The Dissertation Committee for Jie Zhu Certifies that this is the approved version of the following dissertation:

The Dimerization of *Staphylococcus aureus* Sortase A on Cell Membrane

Committee:

Zhiwen Jonathan Zhang, Supervisor

Walter L Fast

George Georgiou

Hung-wen Liu

Jon D Robertus

The Dimerization of *Staphylococcus aureus* Sortase A on Cell Membrane

by

Jie Zhu, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May, 2010

Dedication

To my supporting parents and friends

Acknowledgements

I would like to thank my advisor Professor Zhiwen Jonathan Zhang for providing me with a novel and diverse project area for my graduate research and for creating an exciting environment in which to conduct research. The members of Zhang group have had a tremendous impact on this work. I would like particularly to acknowledge Dr. Changsheng Lu, Dr. Gabrielle Nina Thibodeaux, Dr. Roshani Cowmeadow, and Aiko Umeda for helping me get started. Later, Liang Xiang and Katie Moncivais joined our group. Although we have not worked together for long time, Liang Xiang helped me a lot in project progression and Katie was a super great helper in improving my English writing skills.

Last, but not least, I would like to give a big thank you to my parents, my boyfriend and my great friends here in Austin. My parents and my boyfriend have encouraged and supported me to continue my studies after I already had a master degree, and put up with me when I am in a bad "I am gonna quit" mood, especially the last year of my graduate studies. I also want to thank my friends here in Austin, their company and friendship have made my life full of joy and hope everyday although I was so far from my family by myself.

The Dimerization of *Staphylococcus aureus* Sortase A on Cell Membrane

Publication No._____

Jie Zhu, Ph.D.

The University of Texas at Austin, 2010

Supervisor: Zhiwen Zhang

Staphylococcus aureus sortase A (SrtA) transpeptidase is a prominent membrane bound virulence factor in gram-positive bacteria, which organizes the peptidoglycan cell wall of the organism. Here, we report the first direct observation of the self-association behavior of SrtA. Formation of a SrtA dimer is highly selective *in vitro* in *E. coli* and *in vivo* on the *S. aureus* cell membrane. Quantitative analysis of protein binding affinity indicated a moderate association between two SrtA molecules with an apparent K_d of about 55 μM *in vitro*. Furthermore, to address the importance of dimerization for enzyme function, site-directed mutagenesis on potential target residues was performed to generate monomer only SrtA mutant proteins to completely disrupt dimer formation both *in vitro* and *in vivo*. Finally, an *in vivo* activity assay was performed to evaluate the function of SrtA wild type protein as well as its monomer only mutants. Our data demonstrated that *S. aureus* cells expressing mutant SrtA in a monomer only form are more successful at invading human epithelial cells than those expressing wild type SrtA in dimer-monomer equilibrium. It suggested that the monomeric form of SrtA is more active than the

dimeric enzyme. We also demonstrated the uniqueness of SrtA dimerization by identifying that at least one other sortase family protein, SrtB only exists in monomer form. SrtA dimerization may have significant implications for understanding its biological function at both the cellular and molecular levels, which will lead to the development of new anti-infective therapies against gram-positive pathogens.

Table of Contents

List of Tables

List of Figures

Chapter 1 *Staphylococcus aureus* membrane protein sortase A as a potential drug target

S. AUREUS INFECTION

Gram-positive bacteria are responsible for numerous devastating human diseases, including anthrax, pneumonia, and meningitis (*1*). A gram-positive species of special interest today is *Staphylococcus aureus* (*S. aureus*). It is the major cause of a wide variety human infections, from minor skin infections to severe life-threatening illnesses such as bacterial endocarditis, pneumonia, meningitis, osteomyelitis, toxic shock syndrome (TSS), bacteremia and septicemia (*2*). As a temporal worldwide trend, both hospital-acquired and community-acquired *S. aureus* infection have increased dramatically in the past thirty years. In particular, the community-acquired infection is now increasingly common (*3, 4*).

S. aureus is usually carried on the human skin and mucous membranes, such as in the nose of healthy people (*5, 6*). Based on von Eiff's investigation conducted in 2001, three groups were distinguished by carriage patterns. About 20 percent of healthy individuals always carry a certain strain of *S. aureus*, which are classified as persistently colonized. Sixty percent of the population carries *S. aureus* intermittently with changing strains and are classified as intermittently colonized. Less than twenty percent of the population never encounters or harbors the pathogen. Persons colonized with *S. aureus* are at increased risk for infections, especially those originating from endogenous colonies in nasal mucosa (*7*). However, the biology of colonization by staphylococci, is not completely understood yet (*8*).

The staph bacteria can cause disease, mediated by tissue invasion and/or toxin production (*2*). Since *S. aureus* is a commensal in the nares, vagina, pharynx, or on the

surface of wounded skin, the infection can be easily initiated when there is a break on the skin or mucosal barrier that would allow *S. aureus* to enter adjoining tissues or the bloodstream (*5, 6*). To be internalized into mammalian non-professional phagocytes, such as endothelial cells and epithelial cells is an essential step in the process of *S. aureus* tissue invasion (*9-12*). The internalization not only causes injury in those cells, but also activates a cascade which leads to endovascular disease (*2*). The internalization is mediated by bacterial surface adhesions named MSCRAMMs, Microbial Surface Components Recognizing Adhesive Matrix Molecules (*13*). *In vitro* studies indicated that once adhered to the cells, *S. aureus* is phagocytized into mammalian cells (*6, 14*). Once internalized, the intracellular environment can protect bacteria from antibiotics and host defense response, which may enhance bacterial survival and facilitate development of persistent and recurrent infections (*15*). The phagocytized bacteria will produce proteolytic enzymes to facilitate access to adjoining tissues and the entry of bacteria into the bloodstream. When the bacteria get into the adjoining sub-epithelial tissues, they can induce an inflammatory response that results in abscess formation (*16*).

After phagocytosis, mammalian cells modify their protein expression pattern as well. Fc receptors and adhesion molecules are expressed on cell surface, which may act as a binding site for MSCRAMMs and contribute to enhanced bacterial adherence (*17*). Cytokines such as interleukin-1, interleukin-6, and interleukin-8 are released from endothelial cells invaded by bacteria (*18*). When in contact with intact *S. aureus* cells or peptidoglycan components especially associated with lipoteichoic acids, tissue-based macrophages and circulating monocytes can release interleukin-1, interleukin-6, interleukin-8, and tumor necrosis factor a (TNF-a) (*19, 20*). These cytokines are released into the bloodstream and contribute to the exhibitions of fever, hypotension, multi-organ dysfunction and other phenotypes associated with systemic *S. aureus* diseases (*21-23*).

S. aureus associated toxins may have effects at sites distant from the spots of colonization and usually cause life-threatening diseases such as toxic shock syndrome. Toxins produced by *S. aureus* are superantigens, which are T-cell mitogens. They can bind directly to invariant regions of major histocompatibility complex (MHC) class II molecules. The MHC-bound superantigens can attach to T-cells and cause an expansion of clonal T-cells, resulting in a massive release of cytokines. The pathophysiology of endotoxin shock syndrome is similar to cytokine mediated TSS: that the bacterial products trigger the elevated level of cytokines and cause tissue damage (*21-23*).

Ever since the introduction of penicillin in the 1940s, bacteria have been developing various strategies to circumvent antibiotics. Twenty years after the introduction of penicillin, *S. aureus* and other gram-positive bacteria managed to develop resistance to penicillin and other penicillin-like drugs, such as methicillin, oxacillin and nafcillin. *S. aureus* strains that are resistant to all ß-lactam antibiotics are commonly termed as methicillin-resistant *S. aureus* (MRSA). The proportion of infections caused by methicillin-resistant *S. aureus* has been increasing significantly in the United States. A 2007 report from Centers for Disease Control and Prevention (CDC) estimated that, from 1999 to 2005, the number of MRSA infections has increased from 127,000 to 278,000. In 2005, there were 18,650 hospital stay-related deaths associated with MRSA, indicating that MRSA infections cause even more deaths than AIDS in United States (*24*). In Europe, MRSA caused 1,629 deaths in 2005 in England and Wales, but only 51 in 1993 (*25*). Molecular typing techniques have clearly demonstrated the rapid population of these disease-producing, methicillin resistant strains spreading to diverse geographic areas. Antibiotics, which can effectively treat MRSA these days include glycopeptides, teicoplanin and vancomycin. Historically, these were considered drugs of last resort to treat MRSA. However, ever since the emergence of first case of vancomycin resistant

bacteria in 2002, strains with intermediate or full resistance to vancomycin have been isolated in some countries at an increasing frequency (*26*). Thus, it has become a critical challenge of microbial research to identify novel drug targets and develop new inhibitors to treat MRSA. Several recently developed antibiotics, such as linezolid of the oxazolidinone class and daptomycin of the lipopeptide class have proven to have activity against MRSA.

However, massive production and use of antibiotics in modern industry, agriculture, and clinics plays a significant role in the emergence of antibiotic resistant bacteria by applying an environmental pressure to allow strains with certain traits to survive and reproduce. It creates a persistent demand for identification of novel drug targets and new antibiotics of therapeutic value since the current therapy will fail sooner or later due to the selection pressure (*27*). A newly emergent hypothesis is that instead of inhibiting bacterial growth and reproduction, bacterial virulence could be inhibited, thereby interrupting bacterial infections without necessarily creating an environmental selection pressure. The concept of so-called anti-infective therapies require proof-of-principle and has not been developed for use with human yet. Comprehensive knowledge of the pathogen invasion mechanism, pathogen-immune system interactions and experimental testing would be required to verify its feasibility (*28*). In *S. aureus*, a prominent virulent factor, sortase A (SrtA), has been proven to be a key regulator for pathogen invasion into human tissues and represents a promising target for the development of anti-infective drugs (*29*). The current research progress to establish SrtA's cellular and molecular mechanism is summarized and described below in Chapter 1.

THE STRUCTURE OF THE *S. AUREUS* CELL WALL

The cell wall envelope of gram-positive bacteria, peptidoglycan (PG), is composed of polymers of long glycan chains cross-linked via peptide bridges as shown in Figure 1.1 (*30*). In *S. aureus*, the cell wall synthesis starts from production of a soluble nucleotide precursor, Park's nucleotide (UDP–MurNAc–L-Ala–D-isoGln–L-Lys–D-Ala–D-Ala), in the cytoplasm. The UMP nucleotide is then removed and the precursor is transferred to a bacterial membrane carrier to generate lipid I on the cell membrane (C55-PP–MurNAc–L-Ala–D-isoGln–L-Lys–D-Ala–D-Ala). Lipid I is further modified with GlcNAc (N-acetylglucosamine) and Gly5 (pentaglycine) cross-bridges at the ε-amino of L-Lys to generate lipid II C55-PP–MurNAc–[L-Ala–D-isoGln–L-Lys(Gly5)–D-Ala–D-Ala]– (1-4)-GlcNAc). By an unknown mechanism, lipid II is flipped across the cell membrane, where the MurNAc-GlcNAc subunit is polymerized into glycan strands (repeating disaccharide chains) by transglycosylation. The cross bridge between two disaccharides is formed by cleaving at the terminal D-Ala of pentapeptide precursor (L-Ala–D-isoGln–L-Lys(Gly5)–D-Ala–D-Ala) and forming an amide bond between the carboxyl group of the fourth D-Ala and the amino group of pentaglycine group in the other wall peptide unit. The cell wall synthesized in this manner is rigid enough for bearing stress and maintaining shape. The cross-linked matrix also provides a scaffold for incorporation of other molecules to promote interactions between microbes and their environment (*31*). These molecules, such as teichoic acids, polysaccharides, and proteins, are covalently or non-covalently associated with cell wall peptides (*32-34*). Many of them are MSCRAMMs and are required for successful infection by performing functions such as adherence to host tissues, invasion, and avoidance of the host immune system (*35-38*).

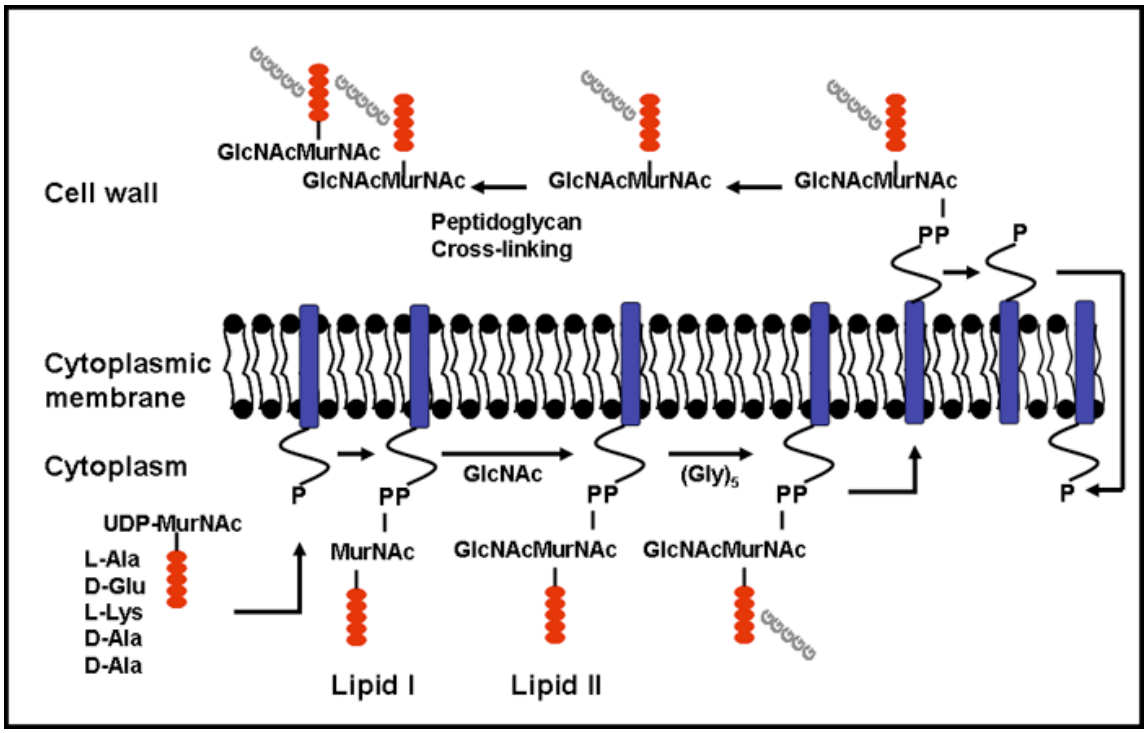

Figure 1.1 Peptidoglycan synthesis in *Staphylococcus aureus*.

CELL SURFACE PROTEINS AND SORTASE A

In gram-positive bacteria, many cell surface proteins are covalently anchored to the cell wall by a 206 amino acid, membrane-bound cysteine transpeptidase with an extracellular catalytic domain, known as surface protein sorting A or Sortase A (SrtA) (*39*). The sequence was first identified in *S. aureus*. Later, the sequence of the prototypical staphylococcal SrtA has been used to identify homologues in other bacterial genomes, revealing an abundance of sortase enzymes in almost all gram-positive bacteria (*40*). SrtA mutants display major defects in virulence, suggesting that the display of surface proteins on the cell wall is essential for the pathogenesis of *S. aureus* and other gram-positive bacterial infections (*41*). Therefore, the inhibition of SrtA's function might

be effective as a novel antibacterial therapy to treat human infections caused by gram-positive pathogens.

The SrtA enzyme displays various *S. aureus* surface proteins, including staphylococcal protein A (Spa), fibronectin-binding proteins, and fibronectin-binding clumping factors (*42*). These proteins are crucial for *S. aureus* tissue invasion. For example, when Spa is displayed on the cell surface, it can bind to the Fc region on host IgG through the heavy chain, so the serum antibodies are in the wrong orientation resulting in the disruption of phagocytic engulfment and opsonization (*43*). All of these surface proteins harbor an N-terminal signal sequence and a C-terminal sorting signal. The sorting signal consists of an LPXTG motif (where X can be any amino acid), followed by a hydrophobic domain, and a positively charged tail. The hydrophobic domain and charged tail temporarily retain the substrate protein in the plasma membrane where SrtA enzyme can work on the substrates in close proximity. Catalysis of SrtA has been proposed through a ping-pong mechanism. After the initial recognition between LPXTG motif and enzyme, the reaction is initiated by nucleophilic attack by the thiol group on active site residue Cys^{184} on the carbonyl carbon on threonine (T) to form a transient thioacyl intermediate. SrtA then amide-links the carbonyl carbon of threonine to the amino-group of a pentaglycine cross-bridge in a peptidoglycan precursor to generate a second tetrahedral intermediate and resolves into the peptidoglycan cell wall anchored final products (Figure 1.2) (*32, 39, 44-46*). The C-terminal 35 amino acids of Spa is shown to be the sorting signal and is sufficient to confer cell wall anchoring to other secreted protein sequences (*32*). This type of surface protein anchoring has been found to be a universal mechanism in gram-positive pathogens.

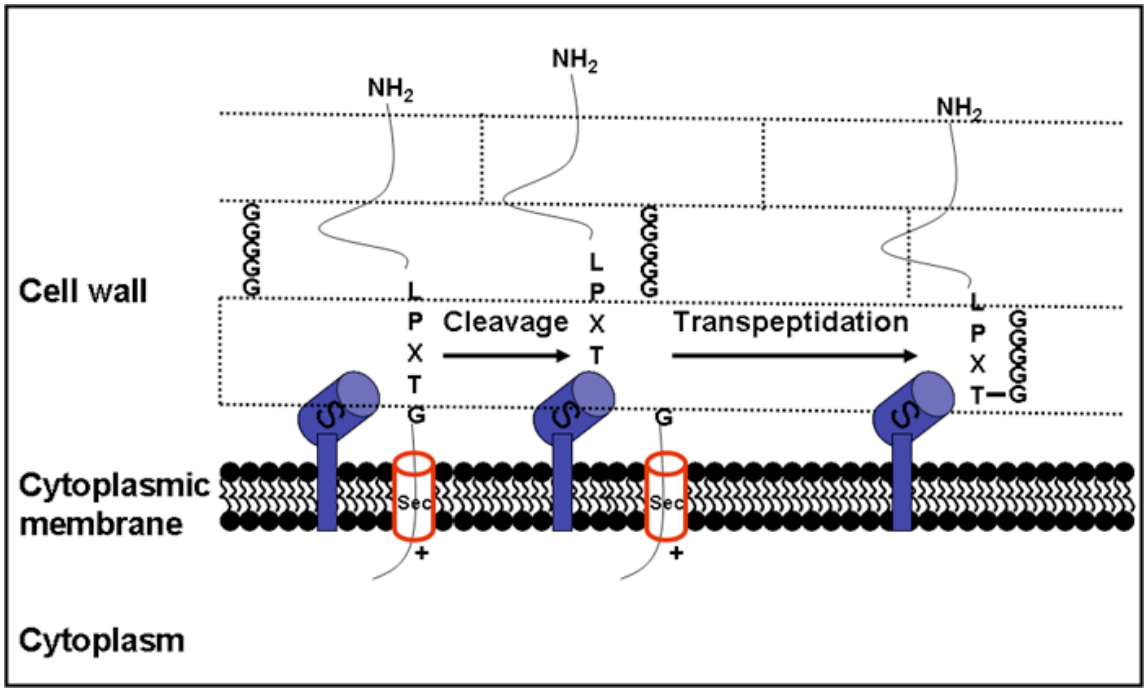

Figure 1.2 Mechanism of protein anchoring by SrtA.

Full length SrtA contains three domains: a transmembrane domain on the N-terminus consists of a stretch of non-polar amino acids (residues 1-24): an unstructured linker domain (residues 25-59) which connects the membrane spanning anchor and the C-terminal catalytic domain (residues 60-206). It has been shown that the catalytic domain itself, expressed and purified from *E. coli*, can independently fold into its catalytic form and can sufficiently perform the transpeptidation reaction with substrate analogs containing a sorting signal *in vitro* (*39, 47-49*). The NMR and X-ray crystallography structures of the SrtA catalytic domain revealed that the enzyme folds into a unique eight-stranded β-barrel connected by several helices and loops (*50, 51*). Similar overall structures are also observed in other sortase enzymes, such as *S. pyogenes* SrtA (*52*), *S. aureus* SrtB (*53, 54*), *Bacillus anthracis* SrtB (*53, 55*), and *S. pneumoniae* SrtC-1 and SrtC-3(*56*). However, since most of the reported structures listed above are apo-forms of

the enzyme without substrate bound or with substrate bound non-specifically, the molecular mechanism of transpeptidation and recognition between substrate and enzyme is not well identified. A recent publication from Suree et. al reported a substantially different structure of *S. aureus* SrtA, which is bound to a substrate analog, Cbz-LPAT (Cbz is a carbobenzyloxy protecting moiety, T is (*2R, 3S*)-3-amino-4-mercapto-2-butanol) to mimic the key catalytic thioacyl intermediate (*57*). This solution structure shows the dynamics of SrtA recognizing its substrate and adjusting its structure by closing and immobilizing the active site β6/ β7 loop. It has been shown that molecular docking using this structure gave more consistent results with SrtA inhibitors' SAR data as compared to the apo-form crystal structure (*58*).

The study of SrtA structures with and without substrate binding has enabled the proposal of a detailed plausible mechanism for sortase catalyzed transpeptidation. Some progress has been made to identify and verify the mechanism with biochemical and kinetic data in the past decade. The proposed active site, comprised of three conserved residues Cys^{184} His^{120} and Arg^{197}, is buried in a hydrophobic cavity formed by two β-stands β6/ β7. The important role of the β6/β7 loop in sorting signal recognition has been established (*50, 51, 59-61*). It is located in an ideal position to interact with the bulky, nonpolar leucine-proline residues at the N-terminal of sorting signal (*50*). Alanine scanning mutagenesis on the β6/β7 loop has indicated that Val^{168} and Leu^{169} may contribute to the hydrophobic interactions (*61*). The β6/β7 loop has been shown to be highly flexible in solution until the addition of calcium to stabilize the loop into the correct position to facilitate binding and cleavage of the sorting signal, which stimulates the enzyme's activity eight-fold (*62*). Binding of the sorting signal also locks the β6/β7 loop to a closed form, that occluding the active site (*57*). When this loop domain was

swapped between SrtA and SrtB, the recognition specificity was also altered, suggesting the function of this loop in substrate discrimination (*60*).

