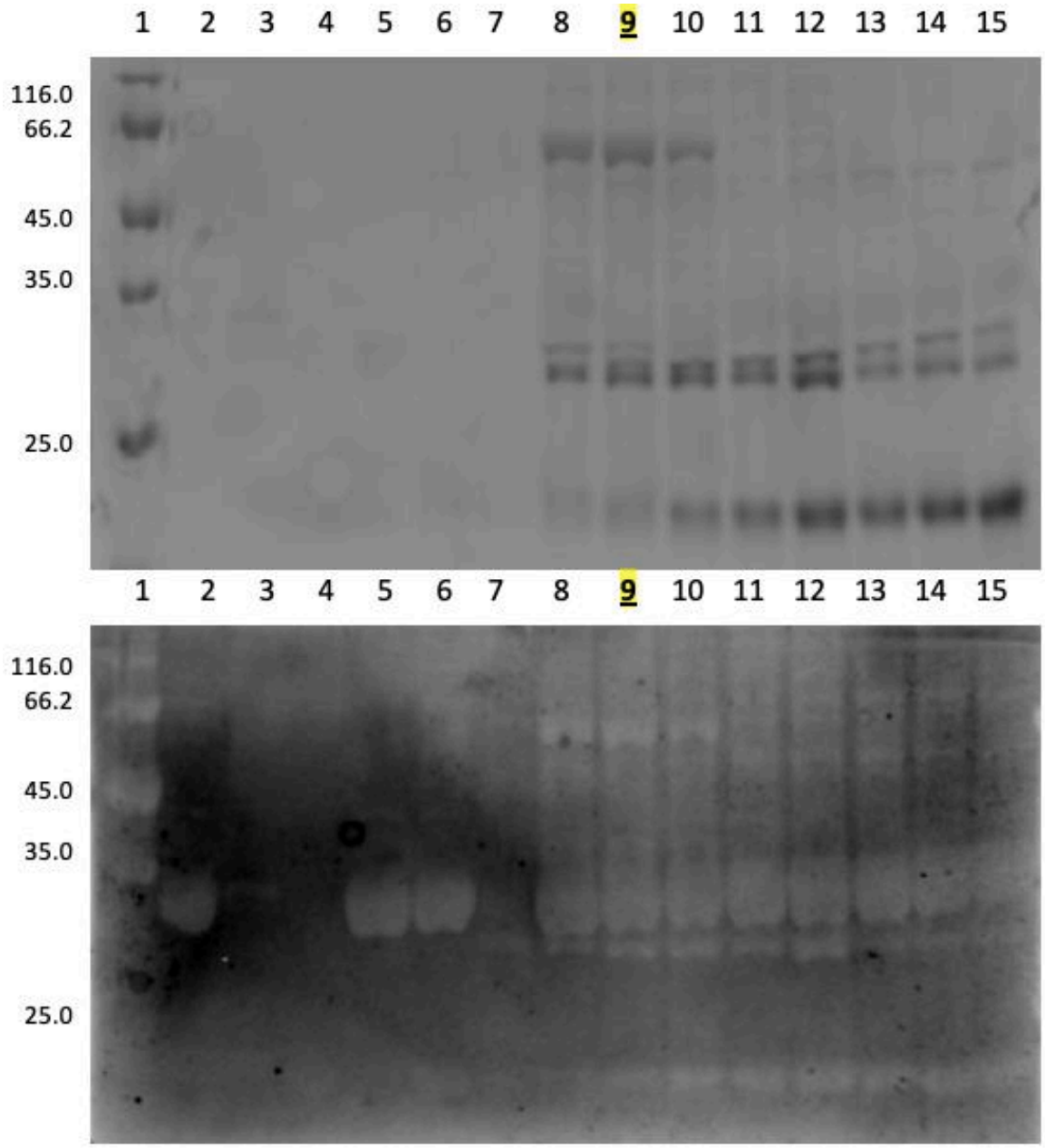


Fig. C-6: After the GGG-mCherry optimization, the SrtA WT concentration was also optimized for the ligation of ELP-LPETG and GGG-mCherry, mediated by SrtA WT. Lane 9, corresponding to 20 μM SrtA WT, was determined by ImageJ to contain the highest protein yield. ELP-LPETG and GGG-mCherry were maintained at 20 μM and 80 μM, respectively. coomassie blue (top) and copper-stained (bottom) gels after one round of ITC are shown for comparison.