The recent identified NMR structure of SrtA with peptide substrate mimic specifically bound has enabled the proposal of a detailed mechanism of SrtA which is compatible with current data (Figure 1.3) (*63*). It is well established that Cys^{184} functions to perform a nucleophilic attack on the scissile peptide bond between the threonine and glycine in the LPXTG motif and forms a covalent linkage between enzyme and substrate. The acyl intermediate is then resolved by nucleophilic attack from the amino group of the pentaglycine cross-bridge within the cell wall. It was originally suggested that His^{120} functions to activate Cys^{184} and forms an imidazolium-thiolate bond (*64*). However, NMR-pH titration experiment was performed to estimate that the pK_a for Cys^{184} and His^{120} is ~9.4 and ~6.3-7 individually. Both of them are not charged at physiologic pH (*47, 63*). Solvent isotope effect measurements proposed that the enzyme in this uncharged form is not active. Only a small amount of the population (0.06%) contains Cys^{184} and His^{120} in their ionized form that is active (*47*). Based on this mechanism, Cys^{184} does not need to be activated and the thiolate can perform nucleophilic attack on the incoming carbonyl carbon on threonine directly. The imidazolium side chain of His^{120} can protonate the amide leaving group of glycine when the covalent bond between threonine and glycine is broken. NMR chemical shift experiments suggested that the deprotonated His^{120} is also located at the entry point of the second substrate, Gly_5 of the lipid II. Now the His^{120} functions as a general base to activate the terminal amine on glycine to facilitate the subsequent nucleophilic attack on the thioacyl tetrahedral intermediate (*47, 57*). Protonation of the second tetrahedral intermediate would produce the final product of transpeptidation anchored on the cell wall and regenerate the enzyme active sites for further catalysis (*44*). However, in all steps, the function of Arg^{197} is controversial with a

variety of different proposals and still under active debate and investigation. Early proposals suggested that Arg^{197} might be involved in proton transfer events, such as deprotonating Cys^{184} (*65*) or lipid II substrate (*51*). A recent study ruled out this possibility (*66*) and implied its function to stabilize either the sorting signal binding (*61, 65*) or oxyanion intermediate binding (*51, 61*). The stabilization of Arg^{197} acts through direct hydrogen bonding with the sorting signal. The possible position for this hydrogen bond is the carbonyl group of the peptide backbone between proline and X on the LPXTG motif. Dynamic NMR structures suggested the side chain of Arg^{197} stabilizes the oxyanion transition state of both tetrahedral intermediates (*57*).

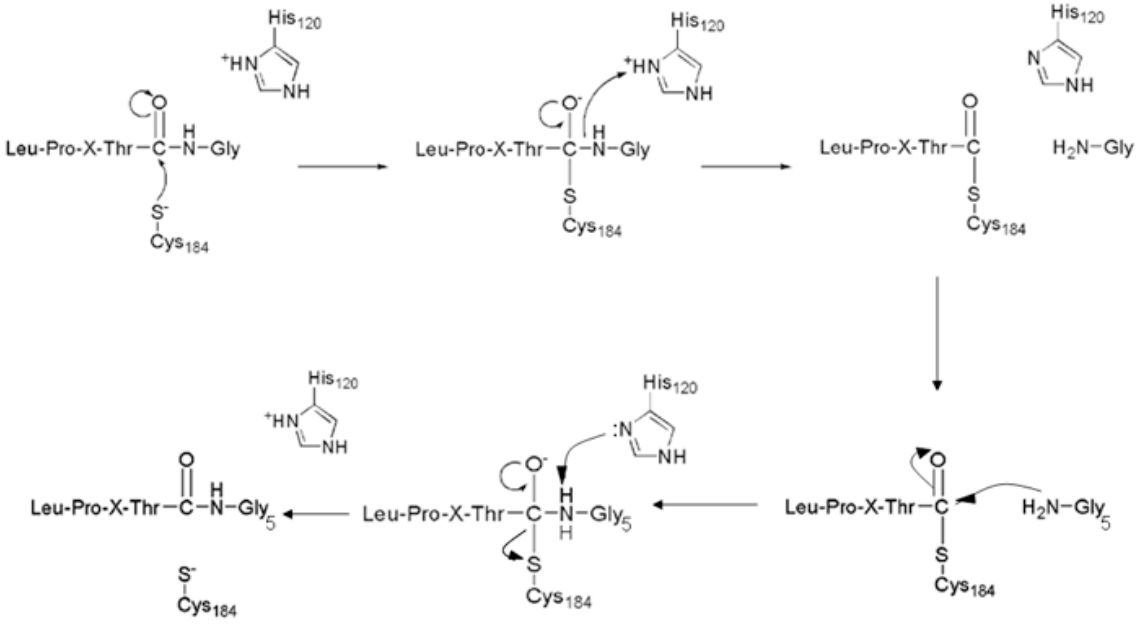

Figure 1.3 Proposed mechanism of SrtA catalysis.

Some investigations have focused on subcellular localization of SrtA and its substrate. A holoenzyme concept has been proposed by Hotje in 1993. It hypothesized that major membrane proteins involved in cell wall synthesis are coordinated spatially and temporally together to form a multi-enzyme complex (*67, 68*). This concept has been

supported by findings from various groups and the complex is shown to localize on the division plane (*69*). As a major enzyme for cell wall synthesis, SrtA is also found to localize to the active division plane or the polar region, the site of previous cell division (*70, 71*). It has been suggested that the positive charges flanking the transmembrane domain of SrtA retains the protein at these specific foci. The investigation on SrtA substrates indicates that the peptide leader sequence leads protein to cross the membrane to polar region, which colocalize with the foci for SrtA anchoring (*72*). The interaction responsible for arranging the localization of the putative cell wall synthesis holoenzyme, include SrtA and its substrates (cell surface protein precursors) is under active research, in the hopes of identifying the general mechanisms and pathways that control the biosynthesis of the cell wall and cell wall associated virulence factors, both of which are potential anti-microbial and anti-infective drug targets.

SORTASE A AS A DRUG TARGET

The ultimate goal of most studies of SrtA has been to identify a potent and selective inhibitor to block the cell surface display of virulence factors, which in turn disrupt the invasion into human tissues. The SrtA knockout strain *srtA*$^{-}$ of *S. aureus* has no defects in cell growth, suggesting that the highly specific inhibitors of SrtA will not generate a selection pressure on the bacteria population which lead to the evolution of multi-drug resistant "super bug" in the same manner as conventional antibiotics (*28*). This new therapeutic strategy may be generalizable to other gram-positive pathogens.

Several different strategies have been employed to develop inhibitors against sortase family proteins. Some identified inhibitors have been illustrated in Figure 1. 4. Inhibition of the sorting pathway was observed even before the identification and cloning of SrtA from *S. aureus*. Vancomycin and moenomycin that inhibit the lipid II biosynthetic pathway can also inhibit SrtA catalyzed transpeptidation, suggesting that

lipid II is a substrate for sorting reaction. Sortase inhibitors such as [2-(trimethylammounium)ethyl]methanethiosulfonate (1 in Figure 1.4) and *p*-hydroxymercuribenzoic acid (2 in Figure 1.4) react with sulfhydryl groups and inhibit cysteine proteases, and so imply the function of sulfhydryl group in the transpeptidation mechanism (*44*).

Small libraries of natural products extracted from medicinal plants were screened *in vitro* for SrtA inhibitors by a Korean group using purified SrtA protein (*73-81*). These studies have generated potential compounds worth further development, including *β*-sitosterol-3-*O*-glucopyranoside (3 in Figure 1.4) from *Fritillaria verticillata* with a IC_{50} of 18 μg/mL (*77*); berberine chloride (4 in Figure 1.4) from *Coptis chinensis* with IC_{50} of 8.7 μg/mL for SrtA, 6.3 μg/mL for SrtB (*76, 79*); psammaplin A1 (5 in Figure 1.4) from *Aplysinella rhax* with IC_{50} of 39 μg/mL for SrtA, 23 μg/mL for SrtB (*79*); bromodeoxytopsentin (6 in Figure 1.4) from *Topsentia genitrix* with IC50 of 19.4 μg/mL (*75*); curcumin (7 in Figure 1.4) from *Curcuma longa* with IC_{50} of 13 μg/mL (*73*); flavonoid phenols (such as morin, 8 in Figure 1.4) from *Rhus verniciflua* with IC_{50} of 37-52 μM for SrtA, 8-36 μM for SrtB (*80*). Some of the compounds have cytotoxic effects on bacterial growth, suggesting targets in *S. aureus* other than SrtA. These compounds showed inhibition on both SrtA and SrtB, suggesting a similar molecular mechanism or binding site shared by the two enzymes. However, flavonoid phenols as well as *β*-sitosterol-3-*O*-glucopyranoside exhibited at least 3-fold difference in the inhibition between SrtA and SrtB, which raises the possibility of active site difference between isoforms in the sortase family. Research continues to determine the mechanism of inhibition by these molecules so that the specificity and effectiveness can be understood and improved.

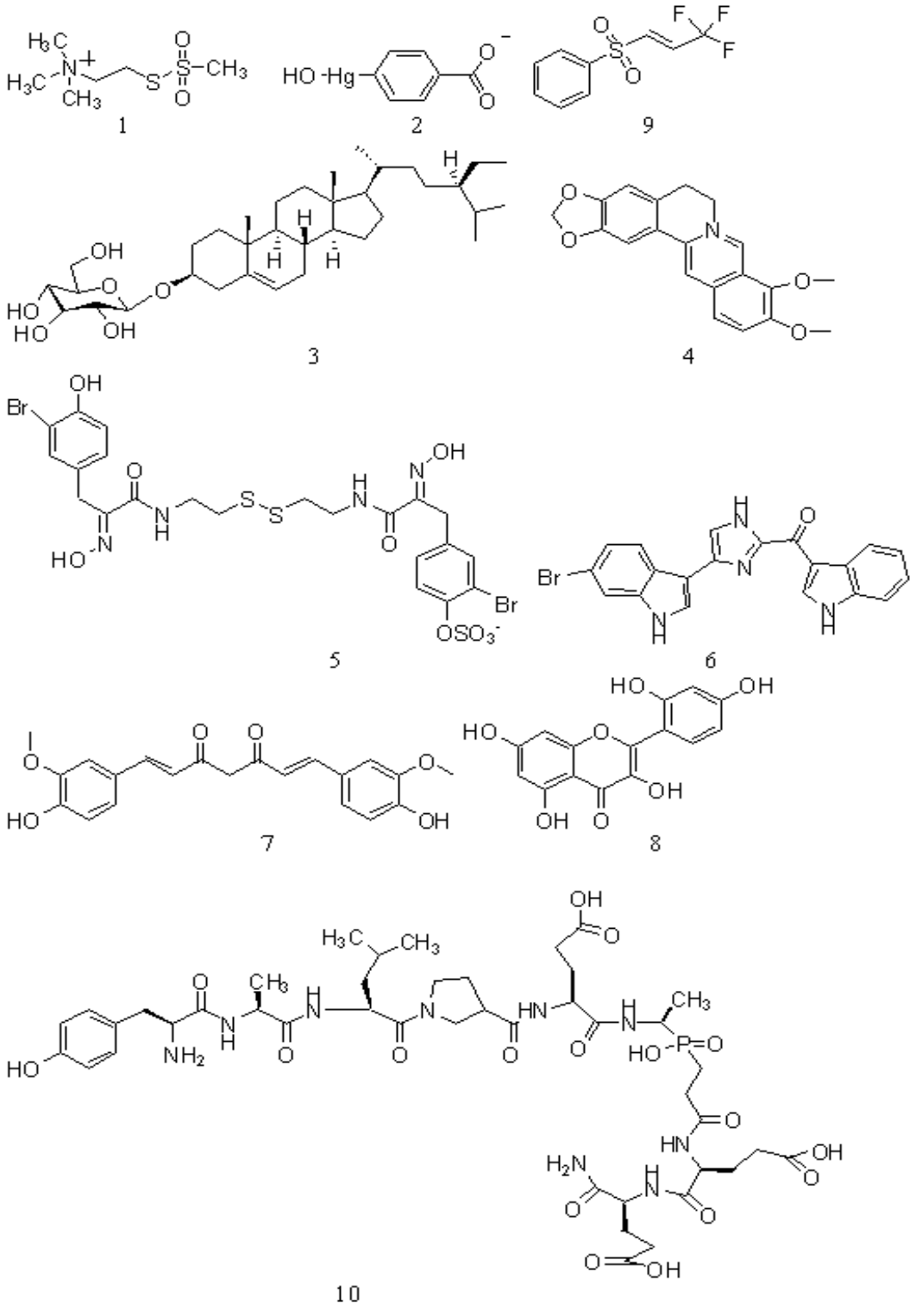

Figure 1.4 Inhibitors of SrtA.

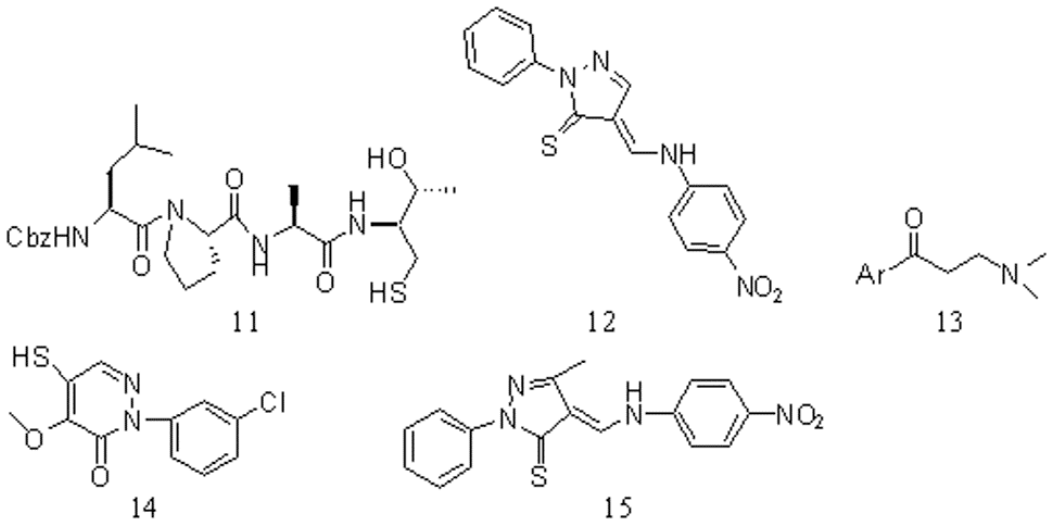

Figure 1.4 Cont.

Rational design of sortase specific inhibitors is based on the conserved cysteine at the active site or the characteristic substrate recognition sequence LPXTG. To target the nucleophilic thiol group at the active site, substrate or transition state derived reactive electrophiles were designed to make a covalent bond with Cys^{184}, irreversibly modifying the active site (*82*). For example, the sorting signal LPXTG derived inhibitors included the peptidomimetics with peptide bond between T and G replaced with a diazoketone or chloromethyl ketone, both of which displayed micromolar inhibition constants (*82*). Known electrophilic cysteine proteases inhibitors were also tested on SrtA, identifying a compound, 3, 3, 3-trifluoro-1-(phenylsulfonyl)-1-propene (9 in Figure 1.4) as an inhibitor of SrtA with a 190 μM IC_{50} (*83, 84*). A nonhydrolyzable peptidomimetic that imitates the transition state of sortase reaction, (NH_2-YALPE-Ala[PO_2HCH_2]Gly-EE-NH_2) (10 in Figure 1.4), was used in a kinetic analysis of inhibition patterns, which revealed a Ping-Pong Bi Bi catalysis model for SrtA, although the in vitro IC_{50} is only 10 mM (*47, 66, 85, 86*). Clubb's group has employed substrate analog inhibitors to map the molecular

mechanism of peptide binding motif on SrtA using solution NMR. Original trials used a partial recognition sequence, LPAT (11 in Figure 1.4) modified by cyanoalkene and sulfhydryl groups, as an inhibitor to identify the substrate-binding site at the hydrophobic groove formed by β4 and β7 (*87*). Further improvement fused this inhibitor to a threonine derivative at the C-terminus that contains a -CH_2-SH group to form disulfide bond with Cys^{184}, structurally mimicking the thioacyl intermediate (*57, 88*). A dynamic NMR structure of SrtA, markedly different from known structures, with a specific substrate bound was obtained, which significantly enhanced our knowledge of SrtA catalysis (*57*).

High-throughput screening with small molecule libraries was performed to identify SrtA inhibitors using an *in vitro* FRET assay. In 2004, the Oh group screened a library of 1000 diverse compounds, and identified a diarylacrylnitrile (12 in Figure 1.4) reversibly inhibit with a IC_{50} of 231 μM (*89*). Structure-activity relationship (SAR) analysis improved the IC_{50} to 28 μM and indicated the importance of a 2,5-dimethoxy group. The possible inhibition mechanism was suggested to be the interaction between the phenyl rings on the inhibitor and the lipophilic residues at the enzyme's substrate binding pocket. Another screen performed in 2007 by the Maresso group used a diverse small molecule library of 135, 000 compounds. The first screen for SrtA inhibition yielded a high hit rate (~4%). A set of 407 molecules was selected based on their "drug-like" characteristics. These molecules were subjected to a second screen aiming to eliminate those non-specific inhibitors, which can also inactivate a structurally related cysteine protease. The molecules that inhibited sortase family enzymes but not other thiol proteases were classified as specific. A class of aryl (β-amino) ethyl ketone (AAEK) (13 in Figure 1.4) inhibitors was selected due to its drug-like structure and high specificity. AAEK molecules are mechanism-based inhibitors, which are activated via β-elimination by sortase. The olefin intermediates generated covalently modify the Cys^{184} thiol group to

inactivate the enzyme with an apparent IC_{50} of 5-50 μM (*55*). The most recent screen was performed with a ~30, 000 compounds library by the Clubb group in 2009. Three potent molecules that reversibly inhibit SrtA were selected for further SAR analysis. This study has reported the most potent SrtA inhibitor identified thus far, with apparent inhibition constants of 1.4 μM for pyridazinone analogs (14 in Figure 1.4) and 0.3 μM for pyrazolethione analogs (15 in Figure 1.4). Those two classes of molecules also demonstrated sub micromolar IC_{50} inhibition on *B. anthracis* SrtA, suggesting the high specificity against SrtA enzyme (*57*).

The identified lead compounds based on *in vitro* activity assays will have to be scrutinized for *in vivo* efficacy. Indirect measurements of the sorting reaction, such as the adhesion of *S. aureus* to mammalian extra cellular proteins, the invasion of *S. aureus* into mammalian epithelial or endothelial cells can provide some evaluation on the effectiveness of inhibitors of virulence factors display. However, various other pathways in protein secretion, folding, and the cell wall biosynthesis can affect the inhibition readout as well. A direct assessment of sortase activity was measured by pulse immunoprecipitation and autoradiography (*65*). The disadvantage of this method is that it is time- and labor intensive, which is not applicable for high throughput screening (HTS). A possible protocol for HTS of inhibitors *in vivo* might be fluorescence-activated cell sorting (FACS) based on the abundance of a GFP substrate that successfully anchored on the cell wall. Nevertheless the final validation of inhibitors should be conducted with animal models of infection by gram-positive pathogens.

The research on cellular and molecular mechanisms of SrtA's biological function has suggested that it is a compelling target for anti-infective drugs that may have an advantage compared to conventional antibiotics due to the avoidance of direct selection pressure. In this thesis, I describe our investigation into the mechanism of this enzyme

and discuss about the significance of our discovery that SrtA exists in a dimer-monomer equilibrium on *S. aureus* cell's membrane, a fact of which plays a key role in SrtA catalysis but has not been recognized until our investigation.

Chapter 2 *Staphylococcus aureus* sortase A is a dimeric protein *in vitro*

INTRODUCTION

S. aureus virulence factor SrtA exists as a dimer *in vitro*

S. aureus, a facultatively anaerobic gram-positive coccus, can cause a wide spectrum of diseases, ranging from minor skin infections, to life-threatening illnesses such as endocarditis, pneumonia, septic arthritis, toxic shock syndrome (TSS) and sepsis (*1*). According to a CDC study, multidrug-resistant *S. aureus* caused infections tripled from 2000 to 2005 in the United States. An estimation of more than 2 million patients contract *S. aureus* infections from hospitals, causing 18,650 people die every year, even more than HIV (*24*). A critical step in *S. aureus* pathogenesis is the adherence of bacterium to host tissues and the subsequent internalization, which is mediated by a collection of adhesins on *S. aureus* cell surface, called microbial surface components recognizing adhesive matrix molecules (MSCRAMMs) (*9-13*). Many of these surface proteins possess a C-terminus sorting signal consisting of a LPXTG motif, a hydrophobic transmembrane segment, and followed by positively charged residues, to the peptidoglycan cell wall (*13*). They are anchored to the cell wall by a 206 amino acid, membrane-bound cysteine transpeptidase, known as sortase A (SrtA) (*39*).

The *S. aureus* SrtA is the prototypical member of the sortase enzyme family conserved in Gram-positive bacteria (*40*). The knockout mutants of SrtA display severe defects in virulence in *S. aureus* (*90-92*), *Listeria monocytogenes* (*37, 93*), *Streptococcus gordonii* (*36*), *Streptococcus pneumonia* (*94*), *Streptococcus pyogenes* (*95*), *Streptococcus mutans* (*96*), *Streptococcus suis* (*97*), *Streptococcus sanguinis* (*98*), and *Streptococcus agalactiae* (*99*), suggesting it as an general drug target for gram-positive

pathogens. Given the interest in designing highly selective inhibitors for SrtA, several classes of molecules that target the active site of the enzyme have been described *in vitro* (*39, 55, 63, 76-78, 82, 83, 86-89*). *In vitro* studies of the *E. coli* expressed and purified SrtA catalytic domain ($SrtA_{\Delta N59}$, residues 60–206) have proposed a possible mechanism of transpeptidation through a Ping-Pong Bi Bi model (*44, 47, 50, 51, 59-65*). However, some discrepancy in the kinetic results from different research groups suggest that there are unsolved problems in this mechanism (*49, 62, 64, 65, 85*). For instance, based on a reverse protonation mechanism, SrtA with uncharged state of active site residues His^{120} and Cys^{184} at physiological pH is in inactive form, which accounts for about 99.94% of the population (*47*).

In this chapter, I described our discovery that the catalytic domain of SrtA exists as a dimer *in vitro* with a moderate K_d of 55 μM, and both the monomer and dimer fractions of the protein are active *in vitro* (*100, 101*). The dimerization between two monomers is highly specific in the context of the entire cellular protein mixture.

It is of interest to understand what physical forces drive the process of dimerization at the molecular level. The interactions between proteins are of central importance in many biological functions. Based upon the current database of protein-protein interactions (PPIs) identified, the association of proteins is intrinsic to virtually every cellular process that maintains life. Inappropriate PPIs would lead to the malfunction of proteins and enzymes, in some cases even the death of cells. It has been established in some membrane proteins and proteases that catalytic function is dependent on dimerization. For example, Outer membrane phospholipase A (OMPLA) of gram-negative bacteria has moderate binding affinity for its homo-dimerization *in vitro* and when present on the outer membrane, dimerization is induced to activate the enzyme (*102*). The catalytic ability of the Epstein-barr virus protease has previously been shown

to be dependent on the extent of dimerization (*103*) as well. An understanding of the SrtA oligomerization state may aid the rationale design of molecules to regulate the extent of dimerization and possibly the enzyme activity as well.

Characterization of protein-protein interactions

The knowledge of protein-protein interactions, including thermodynamic, kinetic and structural properties are important to understand protein interaction networks and their evolution, which may ultimately help to accurately predict novel interactions and re-engineer existing PPIs (*104*). A wide variety of methods have been developed in the past decade to study protein complexes, by firstly detecting the PPIs, analyzing subunit composition, quantifying thermodynamic and kinetic parameters, and deciphering the molecular details of the binding interfaces. Each method has its own characteristics and limitations, therefore, extensive analysis using sets of different approaches is required to generate a comprehensive understanding of individual protein complexes and protein interaction networks (*105*). The methods used to detect SrtA dimerization and quantify the association affinity are introduced in the following paragraphs.

Gel filtration chromatography

Gel filtration chromatography is a simple way to detect and estimate protein-protein binding constants. In gel filtration, the separation of macromolecules and complexes is based on their differences in Stokes' radii (*106*). The column resin is made up of cross-linked polymers forming different pore sizes. Larger proteins or protein complexes having larger Stokes' radius are less likely to go through smaller particle pores and are eluted directly and earlier. Smaller protein complexes or protein monomers can pores and are eluted later. This provides a conceptually straight-forward technique to

detect protein-protein interactions and to characterize the strength of binding between two different proteins.

Gel filtration can be analyzed by three distinct methods to evaluate the affinities of protein-protein interaction. Nonequilibrium small zone gel filtration chromatography is a simple approach containing the protein, the protein binding partner, and their mixture is eluted individually on the same column (*107*). The complex should elute later because of its relatively larger size. Thus, the binding affinity can be estimated by measuring the concentration dependence of protein species required to form the complex. This approach has been successfully used to measure the binding affinity between cAMP-dependent protein kinase with R1 and PKI subunit (*108*). However one disadvantage is that the binding affinity is sometimes underestimated, especially in a fast-equilibrium system (*109*). To overcome this problem, a computer simulation has been developed and used to study the rapidly equilibrating antibody-antigen interactions (*107, 110-112*).

The Hummel-Dreyer method is an equilibrium method based on the assumption that the rate of stabilization of the protein-protein complex is much higher than the speed of chromatographic separation (*113*). In this method, the column is first equilibrated with column buffer containing the same concentration of protein ligand as the sample. When the sample with protein and ligand mixture is applied and run through the column, the concentration of protein ligand will increase where the protein is eluted. The binding affinity is evaluated as a function of equilibrium equation by measuring the free concentration of protein ligand, the concentration of ligand bound with protein and the protein concentration eluted. This method was only applied to very few cases, such as the interaction between glycerol-3-phosphate dehydrogenase and lactate dehydrogenase (*114*) and the interaction between CheW and CheA (*115*). The main weakness of this method is that upon running through the column, the sample containing the protein

complex becomes gradually diluted. Therefore, only an apparent binding affinity value is obtained (*116*).

Large-zone equilibrium gel filtration is another equilibrium method in which, the sample is loaded on a column in very large volume. Although this method is thermodynamically rigorous, it has only received limited attention since significant amount of protein sample is consumed, which is a disadvantage when materials are of limited supply (*107*).

Analytical ultracentrifugation

Analytical ultracentrifugation was first developed in the late 1920s and had been one of the early approaches to characterize proteins and protein complexes in solution. Later in the early 1990s, the instrument has been improved with digital data acquisition and better optics (absorption and interference). Ever since then, analytical ultracentrifugation has become a powerful tool for the investigation of PPIs in ordinary and a wide range of unordinary solutions (*117*).

In addition to a high speed centrifuge and a rotor with cell compartments, the analytical ultracentrifuge instrument has an optical system to measure the concentration gradients of proteins spun under high centrifuge force. The protein samples with different concentrations are loaded into centerpieces with quartz or sapphire windows on both sides so that light can go through the centerpiece to achieve either absorption or interference optical data. An experiment can operate in at least two different modes: sedimentation equilibrium and sedimentation velocity. The sedimentation equilibrium experiment is usually conducted at relatively low centrifuge speed. It takes longer time to complete. Under the centrifugal field, the sedimentation of proteins is balanced by diffusion opposing the concentration gradients. Eventually, an equilibrium would be established, which can be described by thermodynamic theory. The fitting of data by

Boltzmann distributions are widely used to determine the absolute solution molecular weight, association equilibrium constants for monomer-multimer self-associating systems and heterogeneous interacting systems as well (*117*). The sedimentation velocity experiment is completed at high speeds to measure the time-course of sedimentation. Each protein species, with different molecular mass and shape, will sediment at a characteristic speed. The velocity and shape of the moving boundary formed by individual species can then be used to determine the overall size and hydrodynamic shape of protein and protein complexes with high reliability and accuracy (*118*). With recent advances in data acquisition and analysis, analytical ultracentrifugation is routinely applied to study macromolecular systems. However, the limitations also exist in that it is difficult to accurately detect low levels of protein aggregates with weak and transient binding. In addition, the use of analytical ultracentrifugation requires extensive professional training on both the instrument and the data analysis (*119*).

RESULTS AND DISCUSSION

Expression of cloned SrtA

To obtain *Staphylococcus aureus* SrtA protein for *in vitro* study, the SrtA gene was first cloned in pET vector, and then over-expressed and purified from *E. coli*. A His6 tag was fused to C-terminus of the protein to facilitate the purification of the protein using nickel beads. Both the full-length and the truncated form of SrtA were generated for further investigation. The truncated SrtA protein ($SrtA_{\Delta N59}$), which lacks the N-terminal 59 amino acids, was easier to express and get purified than full-length protein. The N-terminal lipophilic region of SrtA is believed to function as a signal peptide for peptide secretion and a stop-transfer signal for retaining in membrane. It has been

demonstrated that $SrtA_{\Delta N59}$ generated and analyzed *in vitro* can catalyze the same transpeptidation reaction as the full-length SrtA (*44, 50*).

SrtA forms a dimer *in vitro*

$SrtA_{\Delta N59}$ was purified under the denaturing conditions, and analyzed by SDS-PAGE. When stained with Coomassie Blue, there appeared to be two bands, one intense band around 20 kDa and the other less intense band near 45 kDa (Figure 2.1 A, lane 1). The calculated molecular weight of a $SrtA_{\Delta N59}$ monomer based on amino acid sequence is 17.9 kDa, while that for a homo-dimer is 35.8 kDa. This data suggested that the two bands observed on SDS-PAGE represented monomeric and dimeric forms of $SrtA_{\Delta N59}$. To provide further evidence for the identities of the bands, Western blot analysis against anti His6 was also performed. As demonstrated in Figure 2.1 A, a His6 tag was existing in proteins of both bands, indicating C-terminal His6 tagged $SrtA_{\Delta N59}$ in both bands. The 40 kDa band matched the molecular weight of either a $SrtA_{\Delta N59}$ homo-dimer, or a $SrtA_{\Delta N59}$ hetero-dimer associating with another, as yet unidentified protein with similar mass. Full-length $SrtA_{WT}$ was also analyzed under the same denaturing experimental conditions and revealed similar results, suggesting the either homo- or hetero-dimerization of SrtA protein (Figure 2.1 B).

Truncated SrtA, $SrtA_{\Delta N59}$ was also purified under a nondenaturing condition and analyzed by native polyacrylamide gel electrophoresis (Native-PAGE) as shown in Figure 2.1 C (lanes 1 and 2). Different from the analysis results achieved under the denaturing condition, when stained with Coomassie Blue, four protein bands were observed in Figure 2.1 C, lane 1. Anti-His6 antibody visualized western blot indicated that all four protein bands contain C-terminal His6 tagged $SrtA_{\Delta N59}$ (Figure 2.1C, lane 2). We speculated that the four protein bands respectively represented the monomeric, dimeric, trimeric and tetrameric forms of $SrtA_{\Delta N59}$ (Figure 2.1C, lanes 1 and 2). Due to

the higher intensity of the putative dimeric band, it was presumed to be the preferred state for $SrtA_{\Delta N59}$ under non-denaturing conditions.

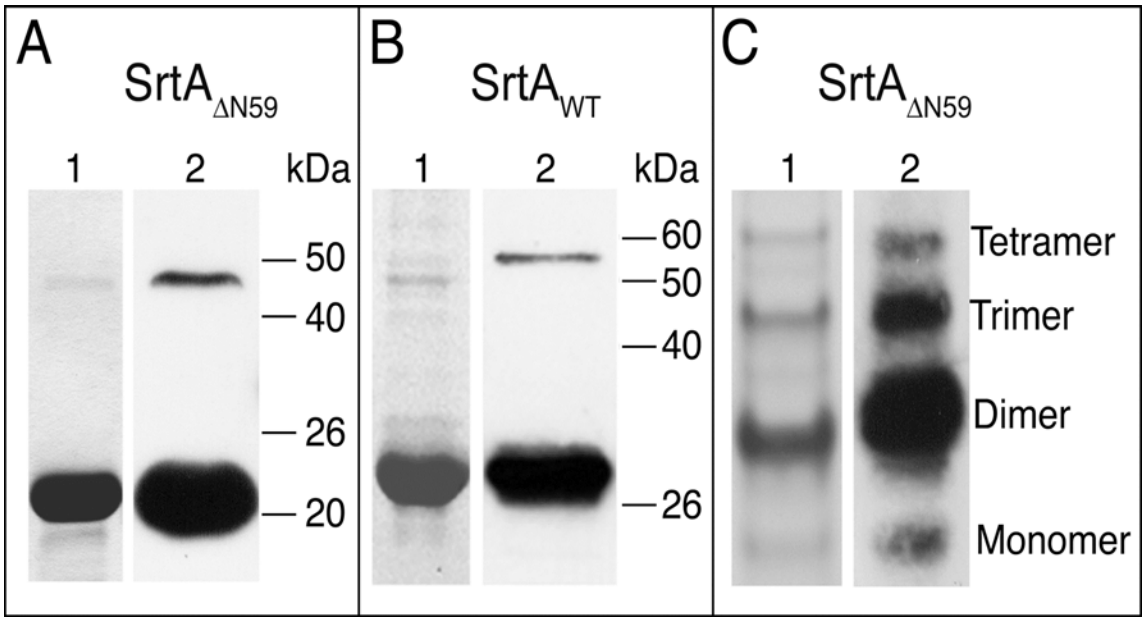

Figure 2.1: $SrtA_{\Delta N59}$ and $SrtA_{WT}$ Proteins expression.

Purified $SrtA_{\Delta N59}$ (*A, C*) and $SrtA_{WT}$ (*B*) proteins were visualized by Coomassie Blue staining (lanes 1) and western blot (lanes 2) using an anti-His6 antibody. Denaturing conditions were applied in panels *A* and *B*, while native conditions were applied in pane l *C*.

Confirmation of dimerization *in vitro*

We next determined that $SrtA_{\Delta N59}$ was the only protein component in the 40 kDa, higher molecular weight band on SDS-PAGE. The protein from this putative dimeric band was extracted from the gel and directly sequenced. A single protein sequence was identified exclusively. The N-terminal first 7 amino acids of this sequence were GQAKPQI, identical to that of $SrtA_{\Delta N59}$. We further analyzed the band by protein in-gel digestion and subjected the digestion sample in peptide fingerprint mapping by mass spectroscopy. The single protein identity detected from the assay has its fingerprint matching with the known primary sequence of $SrtA_{\Delta N59}$ with a confidence value of 83%

(NCBI Protein Database: gi l4277822, Chain A, Structure of Sortase). The homo-dimerization of $SrtA_{\Delta N59}$ can be observed even under denaturing conditions. Addition of various concentrations of dithiothreitol (DTT) was not able to inhibit the dimer formation on a denaturing polyacrylamide gel as well (data not shown), suggesting that the homo-dimer of $SrtA_{\Delta N59}$ is not covalently linked through the disulfide bond formed by the only cysteine residue in protein sequence, the active site Cys^{184}, exposed on protein surface.

Next, size-exclusion chromatography was used to detect the oligomerization of $SrtA_{\Delta N59}$ under native conditions. Three peaks were observed after the $SrtA_{\Delta N59}$ from non-denaturing purification was analyzed on a Superdex 200 gel filtration column (Figure 2.2, peaks *a*, *b* and *c*). The two major peaks have estimated molecular weights of 41 kDa (peak *b*) and 26 kDa (peak *c*) respectively, which correspond to a homo-dimer and monomer fraction of $SrtA_{\Delta N59}$. The eluted fractions containing peaks *b* and *c* were then digested with trypsin, and subjected to peptide mass fingerprinting analysis. The sequence of protein components in these two fractions matched the $SrtA_{\Delta N59}$ sequence with confidence limits of 71% and 99%, respectively (data not shown). It confirms that the major peak b (41 kDa) is composed of the homo-dimeric form of $SrtA_{\Delta N59}$, while peak c (26 kDa) is the monomeric form of the protein. A third, smaller protein peak a has a calculated molecular weight of 56 kDa, which suggests that this peak consists of either a trimeric form or tetrameric form of $SrtA_{\Delta N59}$. Based on the elution profile of $SrtA_{\Delta N59}$ on gel filtration column, the dimeric form of the protein predominates, which is about 62% of the total protein (peak *b*). About 27% of the molecules are monomer (peak c), and about 11% is trimer or tetramer. The protein sample had been incubated over night at 4˚C before analysis on gel filtration column to attain the monomer-dimer equilibrium. When a freshly prepared $SrtA_{\Delta N59}$ sample was used under the same chromatographic condition, the majority of the protein were monomer, which is about 50%, while there were only

about 45% of the protein existed in the dimeric form and less than 5% in the trimeric or tetrameric form.

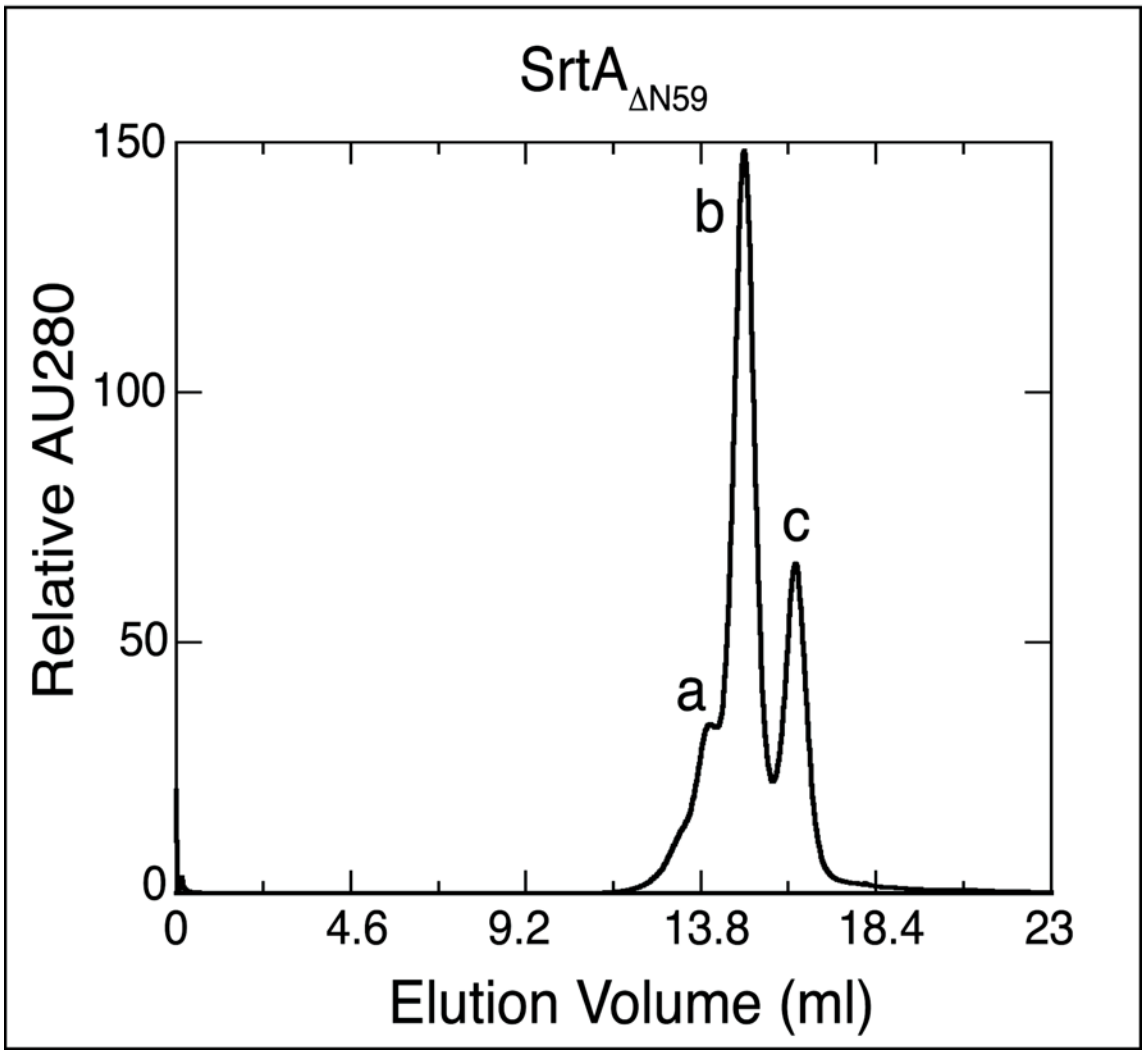

Figure 2.2: Size-exclusion chromatography of SrtA$_{\Delta N59}$ protein.

SrtA$_{\Delta N59}$ protein was purified under native conditions and incubated at 4°C over night before analyzed by size-exclusion chromatography.

SrtA can be cross-linked *in vitro*

To further demonstrate the selectivity of SrtA to form a dimer, *in vitro* chemical cross-linking experiments were also carried out. Purified $SrtA_{\Delta N59}$ and $SrtA_{WT}$ proteins under native conditions were subjected to cross-linking reagent, 1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (EDC), a compound that catalyzes the covalent link between a carboxylic acid group and a primary amine (*120*). When the products of EDC treatment were subjected to SDS-PAGE and western blot analyses (Figure 2.3 A & B, lane 2), a much more intense dimer band was observed, compared to the dimer band from uncross-linked sample (Figure 2.3A & B, lanes 1). Extra higher molecular weight bands in addition to the dimer bands were observed in both the $SrtA_{\Delta N59}$ and $SrtA_{WT}$ samples after EDC cross-linking treatment. In $SrtA_{\Delta N59}$ sample, this extra band has an estimated molecular weight of around 72 kDa, but was not detected in the uncross-linked sample (Figure 2.3A). While in $SrtA_{WT}$ sample, EDC catalyzed cross-linking reaction resulted in a band above 95 kDa other than the monomeric and dimeric bands (Figure 2.3B). Based on their apparent molecular weights from migration pattern on SDS-PAGE, the two bands may be the tetramers for SrtA proteins. This data also verifies the previous observation of SrtA tetramer formation analyzed under non-denaturing conditions, which may be the association of two SrtA homo-dimer complexes.

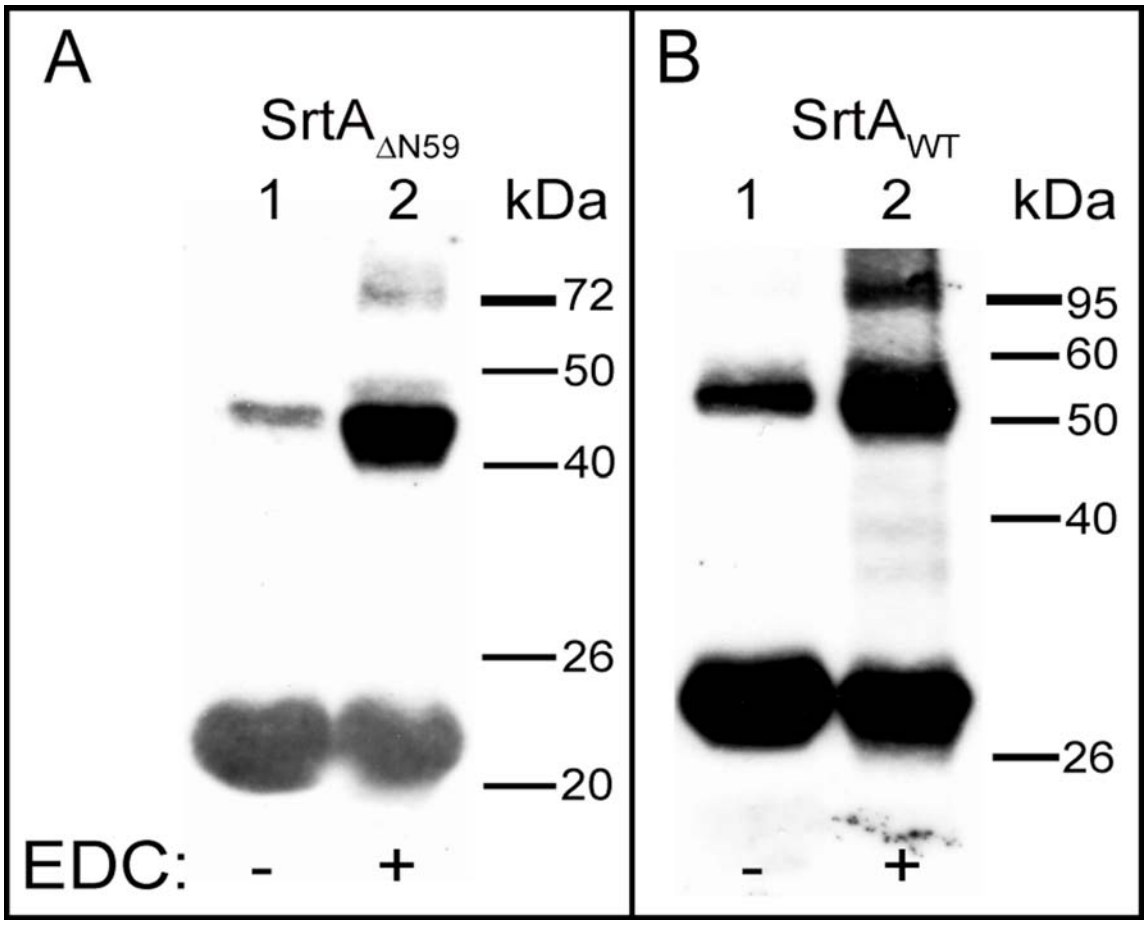

Figure 2.3: *In vitro* EDC mediated cross-linking.

$SrtA_{\Delta N59}$ (panel *A*) and $SrtA_{WT}$ (panel *B*) proteins were subjected to EDC mediated chemical cross-linking *in vitro*. Both cross-linked (lanes 2) and uncross-linked (lanes 1) samples of protein were then analyzed by SDS-PAGE and western blot using an anti-His6 antibody. Equal amount of protein was loaded into each lane.

We repeated cross-linking experiments using another, more efficient and photo-reactive reagent, Sulfo-HSAB (*121*). To prove that the cross-linking between two molecules in dimer is highly specific, egg albumin, a protein that does not normally dimerize, and lysozyme, which has a strong tendency to form dimer were used in this experiment (*122, 123*). After Sulfo-HSAB treatment, the homo-dimer bands were observed in $SrtA_{\Delta N59}$ and lysozyme sample as expected, however, which was not apparent

in egg albumin sample (Figure 2.4). These experiments again support the assumption that SrtA protein forms a homo-dimer naturally.