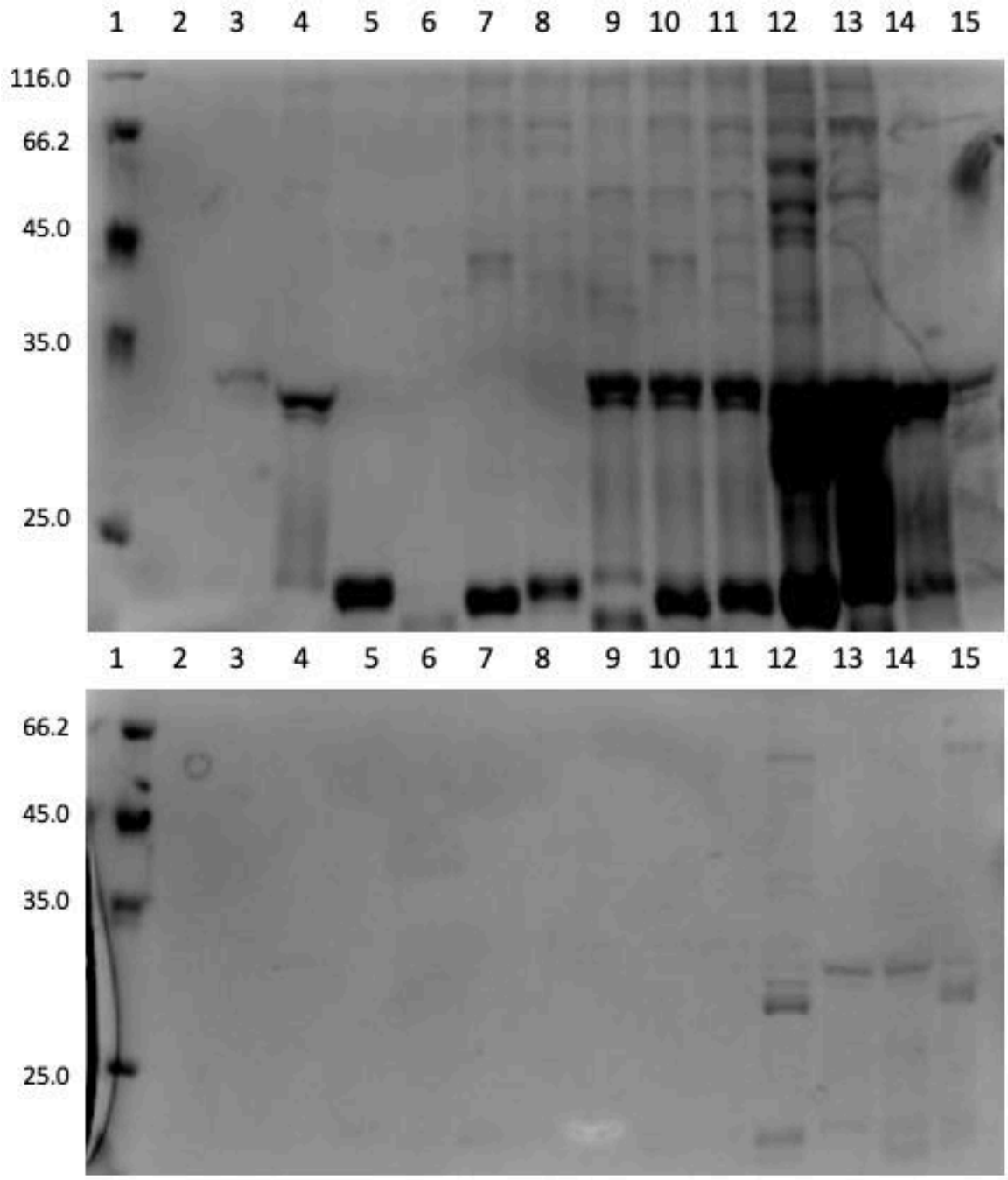


Fig. C-7: 20 μM SrtA WT or SrtA mutant, 20 μM ELP-LPETG, and 80 μM GGG-mCherry were added (lanes 12-15) adjacent to control samples (lanes 2-11). SrtA WT (lane 12) and mutant 8 (lane 15) retained specificity for GGG, whereas GGG transpeptidation activity was lost in mutants 1 (lane 13) and 4 (lane 14).