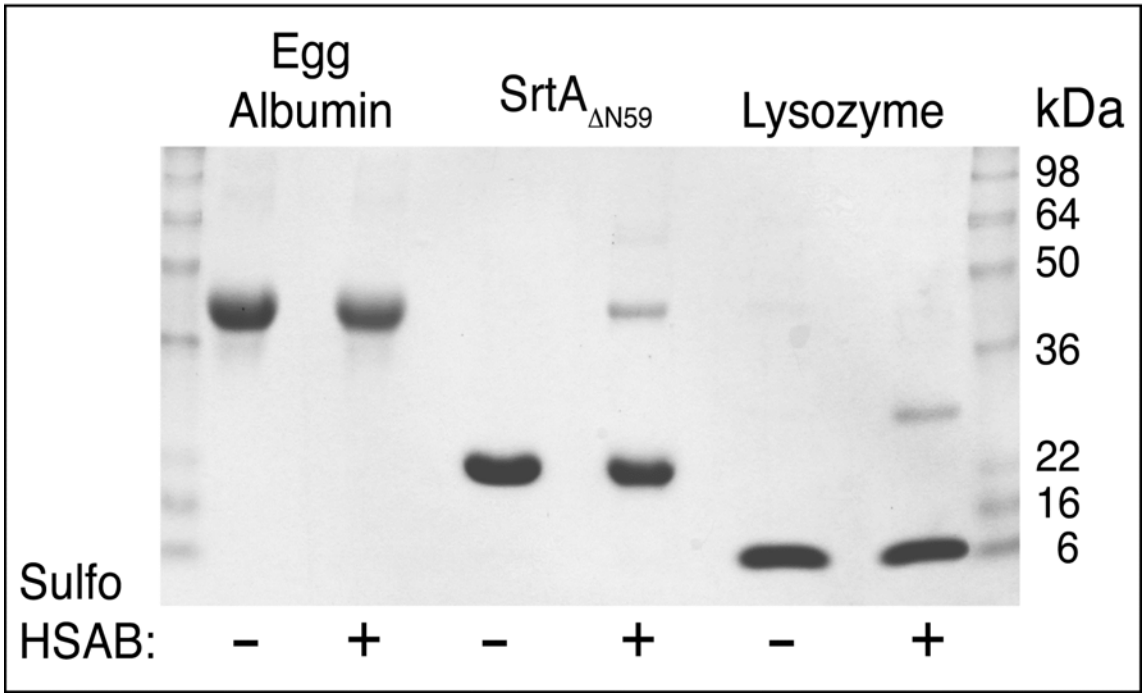

Figure 2.4: *In vitro* Sulfo-HSAB mediated cross-linking.

Egg albumin, SrtA$_{\Delta N59}$, and lysozyme proteins were subjected to Sulfo-HSAB mediated chemical cross-linking, *in vitro*. Both cross-linked (+ lanes) and uncross-linked (- lanes) samples of protein were then separated by SDS-PAGE and visualized by Coomassie Blue staining.

SrtA$_{\Delta N59}$ selectively forms a homo-dimer in *E. Coli*

We also demonstrated that SrtA can form dimer and be cross-linked in *E. coli* cytoplasm. We briefly treated an *E. coli* culture, which expresses SrtA$_{\Delta N59}$, with EDC by adding the reagents directly into cell culture (*124, 125*). SrtA$_{\Delta N59}$ was then purified and examined by SDS-PAGE and Coomassie Blue staining (Figure 2.5 A). Similar results were observed as seen in the *in vitro* experiments: the molecular weights of two bands apparent on SDS-PAGE gel corresponded to a monomeric and dimeric form of SrtA$_{\Delta N59}$.

The dimer band was subjected to in-gel trypsin digestion with trypsin before peptide mass fingerprinting, which suggested with a confidence value of 59% that the only protein present in the sample has fingerprinting map matched with the sequence of $SrtA_{\Delta N59}$ (data not shown). $SrtA_{WT}$ was also expressed in *E. coli* and treated with EDC for cross-linking before harvesting. As shown in Figure 2.5 B, $SrtA_{WT}$ also has two bands corresponded to the size of full-length SrtA monomer and homo-dimer.

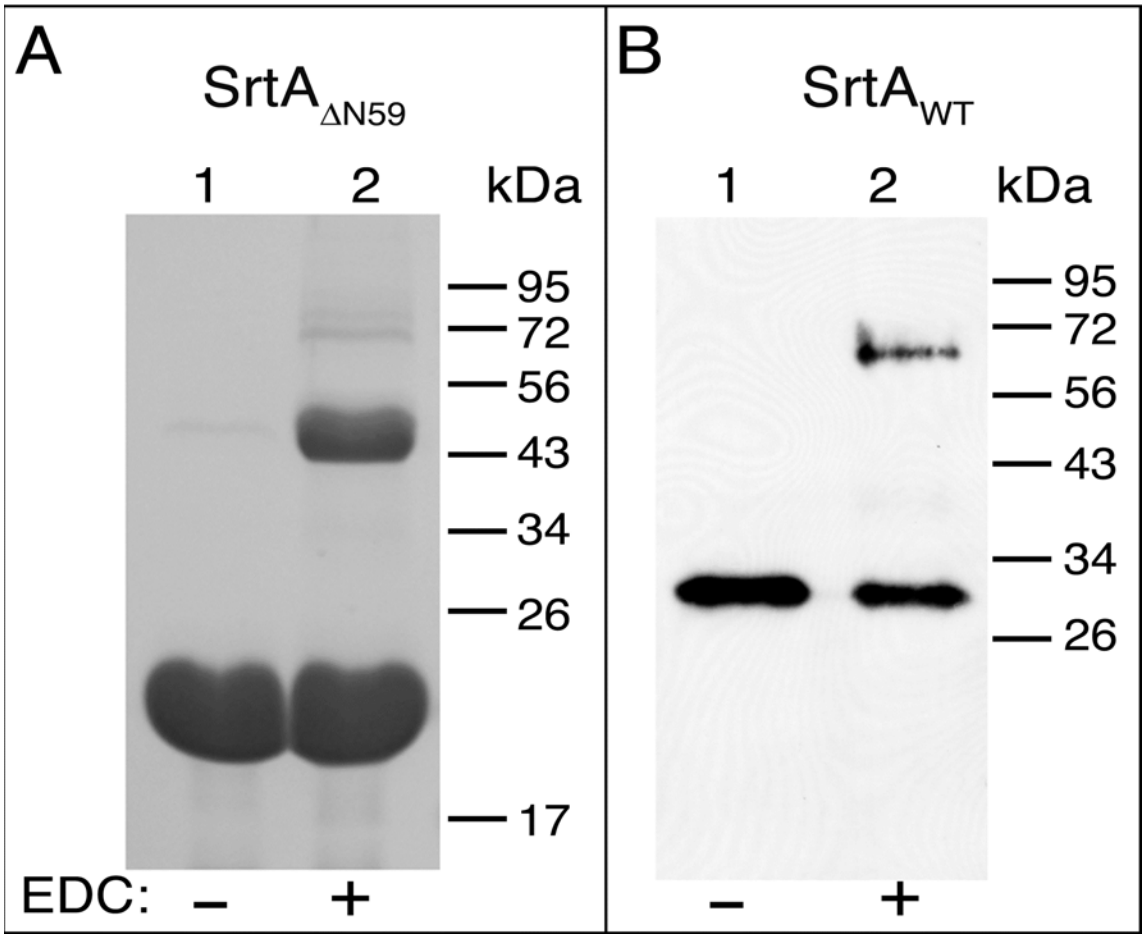

Figure 2.5: EDC-mediated cross-linking in *E. coli.*

$SrtA_{\Delta N59}$ (panel *A*) and $SrtA_{WT}$ (panel *B*) proteins purified from cells after EDC treatment (lanes 2), as well as that lacking EDC treatment (lanes 1), were separated by SDS-PAGE. For purified $SrtA_{WT}$, western blot with an anti-His6 antibody was necessary. Within each panel, equal amounts of protein were loaded into lanes 1 and 2.

Determination of the dissociation constant (K_d) of the SrtA dimer

In order to quantitatively analyze the monomer-dimer equilibrium with $SrtA_{\Delta N59}$, *in vitro* purified protein sample was subjected to an analytical equilibrium sedimentation ultracentrifugation experiment. Representative data collected at different equilibrium state are shown in Figure 2.6 with repeated scans at various time points. Two different protein concentrations of $SrtA_{\Delta N59}$, 26 μM and 52 μM, were used spanning three different speeds in the experimental set up. Collected data points were globally and simultaneously fitted with a monomer-dimer equilibrium model (Figure 2.6 B) in an attempt to give an estimation of dissociation constants. Average dissociation constants (K_d) of 56 ± 7 μM was achieved from 26 μM, and 54 ± 7 μM from 52 μM. Combining the data gives final average K_d of 55 ± 7 μM. This K_d number of 55 μM suggests a moderate binding affinity between $SrtA_{\Delta N59}$ molecules *in vitro*. Some other enzymes under the regulation of dimer-monomer equilibrium have much higher association affinities, with K_d values range from nM for HIV-1 protease (*126*) to high μM for caspase 9 (*127*). However, the measurement of SrtA sample was performed in a dilute solution with the truncated form of enzyme. *In vivo* expressed $SrtA_{WT}$ on *S. aureus* cell membrane may actually have higher tendency to form a dimer, due to its effective concentration on focal site on membrane as well as its tether to cell membrane.

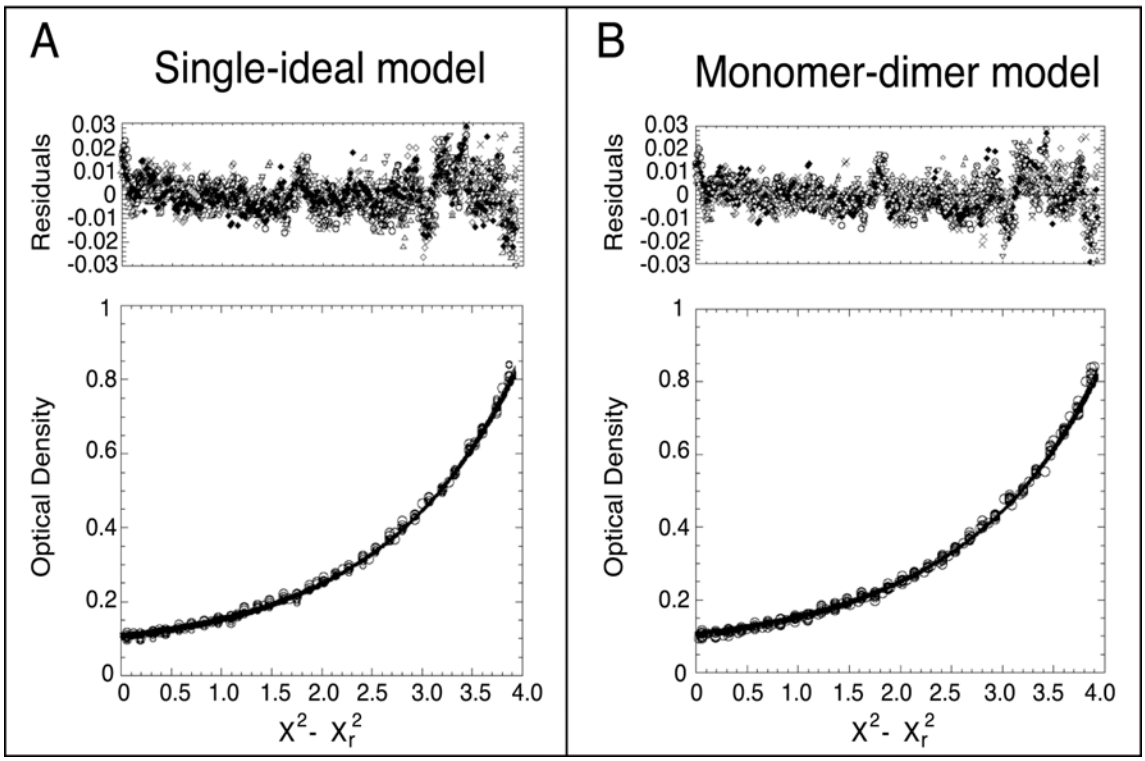

Figure 2.6: Analytical sedimentation equilibrium ultracentrifugation.

$SrtA_{\Delta N59}$ protein was purified under native conditions and subjected to analytical sedimentation equilibrium ultracentrifugation under various conditions. The data was subjected to a single-ideal model (panel *A*) and a monomer-dimer association model (panel *B*).

Kinetic studies of monomeric and dimeric SrtA

In order to determine whether the dimerization can regulate the catalytic function of SrtA *in vitro*, we performed the activity assay to analyze both the monomeric and dimeric forms of $SrtA_{\Delta N59}$ to obtain kinetic constants using a developed fluorescent quenching assay (*44, 48, 49, 64*).

SrtA catalyzes the transpeptidation in two sequential steps following a ping-pong bi-bi mechanism (*49*). The first step is the recognition and hydrolysis of the primary substrate LPXTG sequence and the generation of an acyl-enzyme. In the following step,

the secondary substrate, pentaglycine was ligated into the hydrolyzed LPXT sequence. The first step of LPXTG hydrolysis is believed to be rate-determining step in the reaction (*49*). Therefore, our *in vitro* activity assay has been focused on obtaining the apparent $k_{cat,\ app}/K_{m,\ app}$ of this step but not the complete transpeptidation reaction. Numerous assays have been developed and employed in different research groups, such as the fluoresce quenching assay and HPLC assays (*49, 62, 64, 65, 85*). Although both of the assays measure the hydrolysis a synthetic peptide substrate, the reported kinetic constants fall into a surprisingly wide range, with 8.2 μM to 5.5 mM for K_m, 1.2×10^{-4} s^{-1} to 0.57 s^{-1} for k_{cat}, and 6×10^{-3} $mM^{-1}s^{-1}$ to 4.91 $mM^{-1}s^{-1}$ for k_{cat}/K_m (Table 2.1). We have chosen the fluorescent quenching assay to avoid the usage of high concentrations of the peptide substrate in HPLC method, which may have a problem of precipitation. The synthetic peptide substrate Abz-LPETG-Dap(Dnp) and 2 mM pentaglycine as secondary substrate were used for enzyme kinetics. Abz-LPETG-Dap(Dnp) consists of the LPXTG SrtA recognition motif with modification of a fluorescent group (Abz) at the N-terminus, and a quencher (Dap(Dnp)) at the C-terminus. Saturated concentration of synthetic pentaglycine was included in the reaction as the nucleophile, substituting for the pentaglycine cross-bridge of the cell wall. Under these *in vitro* conditions, SrtA catalyzes the cleavage of Abz-LPETG-Dnp between T and G, releasing a fluorescent signal, and generating the final product Abz-LPET-GGGGG. The released fluorescent signal was monitored in the first 10 minutes after the initiation of reaction with different amount of peptide substrate. Fitting of the velocity of fluorescence enhancement versus peptide concentrations to equation 4 has generated the K_m and k_{cat} for both the monomeric and EDC-treated dimeric forms of $SrtA_{\Delta N59}$. To obtain a dimer-enriched protein sample, we first stabilized the SrtA dimer by chemically cross-linking $SrtA_{\Delta N59}$ using EDC *in vitro*. The mixture of proteins containing covalently stabilized dimer was then separated by

size-exclusion chromatography under a native condition to individually collect the monomeric and dimeric forms. However, due to the equilibrium between two forms, after many rounds of purification by gel filtration, there was always uncross-linked protein in the collect dimeric fraction. According to the analysis by denaturing SDS-PAGE, the putative dimeric fraction contained only about 20% covalently linked SrtA dimer. Therefore, the dimeric and monomeric forms of enzyme used in the assay were the two fractions eluted from gel filtration column. Assay results are shown in Figure 2.7 and summarized in Table 2.1.

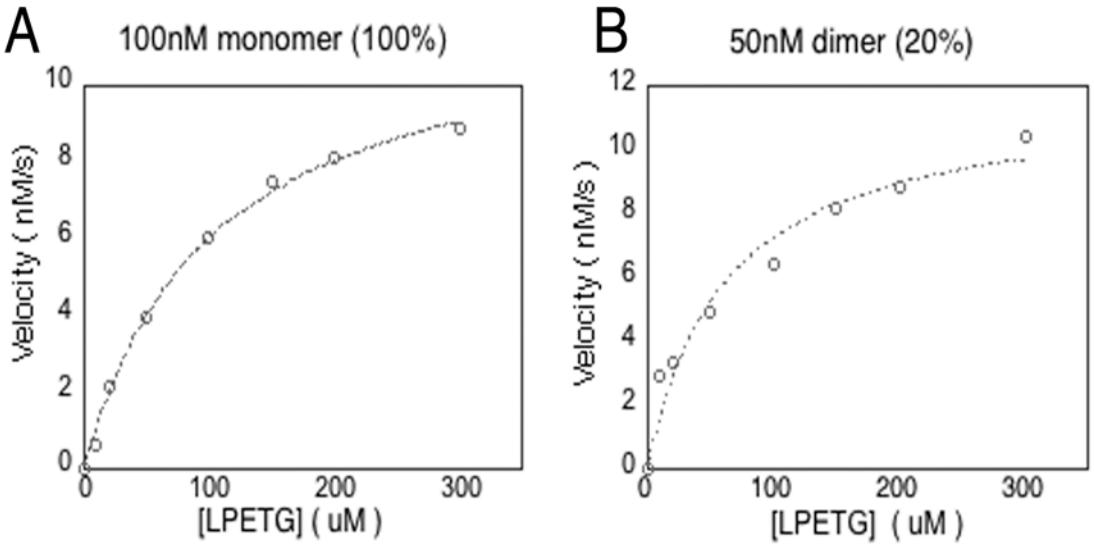

Figure 2.7: Steady-state rate plotted against the concentration of Abz-LPETG-Dap(Dnp).

Purified $SrtA_{\Delta N59}$ was subjected to EDC cross-linking, and then separated by size-exclusion chromatography. A monomeric fraction containing 100% monomer of $SrtA_{\Delta N59}$ and a dimer-enriched fraction containing 20% $SrtA_{\Delta N59}$ dimer were collected. Each fraction was then used to catalyze the SrtA transpeptidation reaction using Abz-LPETG-Dap(Dnp) as its substrate. Kinetic constants were determined using the Michaelis-Menten equation, where the rate is plotted against Abz-LPETG-Dap(Dnp) concentration, for both the monomeric fraction (panel A) and the dimeric fraction (panel B).

Table 2.1 Kinetic parameters for the dimeric and monomeric fractions of $SrtA_{\Delta N59}$.

Enzyme fraction	Monomer (100%)	Dimeric (20%)	Published data
$K_{m,\ app}$ (μM)	100.3 ± 11.3	61.2 ± 17.0	8.2 - 5500 [29, 43-46]
$k_{cat,\ app}$ (s^{-1})	0.121 ± 0.01	0.117 ± 0.01	1.2×10^{-4} - 0.57 [29, 43-46]

Duplicate data sets for each experiment were used to calculate the steady-state velocity at different Abz-LPETG-Dap(Dnp) concentrations for both the dimeric and monomeric fractions of $SrtA_{\Delta N59}$ enzyme after EDC cross-linking and FPLC purification. The initial velocity was then normalized and plotted as in Figure 2.7. $K_{m,\ app}$ and $k_{cat,\ app}$ values were determined by a double-reciprocal plot of the substrate dependence of velocity using Equation 2.4.

Based on the *in vitro* kinetic assay, a $k_{cat,\ app}$ value of 0.121 ± 0.01 s^{-1} for the monomeric SrtA and 0.117 ± 0.01 s^{-1} for the dimer-enhanced fraction were obtained (Table 2.1). Schneedwind *et al*. has reported a k_{cat} value of 0.57 s^{-1} in their kinetic assay using the same setup, which is in the same order of magnitude as ours (*64*). However, Ellestad and coworkers has also reported a k_{cat} value of 0.096 min^{-1} under the same conditions (*49*). It is not clarified yet why such discrepancy occurs (*49*). In our study, the only important purpose is to determine if there is any difference in the $k_{cat,\ app}$ values between the monomeric and dimeric fractions of $SrtA_{\Delta N59}$ (Table 2.1).

The $K_{m,\ app}$ values obtained for the dimer-enhanced fraction and the monomeric fraction of $SrtA_{\Delta N59}$ are 61.2 μM and 100.3 μM respectively. The $K_{m,\ app}$ value for monomeric enzyme is comparable with reported data, a 116 μM K_m from Schneewind and coworkers (*64*) and a 117 μM K_m from Ellestad and coworkers (*49*). The concentration of $SrtA_{\Delta N59}$ enzyme used in the reported assays was 1.5 μM. This is well below the K_d of 55 μM as we measured *in vitro*. At this concentration, the majority of $SrtA_{\Delta N59}$ exists in monomeric form. As shown in our assay, the dimer-enhanced fraction

of $SrtA_{\Delta N59}$ has a slightly smaller K_m than the monomeric fraction. Although K_m does not exactly equal K_d, it measures the dissociation constant for all species bound with enzyme (*128*). The smaller K_m observed suggested that the SrtA dimer might have better binding affinity toward the substrate than SrtA monomer. A possible explanation would be that EDC cross-linked dimeric $SrtA_{\Delta N59}$ adopts has a different conformation than monomeric $SrtA_{\Delta N59}$, which is more favorable for substrate binding (*62*).

The apparent $k_{cat,\ app}/K_{m,app}$ values for the transpeptidation reaction were also calculated and compared between the monomeric fraction of $SrtA_{\Delta N59}$ and dimer-enhanced fraction. The value of 1.21 $mM^{-1}\ s^{-1}$ for monomeric form and 1.91 $mM^{-1}\ s^{-1}$ for dimer-enhanced form are slightly lower than the published data of 4.91 $mM^{-1}\ s^{-1}$ using the same fluorescent assay (*64*). However, the difference between the two different SrtA conformations is not significant.

CONCLUSIONS

Due to SrtA's critical role in MRSA pathogenesis, research to understand its catalytic mechanism has intensified. Several small molecule inhibitors have been identified with submicromolar IC_{50} values *in vitro*. None of them have been confirmed with appreciable inhibition efficacy *in vivo*, suggesting the existence of missing points in current reaction model.

Our studies described in this chapter indicate that the SrtA protein selectively forms a homo-dimer *in vitro*. The high selectivity is demonstrated by specific cross-linking between two SrtA molecules even in a complex mixture of proteins, such as in *E. coli* cytosol. Although the dissociation constant is in the middle-μM range and the kinetic constants obtained from *in vitro* activity assay indicated little or no difference between SrtA dimer and monomer, based on the phenomena that catalytic domain of SrtA physically forms a dimer *in vitro*, we proposed that the equilibrium between SrtA dimer

and monomer might have certain function *in vivo*, which can not be revealed by *in vitro* assay. This function might be related to a possible regulating mechanism on cellular membrane to modify SrtA's function *in vivo* in response to certain internal or external stimuli. It is possible that SrtA dimerization can serve as potential drug target, if the regulator of which would modify SrtA's activity. The knowledge of SrtA dimerization also complements the comprehensive understanding of SrtA catalytic mechanism.

MATERIALS AND METHODS

Construction of $SrtA_{\Delta N59}$ and $SrtA_{WT}$ expression vectors

Primers designated PsrtA59 (5'-CGATCCATGGGCCAAGCTAAACCTCAAATTCC—3') and PsrtA59R (5'—CCGCTCGAGTTTGACTTCTGTAGCTACAA—3') were used to amplify a $SrtA_{\Delta N59}$ sequence (which would express only residues 60-206) from genomic DNA from *Staphylococcus aureus* subsp. *aureus* (ATCC 700699D) by the polymerase chain reaction. Full-length SrtA ($SrtA_{WT}$) was generated in a similar manner using the primers designated PsrtA (5'—CGATCCATGGGCAAAAAATGGACAAATCGA—3') and PsrtA59R. The two DNA fragments were digested with NcoI and XhoI (New England Biolabs, Beverly, MA) and cloned into the pET28b expression vector (Novagen, La Jolla, CA) to generate the constructs pET28-SrtA59 and pET28-FSrtA.