APPENDIX D: PYMOL ANALYSIS

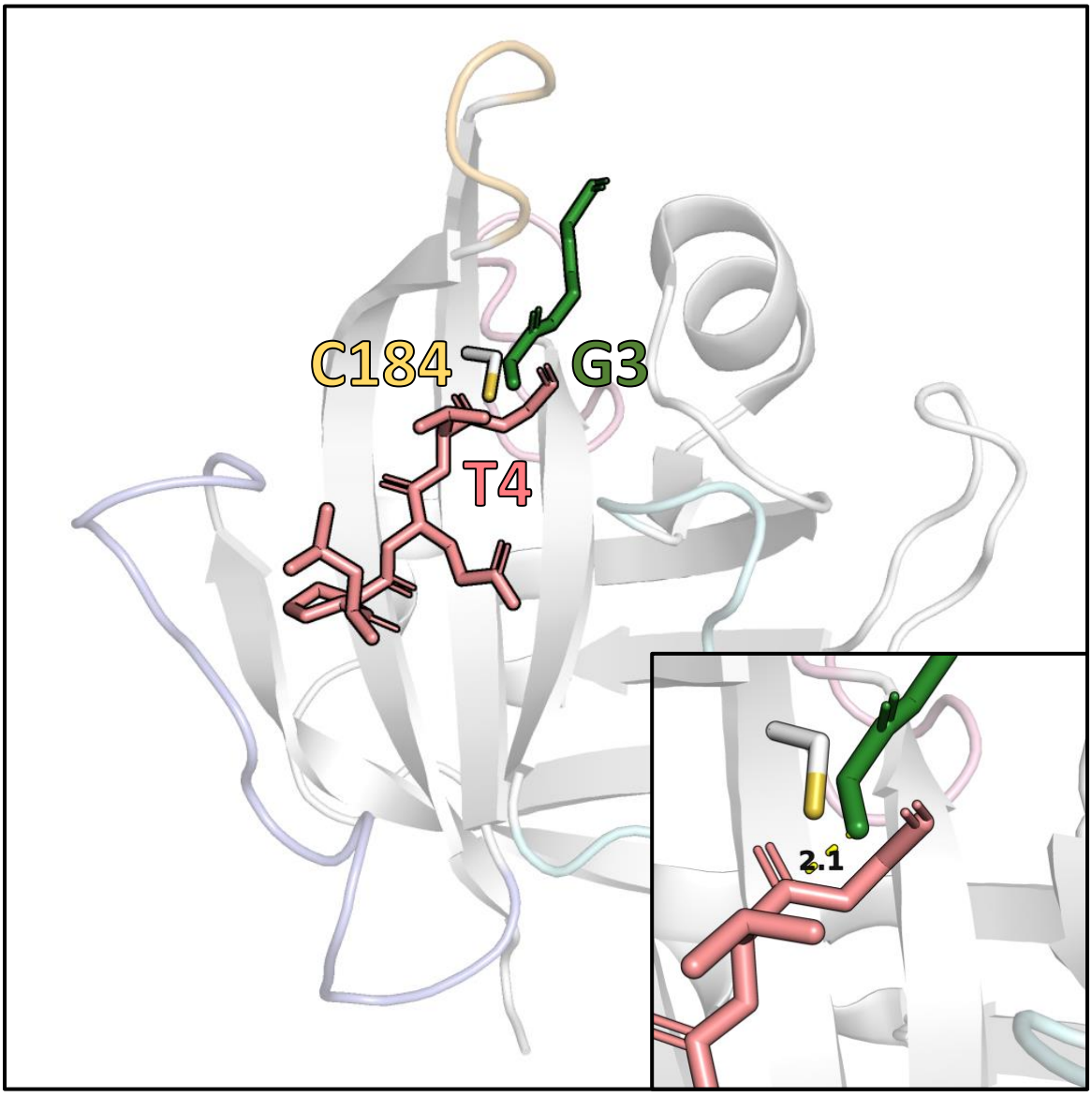


Fig. D-1: The LPET-SrtA complex is shown along with the predicted orientation of GGG in the substrate cleft, as described previously. With the amine nucleophile positioned 2.1 Å away from the thioacyl intermediate, this distance could be feasibly reduced to the proximity required for a covalent bond.

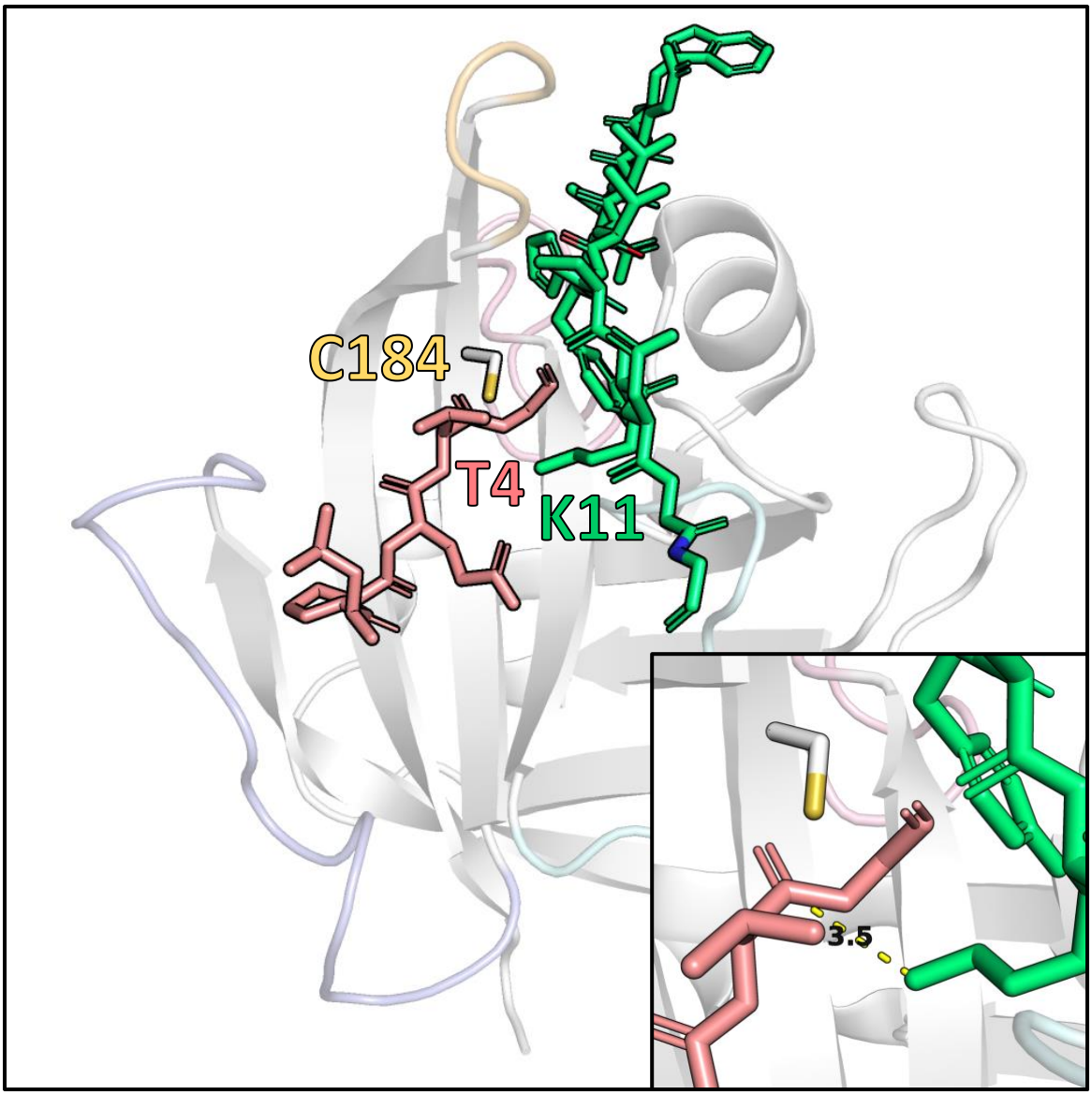


Fig. D-2: The LPET-SrtA complex is shown along with the predicted orientation of P0 in the substrate cleft, as described previously. With the lysine ε-amine positioned 3.5 Å away from the thioacyl intermediate, this distance could be feasibly reduced to the proximity required for a covalent bond.