Purification of $SrtA_{\Delta N59}$ and $SrtA_{WT}$ proteins

The pET28-SrtA59 and pET28-FSrtA constructs were transformed into *Escherichia coli* strain BL21 (Novagen). The pET28-SrtA59 transformed cells were grown in 1 L of Luria broth media at 37 ºC until the OD_{600} reached 0.6. The culture was then induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG, Invitrogen, Carlsbad, CA) and grown for another 8 hours at 37˚C. The cells were harvested and lysed under native conditions using the QIAexpression kit according to the manufacturer's

protocol (Qiagen, Valencia, CA). Cells containing the pET28-FSrtA construct were grown in 6 L of Luria broth media at 37ºC until the OD_{600} reached 0.6. The culture was then induced with 1 mM IPTG and grown for another 9 hours at 37˚C. Since the full-length protein had poor solubility under native conditions, these cells were harvested and lysed under denaturing conditions (8 M urea, 500 mM sodium chloride, 10 mM imidazole, 20 mM 2-mercaptoethanol, 10% glycerol, and 1% Triton X-100, pH 9.5). The lysate was centrifuged and the supernatant was applied to 0.8 mL Ni-NTA agarose beads (pre-equilibrated with 10 mM imidazole). The protein was eluted with 8 M urea (with 100 mM NaH_2PO_4, 10 mM Tris·Cl, pH 4.5). The collected fraction was dialyzed at 4˚C against a linear gradient (6 M - 1 M urea containing 500 mM NaCl, 20 % glycerol, 20 mM Tris·Cl, pH 7.4) for 24 hours to refold the full-length protein. Both the $SrtA_{\Delta N59}$ and $SrtA_{WT}$ proteins were then dialyzed against saline buffer (150 mM sodium chloride, 50 mM Tris·Cl, 5 mM calcium chloride, pH 7.5) for future use. The concentration of protein was determined using Bradford reagent (Pierce Biotechnology, Rockford, IL).

Polyacrylamide gel electrophoresis and western blot analysis

Proteins were analyzed on either a 12% SDS polyacrylamide gel under denaturing conditions, or a 12% native polyacrylamide gel under native conditions, and visualized by Coomassie Blue staining (*129*). For western blot analyses, proteins were transferred to a nitrocellulose membrane (Hybond-C Extra, Amersham Bioscience, Piscataway, NJ) in a Tris-glycine buffered electrophoresis tank (*129*). The membranes were then probed with an anti-His primary antibody (Invitrogen) and an alkaline phosphatase-conjugated goat anti-mouse secondary antibody (Bio-Rad, Hercules, CA). The Phospha GLO AP substrate (KPL Incorporation, Baltimore, MD) was applied to visualize the signals, which were detected by exposing the membrane to BioMax Light Film (Eastman Kodak Company, Rochester, NJ).

In vitro chemical cross-linking

Both the purified $SrtA_{WT}$ and $SrtA_{\Delta N59}$ proteins were dialyzed against 100 mM NH_4HCO_3 buffer (pH 7.6), and lyophilized. Egg albumin and lysozyme were purchased from Sigma-Aldrich (St. Louis, MO) as lyophilized powders. For cross-linking with 1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (EDC, Sigma-Aldrich), the proteins were dissolved in 0.1 M 2-(4-morpholino)-ethane sulfonic acid (Fisher Scientific, Fair Lawn, NJ), pH 4.5, to give a 2.5 mg/mL solution. This was added to a freshly prepared EDC aqueous solution (0.8 mg/mL). The final concentration of the proteins was 1.25 mg/mL. After one hour at room temperature, the reaction was quenched with 2% (v/v) 2-mercaptoethanol. For cross-linking with N-hydroxysulfosuccinimidyl-4-azidobenzoate (Sulfo-HSAB, Pierce Biotechnology), the lyophilized proteins were dissolved in 1X PBS (20 mM NaH_2PO_4, 150 mM NaCl, pH 7.4) to a concentration of 4 mg/ml. Freshly prepared Sulfo-HSAB solution (2 mg/mL in DMSO/H_2O, 1:2 v/v) was then added to the protein solution prepared above to give a final concentration of 3 mg/mL for protein, 0.5 mg/mL for EDC. After incubation at room temperature (protected from light by foil) for 1 hour, the mixture was irradiated by exposure to UV light for 5 minutes. The cross-linked products were analyzed by SDS-PAGE, and visualized by Coomassie Blue staining and western blot as described above.

Chemical cross-linking in *E. coli*

BL21 cells transformed with the pET28-FSrtA construct and cells transformed with the pET28-SrtA59 construct were grown and induced as described above. EDC was then added to the media to a final concentration of 0.7 mg/mL. The cells were harvested 40 minutes later. $SrtA_{WT}$ was purified under denaturing conditions and analyzed by western blot as described above. $SrtA_{\Delta N59}$ was purified under native conditions and analyzed by Coomassie Blue staining as described above.

Protein in-gel digestion and peptide fingerprint mapping by MS

The proposed dimer and monomer bands of $SrtA_{\Delta N59}$ were sliced from an SDS-polyacrylamide gel. The gel pieces were subjected to in-gel trypsin digestion. Protein was then extracted from the gel with formic acid/water/2-propanol (1:3:2 v/v/v) and analyzed by MALDI-TOF mass spectrometry and peptide fingerprint mapping analyses in the Analytical Instrumentation Core Facility at the University of Texas at Austin.

Protein sequencing

The $SrtA_{\Delta N59}$ protein was analyzed by SDS-PAGE and transferred to a PVDF membrane (Immobilon-P, Millipore) using a Bio-Rad Semi-Dry electrophoretic transfer cell. The proposed protein dimer on the membrane was sequenced using the Applied Biosystems 477A pulse liquid phase sequencer (Protein Microanalysis Facility, The University of Texas at Austin) (*130*).

Size-exclusion chromatography

Purified $SrtA_{\Delta N59}$ protein was applied to a Superdex 200 column (300 × 10 mm) pre-equilibrated with 50 mM NH_4HCO_3 buffer. The sample was eluted with the same buffer at 0.5 ml/min at 4°C using an AKTA FPLC (Amersham Pharmacia Biotech). Eluted protein was detected by monitoring at an absorbance of 280 nm. Proteins used as molecular weight standards were cytochrome C (12.4 kDa), carbonic anhydrase (29 kDa), bovine serum albumin (66 kDa), alcohol dehydrogenase (150 kDa), and β-amylase (200 kDa), purchased from Sigma.

Analytical sedimentation equilibrium ultracentrifugation

$SrtA_{\Delta N59}$ protein purified under native conditions was used for these analyses. All experiments were carried out on a Beckman Optima XL-1 analytical ultracentrifuge with the rotor speed set to 20,000 rpm, 30,000 rpm, and 40,000 rpm at 25.0 °C. All

experiments were performed in triplicate using 6-channel centerpieces. Ultracentrifuge cells were assembled containing 26 μM or 52 μM $SrtA_{\Delta N59}$ in saline buffer (150 mM sodium chloride, 50 mM Tris·Cl, 5 mM calcium chloride, pH 7.5). Scans were performed by a monochromator inside the ultracentrifuge with a wavelength of 242 nm, or 246 nm. Ultrascan II version 8.0 software (*131*) was used to estimate the approximate time it would take for the samples to achieve equilibrium at different concentrations and rotor speeds. Thirteen scans were then performed for each cell at the time points 0:30, 18:00, 21:00, 24:00, 27:00, 42:00, 45:00, 48:00, 51:00, 66:00, 69:00, 72:00, and 75:00 (hour:minute). Scanning data was extracted and fitted using Ultrascan II 8.0. A partial specific volume ($\bar{\nu}$) of 0.7349 ml/g (predicted by Sednterp 1.07 software, www.bbri.org/RASMB/rasmb.html) and a buffer density (ϱ) of 1.0 g/ml were used in the analysis. The data were globally and simultaneously fitted to a one component ideal species model and a monomer-dimer equilibrium model using the equations below, where X = radius, Xr = reference radius, A = amplitude of monomer*, M = molecular weight of monomer*, E = extinction coefficient, R = gas constant, T = temperature, B = baseline*, ω = angular velocity, L = path-length, $K_{1,2}$ = monomer–dimer equilibrium constant* (* indicates this parameter can be floated) (*131*).

One component ideal species model (Equation 2.1):

$$C(\mathrm{X}) = e^{\left[\frac{\ln(A)+M\omega^2\left(1-\bar{\nu}\rho\right)\left(X^2-X_r^2\right)}{2RT}\right]} + B$$

Monomer-dimer equilibrium model (Equation 2.2):

$$C(\mathrm{X}) = e^{\left[\frac{2\ln(A)+\ln\left(\frac{2}{EL}\right)+\ln(K_{1,2})+2M\omega^2\left(1-\bar{\nu}\rho\right)\left(X^2-X_r^2\right)}{2RT}\right]} + e^{\left[\frac{\ln(A)+M\omega^2\left(1-\bar{\nu}\rho\right)\left(X^2-X_r^2\right)}{2RT}\right]} + B$$

Determination of kinetic constants

EDC mediated *in vitro* cross-linked $SrtA_{\Delta N59}$ was purified by size-exclusion chromatography with saline buffer (50 mM Tris·Cl, 150 mM NaCl, 5 mM $CaCl_2$, 5 mM 2-mercaptoethanol, pH 7.5). The monomeric and dimeric fractions of $SrtA_{\Delta N59}$ were separated, collected and used for the following enzymatic assays. The peptide substrate *o*-aminobenzoyl-LPETG-(2,4-dinitropheyl)-diaminopropionic acid (Abz-LPETG-Dap(Dnp)) was synthesized based on the well-established Fmoc/piperidine chemistry on a PAL resin (Protein Microanalysis Facility, The University of Texas at Austin). The peptide GlyGlyGlyGlyGly was purchased from BACHEM (Torrance, CA). Assays were performed in a 100 μL assay buffer (50 mM Tris·Cl, 150 mM NaCl, 5 mM $CaCl_2$, 5 mM 2-mercaptoethanol, pH 7.5) containing $SrtA_{\Delta N59}$ (100 nM for monomer, 50 nM for dimer), pentaglycine (2 mM) and varying concentrations of Abz-LPETG-Dap(Dnp) (0 to 300 μM). Reactions were initiated by the addition of enzyme and were monitored by measuring the increment in fluorescence for 10 min (λ_{ex}= 317 nm, λ_{em}= 420 nm) at 37 °C on a Spex FluoroMax-3 spectrofluorometer (Jobin Yuon Co., Edison, NJ). Initial velocities (V_0) were calculated as units of fluorescence per unit time using Equation 2.3, where m is the slope during the linear phase of the cleavage, [S] is substrate concentration, and I_0 and I_{100} are the fluorescence intensities of substrate solution before and after complete cleavage, respectively. The slope (m) was measured in three independent experiments. The velocities (v) were determined from the progress curves (the steady-state rates in the case of biphasic curves) at various substrate concentrations. The obtained velocities were then fit to Equation 4 using software KaleidaGraph (version 3.6) to afford the apparent $K_{m,\,app}$ and $k_{cat,\,app}$ values (*64*).

$$V_0 = \frac{m \times [S]}{I_{100} - I_0}$$ (Equation 2.3)

$$v = \frac{k_{cat}[\mathrm{E}][\mathrm{S}]}{[\mathrm{S}] + K_m}$$ (Equation 2.4)

Chapter 3 Use of site-directed mutagenesis to disrupt sortase A dimerization *in vitro*

INTRODUCTION

In vitro SrtA dimerization

We have shown in Chapter 2 that $SrtA_{\Delta N59}$ exists as a dimer with a K_d of 55 μM and both the monomer and dimer fractions of the protein were active *in vitro* (*100*). However the molecular mechanism of SrtA dimerization and the exact function of the dimer-monomer equilibrium *in vivo* are not clear yet. In this chapter, we describe the investigation to understand what physical forces drive the formation of homo-dimer at the molecular level by alanine scan mutagenesis. Three single point mutated SrtA mutants are generated which are in monomer only conformation in an *in vitro* solution. Based on the understanding of the SrtA dimerization interface, we also developed a site-specific cross-linking reagent, which will allow the rationale design of peptidomimics, which may regulate the extent of dimerization.

Structural basis for protein-protein interactions

The binding of two proteins is determined by various biophysical and chemical properties of interaction interface: First, we will consider size and shape. Most protein-protein interaction interfaces are rather large, on the order of 700-1500 A^2 per protein, and the surfaces can either be relatively flat or twisted (*132, 133*). Second, we will consider complementarity. This property includes both electrostatic and shape complementarity between interfaces of the two associating molecules (*134-138*). Third, we will consider the composition of amino acids on the interface. The chemical properties of these molecules derived from polypeptide backbone or side chains provide

the interactions between two surfaces and stabilize the protein complex. The driving forces are hydrophobic interactions, electrostatic interactions, hydrogen bonds, and the van der Waals contacts (*139*). It has been proposed that hydrophobic interactions drive two proteins to associate together (*140*) while hydrogen bonds and electrostatic interactions determine the specificity (*141*). Energy for regular PPIs, ΔG_d, ranges from 6 to 19 kcal/mol, while a single pairwise interaction between amino acids may yield 4-8 kcal/mol when a salt bridge is formed or 4-5 kcal/mol when a hydrogen bond is formed. Non-polar van der Waals interactions are very weak and no greater than 1 kcal/mol range (*142*). Fourth, we will consider segmentation and modules. Some well-studied very stable protein-protein interfaces are composed of several secondary-structure defined modules that independently provide major association forces (*143*). The cooperative effects between modules, also termed "hot regions", additively contribute to the high affinity for binding (*144, 145*). Fifth, we will consider structural rearrangements on binding. It is not clear to what extent the proteins change their conformation to accommodate binding, since there is a few protein complexes that have molecular structures available both before binding and after, but there must be an energetic price to pay for the movements when they bind (*146-151*).

Alanine-scanning mutagenesis has been developed since 1989 as a simple and widely used technique to identify the specific side chains on proteins that are essential for particular catalytic or functional roles. To map the dimerization interface, single alanine mutations are introduced at every residue contained within the domains that have been implicated in dimer subunits recognition. The substitution with alanine residue eliminates the side chain beside the β carbon and yet retains the main-chain conformation without imposing extreme electrostatic or steric effects on the spot. The substitution of an individual amino acid residue with alanine might disrupt dimerization which indicates a

critical interaction mediated by this site in an interface (*152*). These energetically important residues are called "hot spots", that account for the majority of the binding energy in protein-protein interaction (*153*). It has been suggested that the presence of a small fraction of hot spots on dimerization interface is general in most protein-protein complexes (*154*). To discover both the location and hot spots of dimerization domain of SrtA, we performed an alanine scan mutagenesis on rationally chosen sites suspected to occur at the dimerization interface using both hydrophilic and hydrophobic amino acid substitutes. A successful reduction in SrtA homo-dimerization was realized when any of three key hydrophilic residues: N132, K137 and Y143, was mutated to alanyl residues.

Site-specific protein cross-linking with genetically incorporated 3,4-dihydroxy-L-phenylalanine

To improve the current cross-linking method to study weak and transient protein-protein interactions, we developed a novel site-specific cross-linking reagent, 3,4-dihydroxy-L-phenylalanine (L-DOPA) genetically incorporated into the SrtA primary sequence, to investigate its self-association. The successful selective cross-linking mediated by L-DOPA between two SrtA monomer molecules at the K137 site in turn supported the essentiality of this residue as a hot-spot amino acid. Figure 3.1 has illustrated the principle and scheme for this novel tool.

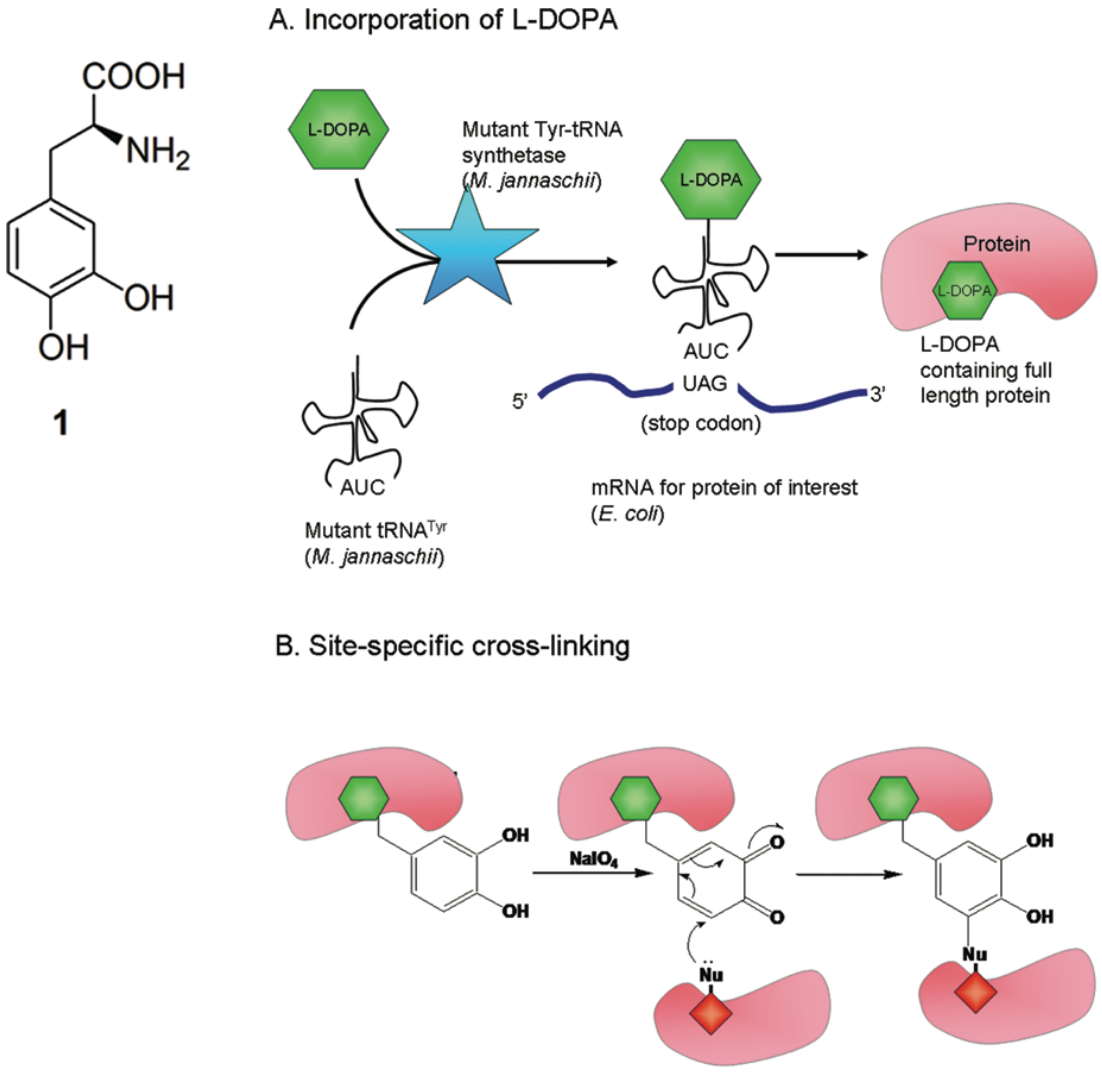

Figure 3.1: Scheme for L-DOPA incorporation.

(A) Established methodology for unnatural amino acid incorporation. (B) L-DOPA mediated cross-linking between two associating molecules.

Previous studies have clearly demonstrated the high specificity of L-DOPA as an ideal cross-linking agent (*155-157*). The insertion of L-DOPA into a peptide ligand can selectively help the recognition and association of this ligand with its binding protein in the context of protein mixture by forming a covalent bond that is oxidized with sodium periodate (*155*). The initial step in the reaction is the generation of an electrophilic ortho-

quinone intermediate, which is attacked by nucleophilic reactive groups in close proximity, such as the terminal amine groups of the polypeptide backbone or nucleophilic side chains such as Asp, Glu, Ser, and Cys (*156*).

L-DOPA is an unnatural amino acid which can be site-specifically incorporated into proteins with an established general methodology (*158*). Briefly, the engineered *Methanococcus jannaschii* $tRNA^{Tyr}$/tyrosyl-tRNA synthetase orthogonal pair can successfully charge L-DOPA on *M. jannaschii* $tRNA^{Tyr}$ and incorporate this L-DOPA into a protein in *Escherichia coli* at the site with amber stop codon (TAG) (*159, 160*). Thus, if the incorporated L-DOPA is located on the direct binding interface of two interacting proteins, it can specifically cross-link two protomer molecules together by forming covalent bond with nucleophiles within close proximity. The formation of L-DOPA mediated covalent dimer can provide us with a novel assay to detect and study protein-protein interactions and map their binding surface.

RESULTS AND DISCUSSION

Generating mutant proteins using site-directed mutagenesis

The crystal structure of $SrtA_{\Delta N59}$ (1T2P) from X-ray diffraction was obtained from the Protein Data Bank (*51*) and analyzed by Glaxo Smith Kline's DeepView / Swiss-Pdb Viewer for various characteristics. The template crystal has three sortase molecules originated from three distinct sides of the crystal in each unit cell. The amino acid residues on protein surface were first distinguished into different groups based on their hydrophilic/hydrophobic characteristics. Then the packing interfaces between different protein subunits among the three chains were analyzed for their proximity. Based on these criteria, all suitable hydrophilic/ionic residues in close proximity to other residues on different protein chains were listed. The five top residues are K62, N132, K137, Y143

and K152. We have generated the five SrtA mutants, with each of these residues mutated to alanine individually. In the same manner, hydrophobic surface residues on different protein chains were also screened based on their close interactions. We did not find any close interacting hydrophobic residues. However, three hydrophobic amino acids, F122, I123 and P126 on each protein subunit, shares a close proximity to each other and forms an extending hydrophobic pocket, and these were chosen as candidates for site directed mutagenesis to disrupt the associations conferred by hydrophobic interactions. In our experiment, the F122G mutation was difficult to obtain, only the I123G and P126G mutant proteins were generated.

The seven mutant SrtA proteins: K62A, I123G, P126G, N132A, K137A, Y143A and K152A, were successfully generated by introducing the point mutation on truncated SrtA $SrtA_{\Delta N59}$ wild type protein individually. Regulated by a T7 promoter in *E. coli,* the expression level of mutant K62A, I123G, P126G and K152A were similar to wild type protein. However, N132A, K137A and Y143A had relatively lower production, yielding only 1/3 to 2/3 of the protein compared to wild type SrtA.

Mutant proteins have different dimerization pattern under non-denaturing FPLC

Individual protein was purified and analyzed on size-exclusion chromatography under native condition to determine the size, molecular mass and oligomerization state of $SrtA_{\Delta N59}$ mutants (*100*). As illustrated in Figure 3.2, elution patterns each mutated SrtA protein were aligned with the pattern of wild type $SrtA_{\Delta N59}$. Wild type $SrtA_{\Delta N59}$ was loaded on a calibrated Superdex 200 10/300 GL column at the concentration of 5 mg/mL (281 μM) and eluted as three peaks. Based on the elution volumes of three peaks, the molecular mass corresponding to each peak was calculated. The major peak came out at 15.2 mL, corresponding to an expected homo-dimer species of 36,507 Da. The second largest peak came out later at 16.6 mL, corresponding to the expected monomer species

of 16,766 Da. The minor peak came out earlier at 14.3 mL was deconvoluted to be a homo-trimer or tetramer of $SrtA_{\Delta N59}$ with a molecular mass 59,498 Da. The integrated area under each peak was calculated and listed in Table 3.1, to indicate the amount of dimer and monomer. A trimer(tetramer):dimer:monomer absorbance ratio of 1:7:1 was observed for wild type protein. Mutant K62A has a similar elution pattern and same dimer:monomer 7:1 ratio as wild type protein (Figure 3.2 B blue curve). The data suggested that K62A mutation does not affect the dimerization of $SrtA_{\Delta N59}$.

Mutant I123G (Figure 3.2 C purple curve) and P126G (Figure 3.2 D red curve) have significant different elution patterns compared to wild type. Both I123G and I126G had two peaks eluted at the same retention times for wild type dimer and monomer peaks. However the size dimer peak significantly decreased. A dimer to monomer peak ratio decreased from a 7:1 ratio to a 2.6:1 ratio for I123G sample. The trimer peak was barely recognized from the chromatography data. The elution pattern for mutant P126G gave the integrated dimer to monomer ratio of 1.9:1. No separate peak for a higher molecular weight protein complex can be detected. Based on the observation, these two mutants at position I123 and P126 individually reduced dimer association of $SrtA_{\Delta N59}$ significantly.

Based on the elution patterns, almost complete disruption of dimer formation was realized by the other three mutants, N132A (Figure 3.2 E pink curve), K137A (Figure 3.2 F light blue curve) and Y143A (Figure 3.2 G green curve). The dimer to monomer ratio of about 1:16 was observed for N132A and K137A, each of which was eluted as a large monomer peak and a very small dimer peak. Mutant Y143A was slightly different. In addition to the large monomer peak and a barely detectable dimer peak, some distinctive higher molecular weight peaks were also observed, which corresponded to molecular mass of 190 kDa, 85 kDa and 47 kDa, with a ratio of 10:1:1.5. To identify the protein composition of those peaks, the eluted fractions were collected and subjected to mass

spectroscopic analysis. Protein ID data suggested that these peaks contain $SrtA_{\Delta N59}$ oligomers (data not shown).

However, the opposite oligomerization result was discovered in another hydrophilic alanine mutation, mutant K152A (Figure 3.2 H orange curve). K152A had a dimer to monomer ratio of 12:1, slightly higher than the 7:1 ratio for wild type protein. In addition, the trimer peak is also larger than the one for wild type. The trimer:dimer:monomer ratio in this mutant was 3:12:1, while a 1:7:1 ratio was observed for wild type protein under similar conditions. Therefore, mutation from lysine to alanine at this site enhanced the dimerization of $SrtA_{\Delta N59}$.

Based on the oligomerization states revealed by gel filtration analysis, we identified three mutant proteins that successfully disrupted dimerization, N132A, K137A and Y143A; two mutants that significantly reduced dimerization, I123G and P126G; and, interestingly, one mutant that substantially enhanced dimerization, K152A.

Table 3.1. Analytical size exclusion FPLC of wild type and mutant $SrtA_{\Delta 59}$ proteins.

	Elution volume (mL)		Integrated area (mL*mAu)		Dimer/monomer ratio
	Dimer	Monomer	Dimer	Monomer	
wildtype	15.16	16.53	71.9	10.8	6.8 : 1
K62A	15.19	16.56	75.5	9.6	7.9 : 1
I123G	15.19	16.52	71.2	27.1	2.6 : 1
P126G	15.01	16.51	56.6	30.3	1.9 : 1
N132A		16.52		55.5	
K137A		16.4	1.1	18.3	1 : 16.0
Y143A		16.47		35.5	
K152A	15.19	16.55	52.1	4.3	12.1 : 1
Estimated MW	36507	16766			

Purified wild type and $SrtA_{\Delta N59}$ mutant proteins were applied to a size exclusion gel filtration column. The samples were eluted with washing buffer (0.1M sodium phosphate, 0.15M sodium chloride, pH 6.8) at 0.75 mL/min at 4°C and the elution pattern was detected by monitoring at an absorbance of 280 nm. The integrated area under each peak was calculated in unit of mL*mAu. The relative ratio of peak area was used to estimate the ratio between dimer and monomer species.

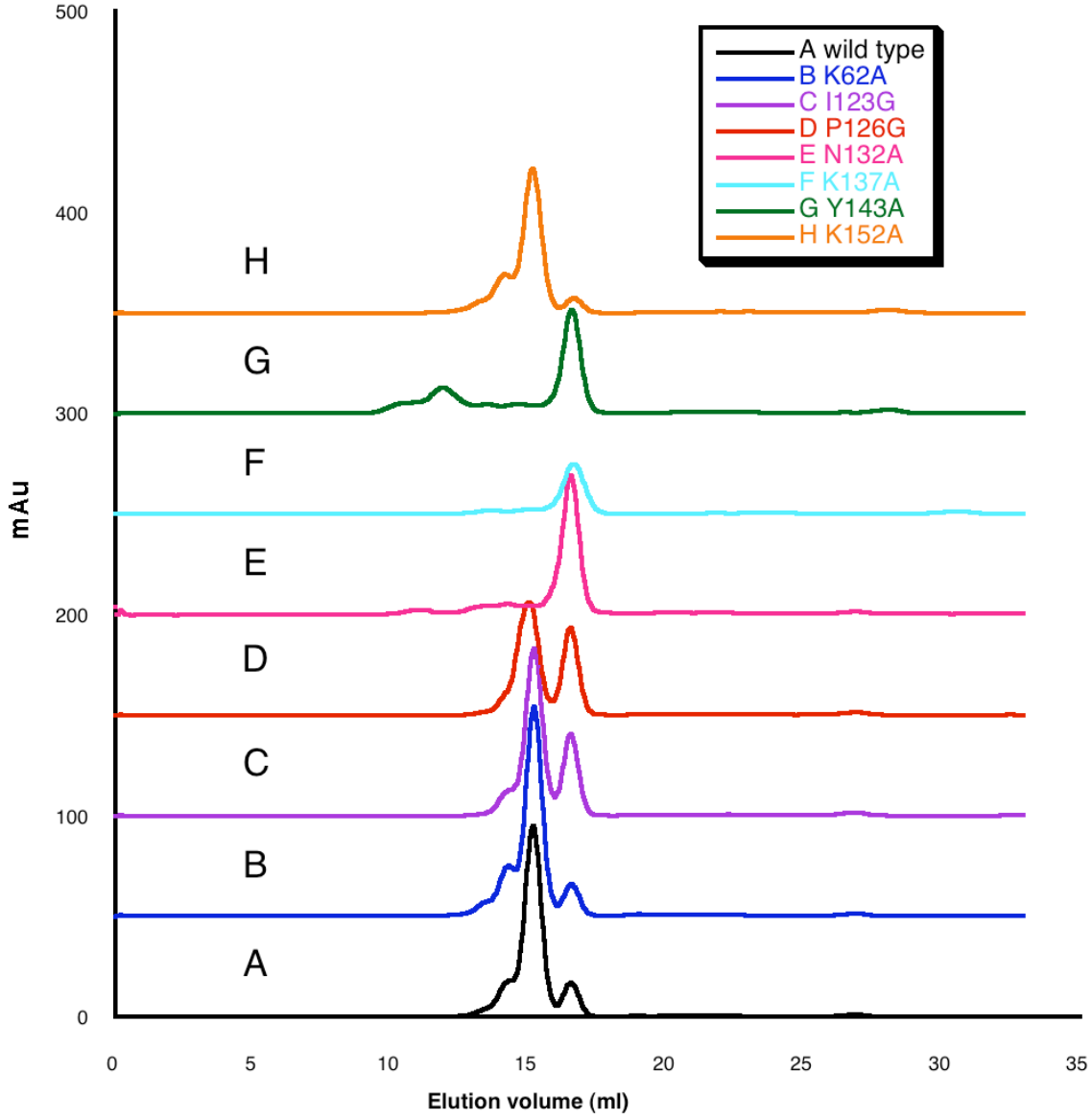

Figure 3.2: Native FPLC with gel filtration chromatography of $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59}$ mutant proteins.

Purified wild type and $SrtA_{\Delta N59}$ mutant proteins were applied to a size exclusion gel filtration column. The samples were eluted with washing buffer (0.1M sodium phosphate, 0.15M sodium chloride, pH 6.8) at 0.75 mL/min at 4°C and the elution pattern was detected by monitoring at an absorbance of 280 nm. Data are over laid with different colors. A Black, wild type protein; B blue, K62→A62; C purple, I123→G123; D red, P126→G126; E pink, N132→A132; F light blue, K137→A137; G green, Y143→A143; H orange, K152→A152.

Mutant proteins on native PAGE show different dimerization and migration pattern

The generated $SrtA_{\Delta N59}$ mutant proteins with different dimerization profiles were subjected to native PAGE analysis (Figure 3.3). Different from denaturing SDS-PAGE, which separates proteins mainly based on their distinct molecular weights, the mobility of a protein in native PAGE is determined by both the overall charges on the protein and its hydrodynamic size (*161*). Purified $SrtA_{\Delta N59}$ proteins, including wild type and mutants, were loaded separately, at a concentration of 169 μM in each lane. The concentration of loaded proteins is higher than the previously calculated K_d (56 μM), which allows the dimer formation in wild type sample. From the data of size exclusion chromatography under a non-denaturing condition, all of the $SrtA_{\Delta N59}$ mutants as well as wild type protein had the same molecular weight for the monomer. Therefore, under the native PAGE condition, only the overall charges associated with the protein's isoelectric point or the hydrodynamic size which correlates to the conformation of the protein will cause mobility difference among protein samples. Demonstrated in Figure 3.3, mutant Y143A, I123G and P126G migrated at a similar velocity as wild type protein. However, mutant K62A, K137A and K152A had a higher relative migration velocity during electrophoresis and moved faster on gel toward the bottom. This pattern can be explained by the charges of those proteins in an experiment buffer with a pH of 7.5. Based on protein calculator (*162*), at pH 7.5, wild type $SrtA_{\Delta N59}$ had a net charge of -0.3, as did mutant I123G, P126G, N132A and Y143A. A higher negative charge of -1.3 on $SrtA_{\Delta N59}$ mutants K62A, K137A and K152A was introduced by the mutation from positively charged lysine to alanine. Since all $SrtA_{\Delta N59}$ proteins had the similar molecular weights for each molecule, if their hydrodynamic sizes are the same as well, the proteins bearing higher negative charges should migrate faster towards the cathode during electrophoresis. In the same

manner, assuming proteins all bears the same net charges, the difference observed in migration pattern can be attributed to proteins' hydrodynamic sizes. The protein with well-folded chains is more compact and smaller than the one with only partially folded chains. The relatively smaller size makes the protein moving faster in the matrix during gel electrophoresis. Although, the overall folding and conformation of these mutant proteins might not be significantly different from the wild type, some changes were observed. For instance, K152A's monomer band moves a slightly slower than the corresponding monomer bands of K62A and K137A, both of which bear the same negative charges of -1.3.

The dimer to monomer ratios of SrtA mutants and wild type protein were also evaluated by quantifying the densities of dimer and monomer bands using ImageQuant. The D/M (dimer/monomer) ratios were calculated and listed in Table 3.2. The estimated dimmer/monomer ratios from native PAGE analysis are in good agreement with the measurements from analytical gel filtration experiments. The discrepancy can be introduced by different buffer systems used in the two methods, interaction of proteins with different matrix systems, temperatures under which experiments are performed, and other factors.

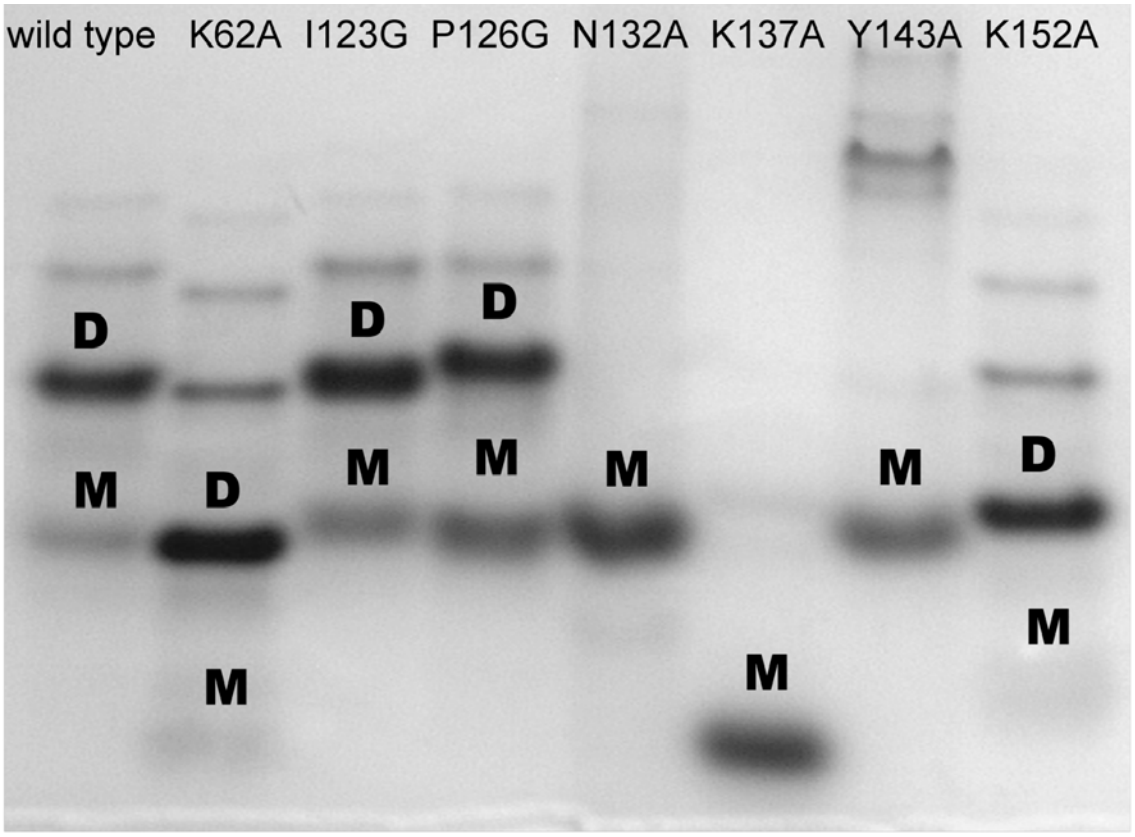

Figure 3.3: Native PAGE of $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59}$ mutant proteins.

Purified $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59}$ mutant proteins were first resolved by PAGE in native condition depending on their molecular mass and hydrodynamic size, and then visualized by Coomassie Blue staining. In each lane, 4 μl of 5 mg/mL protein was loaded except for K137A which has a lower concentration of 2.3 mg/mL. The theoretical mass of a monomer of $SrtA_{\Delta N59}$ is 17.9 kDa, while that of a dimer is 35.8 kDa. The first lane contains $SrtA_{\Delta N59}$ wild type. The lane of K62A contains the mutant $SrtA_{\Delta N59}$ with mutation K62→A62. The lane of I123G contains the mutant $SrtA_{\Delta N59}$ with mutation I123→G123. The lane of P126G contains the mutant $SrtA_{\Delta N59}$ with mutation P126→G126. The lane of N132A contains the mutant $SrtA_{\Delta N59}$ with mutation N132→A132. The lane of K137A contains the mutant $SrtA_{\Delta N59}$ with mutation K137→A137. The lane of Y143A contains the mutant $SrtA_{\Delta N59}$ with mutation Y143→A143. The lane of K152A contains the mutant $SrtA_{\Delta N59}$ with mutation K152→A152. In each lane, the band labeled on top with D indicates the band of dimeric protein, the band labeled on top with M indicates the band of monomeric protein.

Table 3.2: Comparison of the estimation of the dimer/monomer ratios for $SrtA_{\Delta N59}$ mutants between native PAGE and FPLC.

	WT	K62A	I123G	P126G	N132A	K137A	Y143A	K152A
D/M from PAGE	3.6:1	5.52:1	2.82:1	2.07:1	0	9.54:1	0	11.6:1
D/M from FPLC	6.77:1	7.86:1	2.62:1	1.87:1	0	16:1	0	12.06:1

Purified wild type and $SrtA_{\Delta N59}$ mutant proteins were loaded on native PAGE. The intensities of the protein bands were analyzed using ImageQuant software. The relative ratio of band intensities was used to estimate the ratio between dimer and monomer species (D/M). FPLC data were achieved as described in Table 1.

Circular Dichroism spectra of mutant proteins prove proper overall folding

Circular dichroism experiments were performed to analyze the folding behavior and secondary structure differences among $SrtA_{\Delta N59}$ wild type and mutant proteins. Far UV CD spectrum was recorded for each protein sample in buffer PBS and aligned in Figure 3.4. The raw spectra recorded in millidegrees of ellipticity (θ) were first converted to mean residue ellipticity [θ] in deg.cm^2/dmol based on the characteristic of each protein. Mean residue ellipticity is directly correlated with changes of chirality which indicates the changes in protein's secondary structure. The spectrum of wild type SrtA is characterized by a negative band around 210 nm; $(\theta)_{212.3nm}$ =-9453.86 deg.cm^2/dmol.residue, and a positive band near 230 nm; $(\theta)_{232.2nm}$ =651.90 deg.cm^2/dmol.residue. The strong negative band near 210 nm represents the combination of disordered polypeptides and α-helix. The small positive band around 230 nm is the characteristic feature of β-sheet in proteins (*163*). We then used the online program, CDSSTR secondary structure method (*164*) to estimate the contents of secondary structure of folded proteins. Reference data set 4 (*165*) were used. The results were summarized and listed in Table 3.3. NRMSD fit parameter was used to evaluate if the calculated secondary structures are in agreement with the spectra. Very low values of <<0.1 were obtained for our calculation indicating good consistency for our results.

Based on the analysis of CD data, wild type $SrtA_{\Delta N59}$ contains a combination of mainly β-strand, 45%, β-turn, 21%, disordered structures, 31%, and α-helix, 3% (Table 3.3). The crystal structure(*51*) and NMR data(*50*) of $SrtA_{\Delta N59}$ have indicated that among the 145 amino acids of $SrtA_{\Delta N59}$, 64 are involved in β-sheet (44.1%), whereas only 9 are organized in α-helix (6.2%), and the remaining of 72 are connected in turns and random loops (49.7%). Our secondary structural data derived from CD spectrum of wild type $SrtA_{\Delta N59}$ is consistent with the NMR and X-ray crystal data that approves the validity of experimental setting up. Spectra for mutant proteins were also calculated to produce very similar secondary structures. As demonstrated in Figure 3.4, when the spectra of wild type $SrtA_{\Delta N59}$ and mutants were overlaid, there were no significant changes in β-sheet features observed, suggesting no substantial difference in the overall folding of either protein. Although the equivalent peak at 230 nm and negative band near 210 nm were presented in all spectra, the magnitudes were shifted in some cases, which provided the evidence for alterations in secondary structure. The minor difference in secondary structure composition introduced by a single amino acid mutation on protein surface might cause changes in protein's quaternary structure (protein-protein interactions). However, there was no significant change of the predicted secondary structure content among protein samples. Therefore, it is safe to conclude that all of the mutant proteins are well folded in the sense of producing recognizable $SrtA_{\Delta N59}$ proteins although the mutation of certain amino acid on the predicted protein-protein interaction surface area to alanine or glycine was introduced.

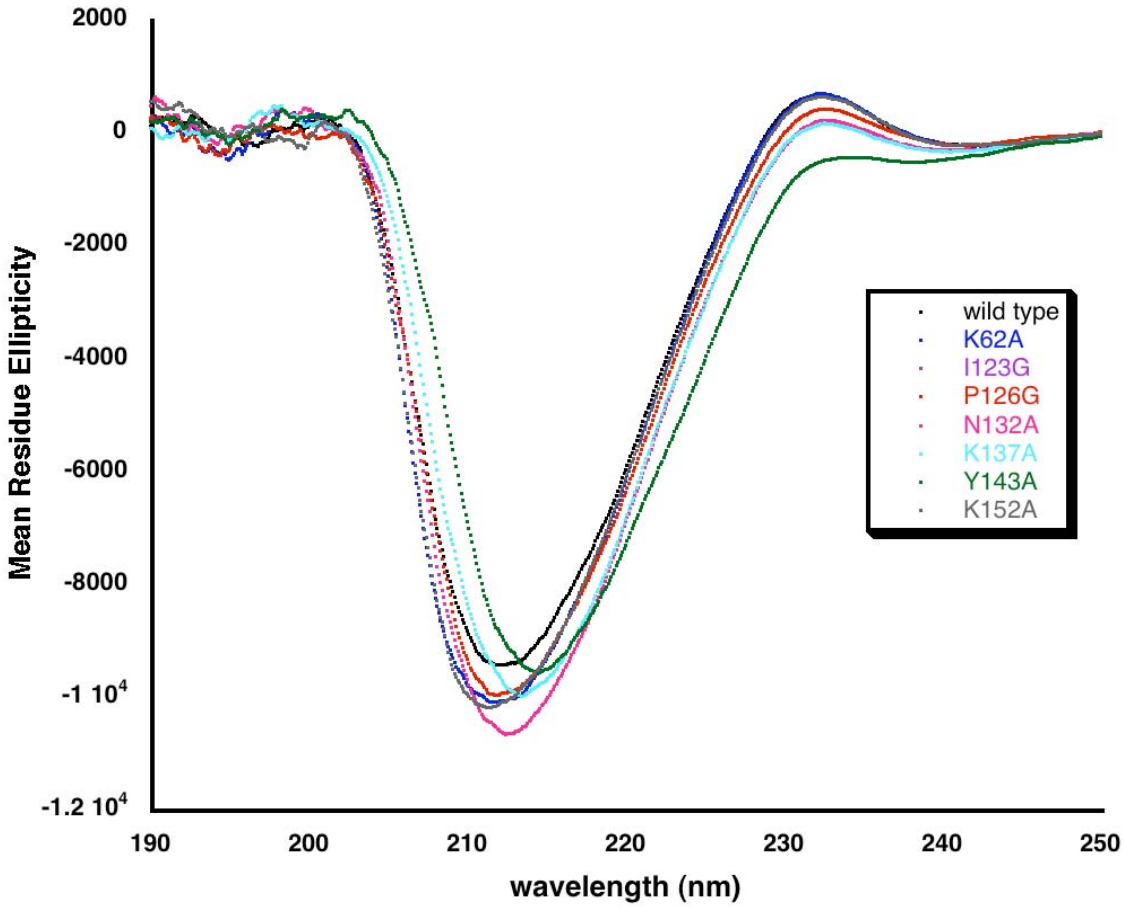

Figure 3.4: CD spectrum of $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59}$ mutant proteins.

Purified $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59}$ mutant proteins were scanned on circular dichroism spectroscopy. For each sample, 38 μM protein in activity assay buffer was scanned from 190 nm to 250 nm. Data are over laid with different color. Black, wild type protein; blue, K62→A62; purple, I123→G123; red, P126→G126; pink, K132→A132; light blue, Y137→A137; green, K143→A143; grey, K152→A152.

Table 3.3. Estimated secondary structure conformation of mutant and wild type of $SrtA_{\Delta 59}$ proteins.

Fraction Ratio	Wildtype	K62A	I123G	P126G	N132A	K137A	Y143A	K152A
NRMSD	0.024	0.022	0.022	0.022	0.018	0.019	0.019	0.025
Helix:	3%	5%	4%	4%	5%	4%	4%	4%
Beta:	45%	45%	50%	50%	50%	47%	46%	47%
Turn:	21%	22%	20%	20%	20%	20%	22%	21%
Random:	31%	27%	24%	25%	26%	27%	28%	27%
Total:	100%	99%	98%	99%	101%	98%	100%	99%

Purified $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59}$ mutant proteins were scanned on circular dichroism spectroscopy. Estimated secondary structure conformation was calculated by Dichroweb server.

In vitro activity assay indicated all mutants are active

To testify if the mutant proteins were still active, an *in vitro* activity assay was set up according to the protocols described in chapter 2 (*85*). Upon the addition of active enzyme, the specific LPETG sequence in Abz-LPETG-Dap(Dnp)-NH_2 substrate would be recognized, digested between T and G and the LPET part transfered to Gly_5. We used a HPLC program to separate the mixture and monitor the signal from released AbZ upon cutting at 355 nm by a UV detector. The recorded spectrum with the relative integration areas of substrates and products were used to calculate the conversion rates. A volume of 100 μL reaction system was set up by incubating two substrates Abz-LPETG-Dap(Dnp)-NH_2 and pentaglycine with 8.4 μM of enzyme ($SrtA_{\Delta N59}$ or respective mutant proteins) at 37˚C for 4 hours before injected into a fast analytical HPLC column. The apparent average rate in 4 hours was calculated. Wild type $SrtA_{\Delta N59}$ and most of the mutant proteins had the same average conversion rate except for I123G, which had a lower conversion rate that is about only 33% of the wild type protein. This data demonstrated that most mutant proteins folded into normal SrtA conformation in the sense of correct orientation of active sites except for I123G. It is possible that I123 is located in close

proximity to His^{120}, which is the key residue for SrtA catalytic reaction (*47*), so that the replacement of Ile with Gly distorts the local folding near the active sites. Although there was no difference in *in vitro* hydrolysis activity detected among various mutant proteins and wild type $SrtA_{\Delta N59}$, it did not imply that $SrtA_{\Delta N59}$ dimer has no effect on function since the enzyme concentration used in this method was much lower than the K_d and was not enough to form dimer. Due to the big controversy of *in vitro* assay set up in the field, we believed that *in vivo* assays under biologically relevant condition would provide more accuracy in detecting the true differences among dimer and monomer form of SrtA.

L-DOPA incorporated at K137 specifically crosslinks SrtA homo-dimers

Based on dimerization disruption results *in vitro*, we hypothesized that the dimerization of SrtA is driven by electrostatic interactions, conferred by the polar and charged amino acid side chains on N132, K137 and Y143. Therefore, substituting these three individual residues with L-DOPA is likely to result in the cross-linking between two protomer molecules, given that these sites are directly interacting with nucleophilic groups on its association partner. Residue K137 was selected as a candidate site for L-DOPA incorporation in positive cross-linking assay. Another candidate site, K206, which is unlikely to participate in the dimerization of $SrtA_{\Delta N59}$ was selected as negative control. It has been shown that he mutation of K206 to alanine does not change the dimerization status of SrtA (data now shown).

First, total protein *E. coli* culture expressing $SrtA_{\Delta N59}$ K137DOPA was analyzed for the efficiency of cross-linking between two SrtA molecules mediated by L-DOPA on its potential binding site. The result shows excellent yields for cross-linking reaction with a very strong band that corresponds to the molecular weight of $SrtA_{\Delta N59}$ dimer (35.8 kDa). The band was subjected to mass-spectrometric analysis for its identity. It has been confirmed that the only protein components in the band is SrtA protein (data not shown).

When L-DOPA incorporation step was omitted, the expression of $SrtA_{\Delta N59}$ was hardly detected as a weak momomer band, suggesting the minimal basal level incorporation of natural endogenous amino acids at this K137TAG site.

We then performed western blot analysis on K137DOPA and K206DOPA samples visualized by anti-His antibody to demonstrate the site-specificity of L-DOPA mediated cross-linking. As shown in Figure 3.5, three bands were observed in K137DOPA sample when L-DOPA was added for incorporation and sodium periodate for oxidation. The monomer band was intense only when L-DOPA was present in the solution, indicating the successful incorporation of L-DOPA at K137TAG site. The band corresponding to the size of $SrtA_{\Delta N59}$ dimer (35.8 kDa) was only observed in the sample with both L-DOPA and sodium periodate, demonstrating the high yield of the covalently stabilized homo-dimer of $SrtA_{\Delta N59}$ by L-DOPA mediated cross-linking. Since the production of covalent dimer is mainly catalyzed by sodium periodate, only a small amount of the dimer band was observed in the absence of sodium periodate, which might come from non-covalently formed $SrtA_{\Delta N59}$ dimer as shown in Chapter 2, Figure 2.1 and/or covalent dimer catalyzed by atmospheric oxygen (*157*). The third band observed corresponded to the molecular mass of $SrtA_{\Delta N59}$ 1.5-mer. Mass spectrometry analysis also confirmed that SrtA is the only protein component in the 1.5-mer band. This 1.5-mer band could be the covalent complex between $SrtA_{\Delta N59}$ monomer with L-DOPA and the truncated imcompleted sequence of $SrtA_{\Delta N59}$ resulting from stop codon TAG caused early termination of protein translation. K206DOPA sample, served as a negative control, produces strong band of $SrtA_{\Delta N59}$ monomer, suggesting the successful incorporation of L-DOPA on K206TAG site. However, only a small amount of dimer was observed after the addition of sodium periodate to oxidize L-DOPA. This trace amount of dimer could be the naturally occurring non-covalent dimer for $SrtA_{\Delta N59}$. This observation proved the high

site-specificity of incorporated L-DOPA as a novel cross-linking reagent, which can be used to study weak protein-protein interactions at particular site on the interface. It also supported our hypothesis about SrtA dimerization that K137 is located on the dimerization interface, which is in the close proximity to, or directly interacting with unidentified nucleophiles, such as His, Cys, in partner molecule. However, K206 is not likely located on the dimerization interface or close to any nucleophilic side chains from residues on partner molecule.

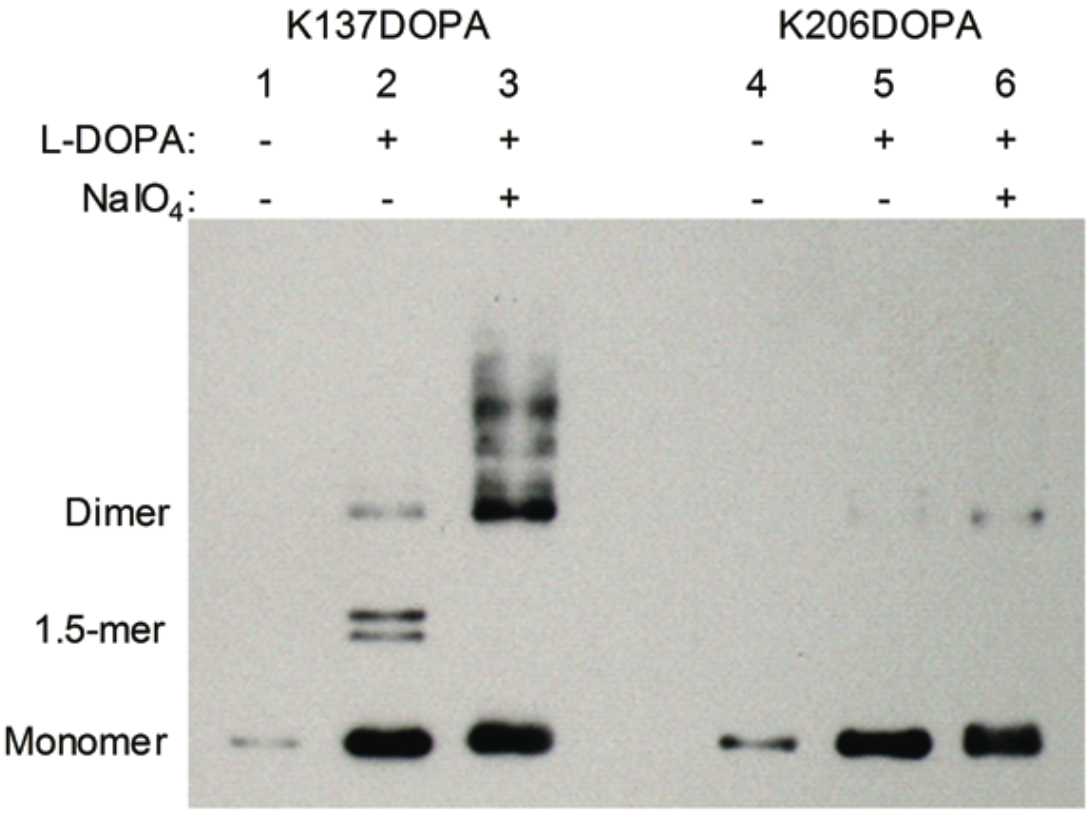

Figure 3.5: Western blot of L-DOPA containing proteins, $SrtA_{\Delta N59}$137DOPA and $SrtA_{\Delta N59}$206DOPA

His6-tagged and purified $SrtA_{\Delta N59}$137DOPA and $SrtA_{\Delta N59}$206DOPA proteins were resolved by SDS-PAGE and visualized by anti-His antibody. L-DOPA and $NaIO_4$ were added or omitted to observe their effects on dimer formation.

The possible role of amino acids involved in SrtA dimerization

By performing the alanine scan mutagenesis, we have generated several $SrtA_{\Delta N59}$ mutant proteins, each of which has a single site mutation on the sequence. Based on the dimerization status of those mutant proteins, we have tried to establish the roles of specific amino acids in forming SrtA homo-dimer. Two of the mutant proteins, K62A and K206A did not change the dimer content significantly. One mutant K152A enhanced the dimerization. Three mutants, K137A and Y143A completely disrupted the dimer formation. Based on far UV CD spectra, all mutant proteins are correctly folded into SrtA wild type conformation. Therefore, at least residues at K137 and Y143 are partially responsible for self-association of $SrtA_{\Delta N59}$. L-DOPA mediated site-specific cross-linking at 137TAG also suggested that K137 is on the dimerization interface, interacting with the residue with nucleophiles from the partner molecule to form electrostatic interaction to stabilize the dimer between two SrtA molecules. The other possible explanation for the function of K137 is based on the crystal structure. Y143 residue on chain A and K137 in chain C are 5.57 Å away with each other. The distance fits the criteria of cation-pi interaction(*166*), which has a moderate energy of about 5 kcal/mol, and is suggested to dictates recognition specificity in many protein complexes. For instance, in the cytochrome c2/ reaction center in *Rhodobacter sphaeroides* (*167*), when the key tyrosine residue participate in the cation-pi system was replaced by point mutation, the dissociation constant increased three fold (*168*). It has been suggested that in 60% of cation-pi involved protein complexes, there are at least one additional electrostatic interactions on the same interface to improve the association (*169*). Therefore, we speculated that K137 and Y143 improve SrtA dimerization by participating in cation-pi interaction as well. After the initial selective recognition between two residues, the cation-pi interaction will be formed with a favorable enthalpy change. The relative

orientations of the two monomers would then be reorganized so that the residues on the interface can reach to form more hydrogen bonds, electrostatic interactions and hydrophobic interactions. These interactions would further improve the optimal binding configuration and enhance the binding affinity of protein-protein association. Replacement of residues on either of these two sites interferes with the initial recognition and obstructs the dimerization.

The other mutation that can disrupt 93% of the dimerization is N132A. In the crystal structure with packing effect, this residue is not on the surface interacting with any other monomer molecules. It is located within the long loop and helix connecting β4 and β5. The loop structure is usually flexible in response to stimuli such as the long loop between β6 and β7 in SrtA, which can flip around and regulate enzyme activity (*59, 60*). Therefore, it is possible that when N132 is replaced with alanine, the loop structure between β4 and β5 changes in a way that the dimerization interface is reorganized and no longer positioned accordingly. Although N132 is not located on the possible surface of interacting domains, it regulates the folding and overall orientation of β5, where reside the other two key residues K137 and Y143, to ensure the recognition between interacting residues on binding interface.

Hydrophobic residues on protein surface from different chains are also analyzed for close interaction. Three hydrophobic amino acids, F122, I123 and P126 are clustered to form a hydrophobic pocket on SrtA surface. The individual mutation at these sites to glycine excluded the hydrophobic interaction and leaded to a moderate decrease in dimer formation but not the complete disruption. The hypothesis is that the mutation on the residue of I123 and/or P126 can lower the total energy released from forming dimer, and shift the dimer-monomer equilibrium towards monomer. Based on a prevalent theory about the molecular mechanism of protein-protein interaction, the specificity between

two interacting is from the hydrophilic interaction among amino acid side chains (*141*), while the energy drive the dimer formation is from the energy released by burial of hydrophobic residues on the interface when the system is in solution (*140*). The reduced dimer fraction observed in I123G and P126G mutant proteins provided further supports to this theory in SrtA model.

CONCLUSION

Three single point mutated proteins, N132A, K137A and Y143A are generated based on a rational analysis of the current crystal structure of SrtA. They successfully disrupt approximately 93% of the dimerization on $SrtA_{\Delta N59}$ molecules. The disruption may be a result of the direct disturbance of the electrostatic interactions between those residues. The other possibility is that the mutated residues are not directly contacting any partner residues, but instead cause a change in overall protein folding, which leads to the incorrect positioning of contacting residues, and a lower level of protein association between two SrtA molecules. The exact molecular mechanism to mediate SrtA dimerization may be found by X-ray structure of wild type and mutant proteins. However it is possible that the high salt concentration generally used in extreme conditions for crystallization might prevent native SrtA dimer to form and instead, favor a physiologically different and biologically irrelevant structure. A solution NMR structure might provide more insights when comparing the different folding and molecular contact between SrtA dimer and monomer only mutants. The analysis of dimerization/oligomerization of SrtA catalytic domain $SrtA_{\Delta N59}$ obtained *in vitro* and in *E. coli*, suggested that the full-length native SrtA from *S. aureus* cell membrane should behave similarly or form association with higher affinity due to the nature of hydrophobic membrane spanning segment. Although all the mutant proteins exhibited similar reactivity as wild type SrtA based on the apparent velocity obtained from simple *in vitro*

assay, the difference in the efficiency in sorting reaction might be observed when a more biologically relevant function assay can be designed and performed *in vivo* so that the facts of membrane association of both SrtA enzyme and substrates can be taken into consideration.

MATERIALS AND METHODS

Construction of $SrtA_{\Delta N59}$ and mutant protein expression vectors

Primers designated PsrtA59 (5'—CGATCCATGGGCCAAGCTAAACCTCAAA TTCC —3') and PsrtA59R (5'—CCGCTCGAGTTTGACTTCTGTAGCTACAA—3') were used to amplify a $SrtA_{\Delta N59}$ sequence (which would express only residues 60-206) from genomic DNA from *Staphylococcus aureus* subsp. *aureus* (ATCC 700699D) and cloned into the pET28b expression vector to generate the constructs pET28-SrtA59.

The following nine primers were ordered from Invitrogen for the mutations K62A, I123G, P126G, N132A, K137A, Y143A and K152A respectively.

K62A: 5' CGA TCC ATG GGC CAA GCT GCA CCT CAA ATT CC 3'

I123G: 5' GCA GGA CAC ACT TTC GGT GAC CGT CCG AAC TAT 3' and

5' ATA GTT CGG ACG GTC ACC GAA AGT GTG TCC TGC 3'

P126G: 5' ACT TTC ATT GAC CGT GGG AAC TTAT CAA TTT ACA 3' and

5' TGT AAA TTG ATA GTT CCC ACG GTC AAT GAA AGT 3'

N132A: 5' CCG AAC TAT CAA TTT ACA GCT CTT AAA GCA GCC AAA AAA 3' and 5' TTT TTT GGC TGC TTT AAG AGC TGT AAA TTG ATA GTT CGG 3'

K137A: 5' CAA CTT TAA AGT ACA CCA TAC TAC CTT TTG CGG CTG CTT TAAG 3'

Y143A: 5' CAA CTT TAA ATG CCA CCA TAC TAC CTT TTT TG 3'

K152A: 5'GTA CTT TAA AGT TGG TAA TGA AAC ACG TGC GTA TAA AAT G 3'

and 5'TCA GTG GTG GTG GTG GTG GTG 3'

Site directed mutagenesis for I123G, P126G and N132A was performed as described in the Stratagene protocol (*170*) using pET28-SrtA59 as template plasmid. Mutants K62A, K137A, Y143A and K152A were generated with the following protocol. The DNA fragments containing individual mutagenesis were amplified by polymerase chain reaction using forward primer listed above, reverse primer PsrtA59R and pET28-SrtA59 as template plasmid. After double digestion with Nco I and Xho I, the amplified fragment and pET28-SrtA59 plasmid were then ligated to afford the new mutant vectors (*171*). All plasmids were sequenced to confirm the mutation at their respective site (University of Texas at Austin, ICMB core DNA facility).

Purification of $SrtA_{\Delta N59}$ and mutant proteins

The pET28-$SrtA_{\Delta N59}$ and mutant constructs were transformed into *Escherichia coli* strain BL21 (Novagen, La Jolla, CA). The transformed cells were grown in 1 L of Luria broth media at 37 ºC until the OD_{600} reached 0.6. The culture was then induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG, Invitrogen, Carlsbad, CA) and grown for another 6 hrs at 37˚C. The cells were harvested and incubated on ice for 30 minutes with lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0) containing 1 mg/mL lysozyme. After a brief sonication, the lysate was centrifuged and the supernatant was applied to 0.8 mL Ni-NTA agarose beads pre-equilibrated with lysis buffer (Qiagen, Madison, WI). After washing off the unbound contaminating proteins, the His-6×-tagged protein was eluted with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0). All samples were analyzed using Coomassie Plus Protein Assay (Pierce Biotechnology, Rockford, IL) measuring the absorbance at wavelength 595 nm to determine the protein concentration.

Native FPLC with gel filtration column

Purified wild type and $SrtA_{\Delta N59}$ mutant proteins were also applied to a Superdex 200 10/300 GL column (Tricorn, Piscataway, NJ) pre-equilibrated in PBS overnight. For each sample, 100 μL of 5 mg/mL sample was injected except for K137A, which had a lower concentration of 2.3 mg/mL. The samples were eluted with washing buffer (0.1 M sodium phosphate, 0.15 M sodium chloride, pH 6.8) at 0.75 mL/min at 4°C using an AKTA FPLC (Amersham Pharmacia Biotech, Piscataway, NJ). Eluted protein was detected by monitoring at an absorbance of 280 nm. Biomolecules used as molecular weight standards were vitamin B12 (1.35 kDa), horse myoglobin (17 kDa), chicken ovalbumin (44 kDa), bovine γ-globulin (158 kDa), and bovine thyroglobulin (670 kDa), purchased from BioRad (Hercules, CA).

Native PAGE

Appropriate dilutions were made for each protein to a total protein concentration of 5mg/mL prior to analysis by native PAGE using 12% Tris-Glycine Gel (Invitrogen, Carlsbad, CA). Approximately 20ug of each protein was loaded to each lane, along with a NativeMark (Invitrogen, Carlsbad, CA) protein ladder ranged from 20 kDa to 1236 kDa. The gel was visualized by Coomassie Blue staining. For western blot analyses, proteins were transferred to a nitrocellulose membrane (Hybond-C Extra, Amersham Bioscience, Piscataway, NJ) in a Tris-glycine buffered electrophoresis tank. The membranes were then probed with an anti-His C-terminal alkaline phosphatase conjugated antibody (Invitrogen, Carlsbad, CA). The Phospha GLO AP substrate (KPL Incorporation, Baltimore, MD) was applied to visualize the signals, which were detected by exposing the membrane to BioMax Light Film (Eastman Kodak Company, Rochester, NJ). We repeated this experiment 3 times, as well as different protein concentrations (data now shown).

Circular dichroism spectrum

All the measurements have been repeated three times. Each data point was recorded in triplet. Circular dichroism (CD) studies were conducted with a Jasco J-810 spectropolarimeter (Jasco International Co., Ltd. Tokyo, Japan), and equipped with a Jasco PFD-425S temperature control system. For measurements, all protein samples are dialyzed against activity assay buffer (50 mM Tris-Cl, 150 mM NaCl, 5 mM $CaCl_2$, pH 7.5) and diluted to a final concentration of 38 μM. CD spectra were recorded from 190 nm to 250 nm at 25°C, using a 1 mm cell, a path length of 0.1 nm, a scan rate of 50 nm/min and a response time of four seconds. Spectra were smoothed, and baseline corrected by solvent subtraction before analysis. The recorded spectra in millidegrees of ellipticity (θ) were converted to mean residue ellipticity [θ] in deg.cm^2/dmol.residue by following equation (Equation 3.1).

Equation 3.1 $$[\theta] = \frac{\theta \times 100 \times M_r}{c \times l \times N_A},$$

where c is the protein concentration in mg/mL, l is the pathlength in cm, M_r is the protein molecular weight and N_A is the number of amino acids in the protein. The mean residue molecular weights of 114.4 for wild type SrtA, 114.04 for K62A, K137A and K152A, 114.05 for I123G, 114.15 for P126G, 114.13 for N132A, 113.82 for Y143 were used for calculation. The mean residue ellipticities were then plotted against wavelength by the software KaleidaGraph (version 3.6).

The recorded spectra were also analyzed using the Dichroweb server(*165*). Different algorithms were used for the structure calculations: SELCON3(*172*), CONTIN(*173, 174*), and CDSSTR(*164*). Reference data 4 and 7 contain the wavelength 190 nm-240 nm were used to be compatible with our reading wavelength range(*164*).

NRMSD parameter was used to evaluate the fitness of various methods(*175*). NRMSD value was calculated as following equation:

NRMSD=$\sum[(\theta_{exp}-\theta_{cal})^2/(\theta_{exp})^2]^{1/2}$, (Equation 3.2).

Where θ_{exp} is the experimental ellipticities, θ_{cal} is the ellipticities of the back-calculated spectra for the derived structure. A NRMSD value less than 0.1 means the predicted spectra and the experimental spectra are in close agreement(*176*).

***In vitro* activity assay**

Purified wild type and mutant proteins of SrtA were further studied for their activity. The peptide substrate *o*-aminobenzoyl-LPETG-(2,4-dinitrophenyl)-diaminopropionic acid (Abz-LPETG-Dap(Dnp)) was custom synthesized on PAL resin based on the well established Fmoc/piperidine strategy (Protein Microanalysis Facility, The University of Texas at Austin). The peptide was cleaved using TFA:water (95/5, v/v) and precipitated using cold diethyl ether. It was then filtered with a fine-porosity fritted glass filter, dissolved in water and lyophilized to dryness. The crude peptide product was purified by HPLC using a preparative C18 column to give ≥ 90% purity. The peptide Gly_5 was purchased from BACHEM (Torrance, CA). Assays were performed in a volume of 100 μL assay buffer (50 mM Tris-Cl, 150 mM NaCl, 5 mM $CaCl_2$, pH 7.5) containing 8.4 μM $SrtA_{\Delta N59}$ (or respective mutant proteins), 0.1 mM Abz-LPETG-Dap(Dnp)-NH_2, and 2 mM pentaglycine. Reactions were initiated by the addition of enzyme and incubated at 37˚C for 4 hours. Reactions were quenched by the addition of 50 μL 1N HCl prior to injection into a reversed-phase C18 fast analytical HPLC column. The peptides of substrate and product were separated using a linear gradient from 10% to 90% CHCN/0.1% TFA over 10 minutes. Dnp-containing peaks were detected by absorbance at 355 nm, and the percentage of reaction product was calculated by integrating the area under the HPLC trace(*85*). To confirm the composition and identity of each product, the

peaks were collected and analyzed by MALDI-TOF mass spectrometry and peptide fingerprinting analyses (Analytical Instrumentation Facility Core, The University of Texas).

L-DOPA incorporated protein cross-linking assay

The following three pairs of primers were ordered from Invitrogen to generate the mutations K137TAG, K206TAG and K206A respectively.

pET28-SrtA59K137TAG: 5'-CGATCCATGGGCCAAGCTAAACCTCAAATTCC-3' and 5'-CAACTTTAAAGTACACCATACTACCTTTCTAGGCTGCTTTAAG-3'.

pET28-SrtA59K206TAG: 5'-GTAGCTACAGAAGTCTAGCTCGAGCACCACCACC-3' and 5'-GGTGGTGGTGCTCGAGCTAGACTTCTGTAGCTAC-3'.

pET28-SrtA59K206A: 5'-GTAGCTACAGAAGTCGCGCTCGAGCACCACCACC-3' and 5'-GGTGGTGGTGCTCGAGCGCGACTTCTGTAGCTAC-3'.

Site directed mutagenesis for K206TAG and K206A was performed as described in the Stratagene protocol(*170*) using pET28-SrtA59 as template plasmid. Mutant K137TAG, was generated with PCR based approach as described earlier(*171*). The vector pAC-DHPheRS-6TRN that harbors one copy of mutant *Methanococcus jannaschii* Tyr-tRNA synthetase and six copies of mutant *M. jannaschii* $tRNA^{Tyr}$ was a gift from Dr. Peter Schulz (The Scripps Research Institute, La Jolla, CA).

Escherichia coli BL21 (Novagen, La Jolla, CA) was transformed with two plasmid. One is for the expression of SrtA protein, the other plasmid is pAC-DHPheRS-6TRN to provide orthogonal tRNA and tRNA synthetase. 50 μg/mL kanamycin and 12.5 μg/mL tetracycline were used to select the colony with both plasmids. Single colony selected was subcultured in LB broth supplemented with antibiotics and 1 mM glucose, and grown overnight at 37°C. The overnight subculture was collected and resuspended in sterile deionized water, and then inoculated into glucose minimal media supplemented

with antibiotics. Cultures were grown at 37°C until OD_{600} reached 0.7. Tetracycline concentration was increased to 18.75 μg/mL and 1 mM 3,4-dihydroxy-L-phenylalanine (L-DOPA, Acros Organics, NJ) was added. After the addition of L-DOPA, the cultures were protected from direct exposure to light to prevent the oxidation on L-DOPA. Cultures were then incubated at 30°C for another 40 minutes before the addition of 1 mM IPTG. Cultures were grown for additional 5 hours at 30°C and the cells were harvested and frozen at -80°C. The frozen cells obtained were thawed on ice and resuspended in lysis buffer with 1 mM sodium periodate (Acros Organics, NJ). The mixtures were incubated on ice for 1 hour. 1 mg/mL Lysozyme was added and the samples were incubated on ice for additional 30 minutes. The concentration of sodium periodate was varied between 1 mM to 2.5 mM depending on the experiment. The incubation time after adding the sodium periodate and lysozyme was also varied between 20-70 minutes and 30-60 minutes, respectively. The procedures for subsequent protein purification, SDS-PAGE analysis and western blot were the same as previously described. The wash buffer in purification was modified to contain one or all of the following: 10% glycerol, 0.1% Tween-20, 20 mM β-mercaptoethanol, 2 M NaCl, and 30 mM imidazole.

Chapter 4 Sortase A dimerization *in vivo*

INTRODUCTION

In Chapter 2 and 3, we demonstrated various genetic and biochemical data from our lab to prove that the C-terminal catalytic domain of SrtA specifically self-associates to form a homo-dimer *in vitro* and in *E. coli* cytoplasm with an estimated K_d of 55 μM (*100, 101, 177, 178*). In this chapter, we provide evidence that the full-length SrtA on the *S. aureus* cell membrane exists as a homo-dimer with a higher affinity *in vivo* than the truncated SrtA can form *in vitro*. In addition, we have identified a SrtA mutant protein that is incapable of homo-dimerization *in vivo*. *S. aureus* strains harboring a SrtA monomer only mutant and a wild type protein were tested for their capability to invade human epithelial cells. Our data suggested that the strain with a SrtA monomer only mutant has higher invasion capability. The discoveries in this chapter provided a possible function of SrtA homo-dimerization *in vivo*, which lead to a more comprehensive and accurate understanding of the cellular and molecular mechanism of SrtA function and its regulation. New inhibitors for SrtA function may be designed based on this mechanism to target SrtA self-association.

RESULTS AND DISCUSSION

Expression of wild type SrtA protein in *S. aureus* RN4220 and a SrtA knockout strain

Endogenously expressed full-length SrtA protein from *S. aureus* RN4220 was first analyzed by using SDS-PAGE to resolve the cell lysate followed by western blot. A commercial anti-sortase A polyclonal antibody was used to detect three bands (Figure 4.1). The fastest eluted band, approximately 28 kDa, corresponded to the estimated size

of the full-length SrtA monomer. However, the two additional bands with sizes about 50 kDa and 55 kDa cannot be attributed to SrtA. In a separate experiment, the anti-sortase A antibody also detected at least three random bands from *E. coli* cell lysate suggesting the poor selectivity of this antibody towards SrtA (data not shown). Therefore, a specific epitope tag such as a His6 and/or FLAG sequence fused to the C-terminal end of SrtA would be necessary to monitor its distribution in *S. aureus* and facilitate its *in vitro* purification.

A SrtA knockout strain, *srtA*$^-$, was generated from *S. aureus* strain RN4220. TargeTron™ Gene Knockout System from Sigma-Aldrich was used to generate the knockout strain. It allows for the permanent and site-specific disruption of target genes on bacterial genomic DNA through the insertion of group II introns (*179*). Southern blot with genomic DNA extracted from the knockout strain confirmed the site-specific insertion of the intron into the srtA gene and only srtA gene (data not shown). The absence of SrtA expression in *srtA*$^-$ was also confirmed using western blot. As shown in Figure 4.1, the fastest eluted band of approximately 28 kDa was absent from the cell lysate of the knockout strain while it was present in the cell lysate from RN4220 strain, suggesting that the absent band in *srtA*$^-$ lysate was the monomer of full-length SrtA. The two non-specific bands were also detected from both RN4220 and *srtA*$^-$.

To express full length SrtA, we generated a *srtA*$^-$*+A* strain by using the knockout strain *srtA*$^-$ as the host. A plasmid, pT-P-srtA-H-FLAG, was first created to contain the endogenous srtA promoter, full-length srtA gene, and followed by a sequence of His6 and FLAG tags on its C-terminal end. Next, the plasmid pT-P-srtA-H-FLAG was transformed into the knockout strain *srtA*$^-$ thereby allowing us to express His6 and FLAG tagged full-length SrtA (SrtA-His6-FLAG). The expression of SrtA-His6-FLAG was probed using anti-FLAG antibody as shown in Figure 2A. We observed a single SrtA

monomer band at the size of 28 kDa and a faint band at around the size of 50 kDa. The 50 kDa band was identified as a nonspecific endogenous protein recognized by various antibodies including a monoclonal anti-S-tag antibody (data not shown). We also performed immunoprecipitation experiment to analyze this 50 kDa protein, it was indicated to be *S. aureus* surface protein A (SPA), which binds to the constant region of the IgG (*180*) (data not shown). SrtA-His6-FLAG expression was also probed with the anti-sortase A antibody (Figure 4.1). The 28 kDa band of SrtA monomer was successfully restored into the knockout strain, and no additional bands were introduced. These experiments indicated that the wild type SrtA was successfully knocked out in the *srtA*$^-$ strain and replaced with a SrtA-His6-FLAG protein in *S. aureus srtA*$^-$*+A* strain.

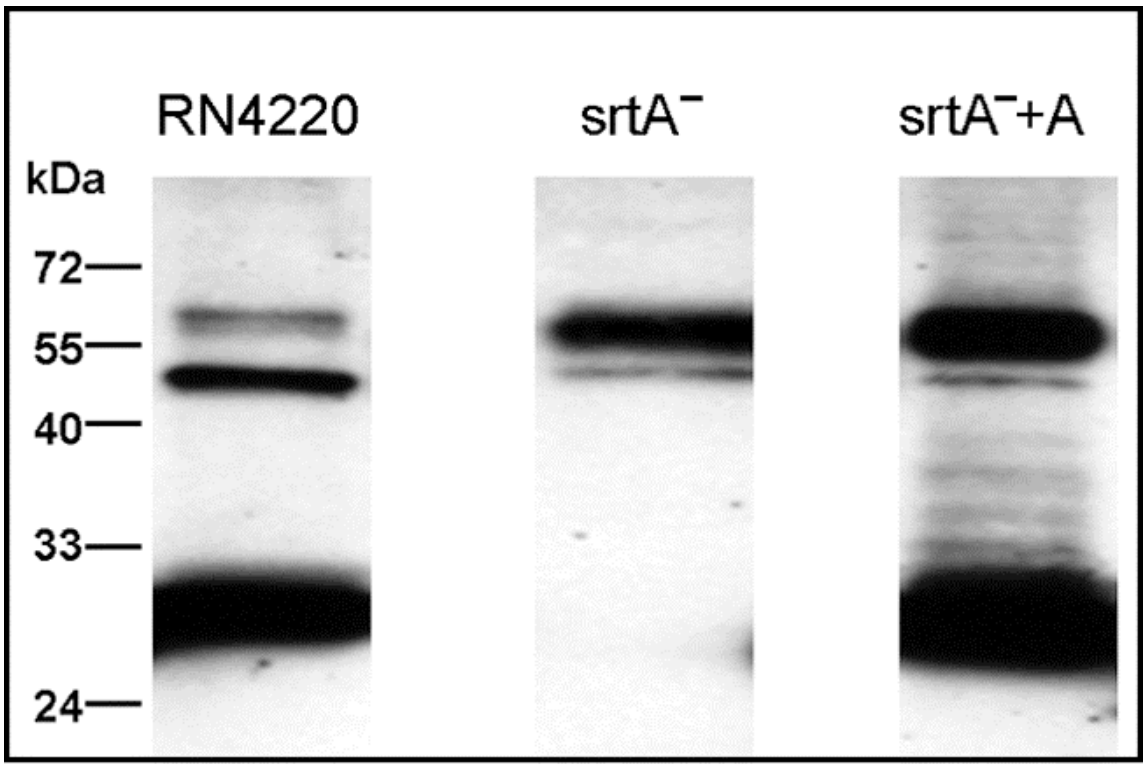

Figure 4.1: Full-length SrtA protein expression in *S. aureus* strains.

Cell lysate from wild type RN4220, SrtA knockout strain *srtA*$^-$ and SrtA expressing strain *srtA*$^-$*+A* were subjected to SDS-PAGE analysis. Western blot using anti-sortase A

antibody was used to visualize the protein bands. Three major bands were observed in RN4220, among which, the fastest eluted bottom band around 28 kDa corresponded to SrtA monomer. This band was eliminated in SrtA knockout strain, and restored in SrtA expressing stain.

***In vivo* chemical cross-linking of SrtA expressed in *S. aureus*.**

An *in vivo* chemical cross-linking experiment was performed to demonstrate that SrtA specifically forms a dimer *in vivo*. 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), a compound that couples carboxylic acid groups with primary amines in close proximity was used to stabilize the SrtA dimer by cross-linking two molecules together (*120*). The cell culture of *srtA*$^{-}$*+A* was briefly treated with EDC prior to harvesting. Next, the cell lysate was resolved by SDS-PAGE and the SrtA-His6-FLAG protein was visualized by western blot using the anti-FLAG antibody (Figure 4.2A). In addition to the 28 kDa monomer band, another strong 56 kDa band was detected which corresponded to the size of SrtA homo-dimer.

Full-length SrtA-His6-FLAG, expressed and cross-linked *in vivo,* was purified from *S. aureus* by using Ni-NTA agarose beads followed by mass spectrometry for protein identification. The purified protein was first resolved based on size under denaturing conditions on SDS-PAGE and the resulting Coomassie Blue stained gel was shown in Figure 4.2B. Two protein bands of interest with molecular weight of 28 kDa and 56 kDa were excised from the gel and subjected to in-gel trypsin digestion followed by MALDI-TOF and peptide fingerprint mapping analyses. Mass spectrometry results identified both the 28 and 56 kDa protein bands as SrtA with Mascot Search scores of 607 and 357 respectively. This suggested that these species represent the monomeric and dimeric forms of the enzyme. Within the 56 kDa band, there was another protein identified as a putative acetyl-CoA transferase with a Mascot Search score of 524 and an estimated molecular weight of 58 kDa. We did not observe a higher molecular weight

band that would have corresponded to the size of the hetero-dimer between this protein and SrtA. Due to common technical challenges associated with the purification of membrane proteins, such as their adhesion to lipid and other lipid associated proteins, we believe that the acetyl-CoA transferase is most likely a contaminant that happened to be co-purified and co-migrate with the SrtA dimer on SDS-PAGE. Current data suggested with high confidence that the 56 kDa band was the homo-dimer band of SrtA.

Mass spectrometry data also confirmed significant SrtA degradation in truncated form (Figure 4.2B), with a Mascot Search score of 557. The degradation was not observed previously when we had denatured the cell lysate immediately after lysing the cell (Figure 4.2A). Because this degraded form of SrtA was recognized by the anti-FLAG antibody (Figure 4.4B), its C-terminus should still be intact, suggesting that the possible cleavage site may be on its N-terminus. The N-terminus of SrtA contains a putative signal peptide cleavage site between amino acids 24 and 25. Since SrtA is a trans-membrane protein, the signal peptide sequence also functions as a stop transfer signal for membrane anchoring, which is not removed (*92*). Degradation was only observed in full-length SrtA expressing in *S. aureus*, but not in *E. coli* strain (data not shown). These data suggested that there might exist an *S. aureus* and/or gram-positive bacterial regulatory mechanism for controlling the membrane content of SrtA, which is similar to another quality control mechanism for membrane proteins, "regulated intramolecular proteolysis" (RIP) (*181, 182*), so that the large amount of full-length SrtA present in *S. aureus* cytoplasm would be subjected to proteolysis to prevent the potential damage to the cells.

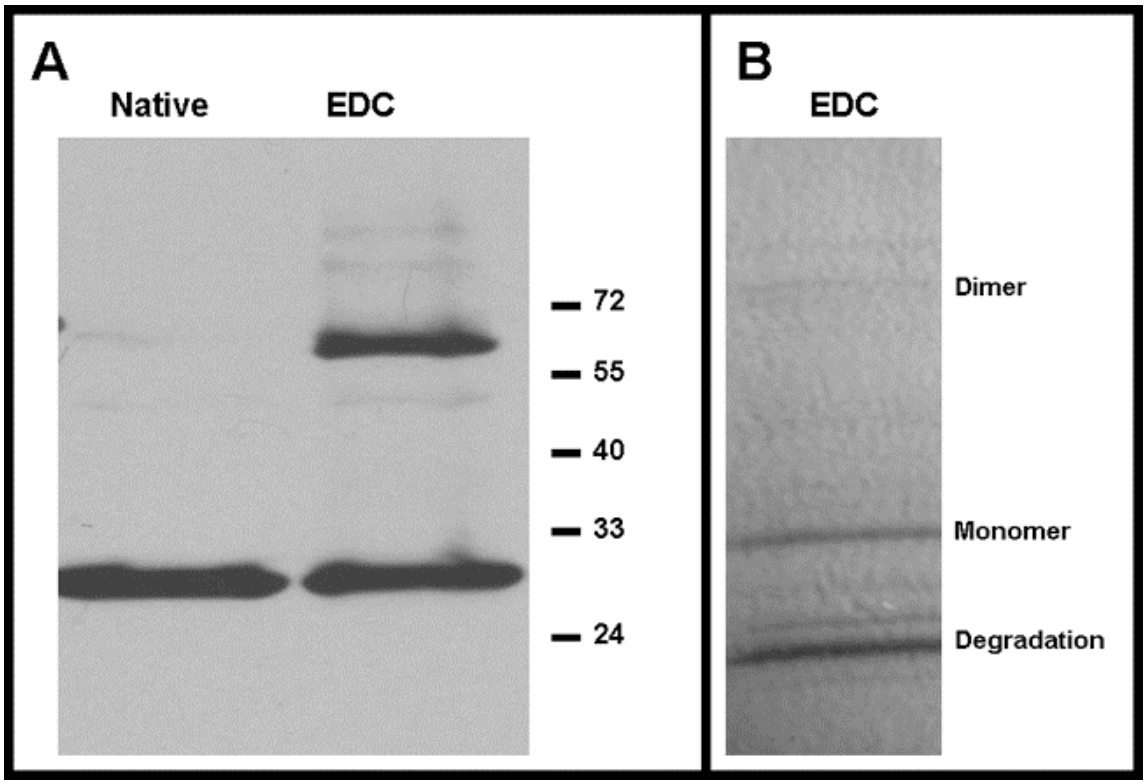

Figure 4.2: Chemical cross-linking of SrtA protein in *S. aureus*.

Full-length SrtA was fused with a His6/FLAG tag on C-terminus and exogenously expressed in *S. aureus srtA*$^{-}$. Cell lysate was analyzed by SDS-PAGE and proteins were visualized by western blot with anti-FLAG M2 antibody (Panel A). Cells were treated with EDC cross-linking reagent (labeled with EDC) or not (labeled with native) before harvesting. SrtA protein purified from EDC treated cells were resolved by SDS-PAGE and visualized by Coomassie Blue staining (Panel B). Dimer, monomer and degradation bands as labeled were identified by mass spectrometry data.

Cellular localization of SrtA

The exogenously expressed SrtA-His6-FLAG was found to be mainly localized on the cell membrane and exist in dimer form. Cell lysate from *S. aureus* was centrifuged at 21, 000 x g for 1 hour at 4°C to separated into a cytosolic fraction (lane C, Figure 4.3) and a cell membrane fraction (lane M, Figure 4.3). Each fraction was separated by SDS-PAGE and analyzed using western blot. The cytosolic fraction contained only trace

amount of the SrtA monomer band (lane C, panel A, Figure 3). A more intense SrtA monomer band was detected in the membrane fraction. This suggested that the cell membrane fraction contained the majority of the exogenously expressed SrtA-His6-FLAG (lane C, panel B, Figure 4.3). Another cell culture was treated with EDC prior to harvesting, and fractionated in an attempt to stabilize the dimer form *in vivo* on cell membrane. A cross-linked dimer band was only observed in the membrane fraction but not in the cytosolic fraction (panel B, Figure 4.3). These data indicated that SrtA protein formed a dimer when expressed and anchored on the *S. aureus* cell membrane. Exogenously expression and the presence of a His6/FLAG tag on its C-terminal end did not seem to interfere with the cellular localization of SrtA thereby potentially retaining its biological integrity.

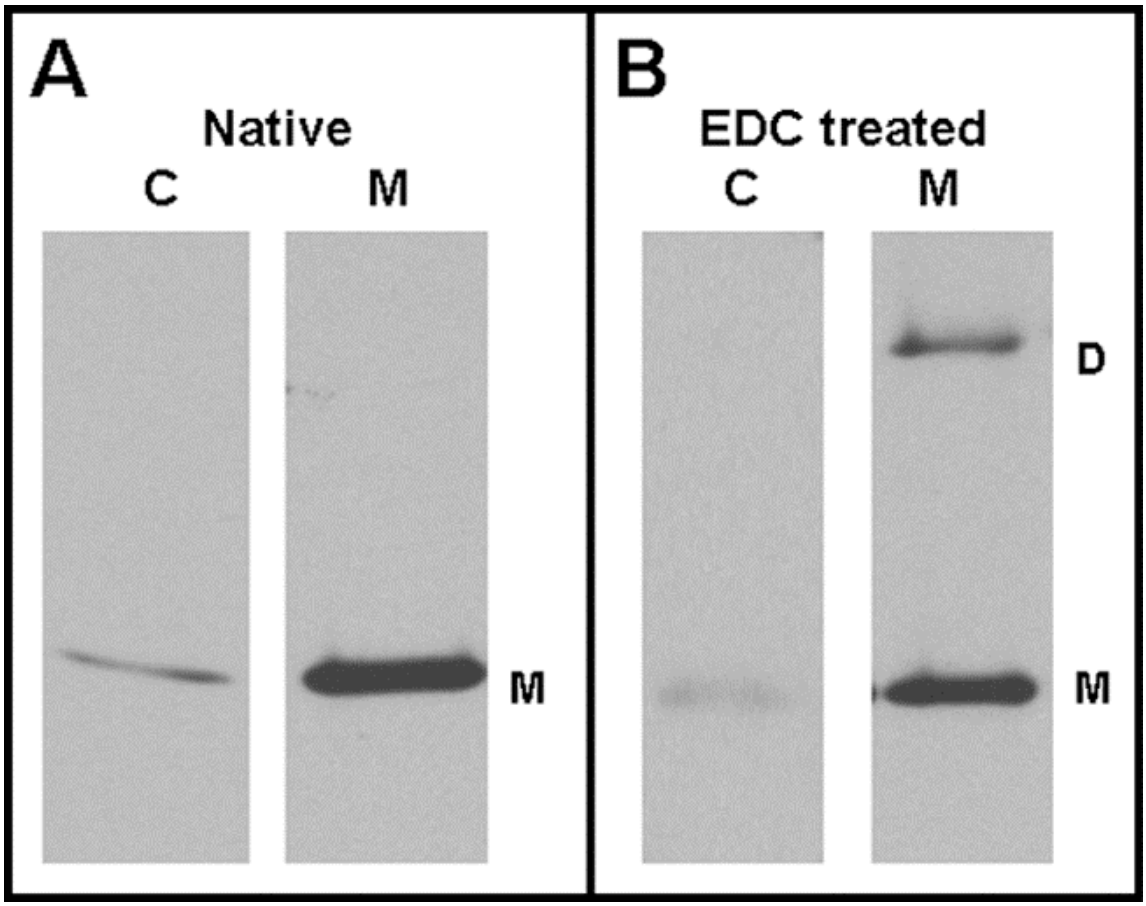

Figure 4.3: Cellular localization of over-expressed SrtA and its dimerization.

S. aureus cells exogenously expressing SrtA were treated with EDC cross-linking reagent (Panel B) or not (Panel A) before harvesting. Cell lysates were then separated into cytosol fraction (labeled with C) and membrane fraction (labeled with M). Proteins were resolved by SDS-PAGE and visualized by western blot with anti FLAG antibody. SrtA was mainly located on cell membrane, as well as its dimerization.

If SrtA is transported through Sec secretion system (*183*), similar as the other membrane proteins, a signal peptide sequence directs unfolded SrtA-His6-FLAG from the cytosol to the cell membrane. Once the nascent protein is translocated across the membrane, it starts folding into its regular conformation, which is favoring dimerization. However, the localization of *in vivo* expressed SrtA onto the cell membrane is actually a complicated process that might regulate SrtA's homo-dimerization and the dynamic

display of surface virulence factors as well. This process is regulated by the secretion pathway and its accessory proteins. Currently, much effort has been put forth towards studying how cellular mechanisms determine the spatial distribution of SrtA within the membrane so that it is in close proximity to its surface protein substrate (*71, 184, 185*). It has been found in *Streptococcus pyogenes,* that SrtA co-localizes with the anchoring areas of one of its substrates, M protein, which is within distinct membrane foci associated with the division septum (*71*). Further study focusing on the mechanism of SrtA molecule deposition onto the membrane would be of great importance and could be employed by the cell as a potential pathway to regulate SrtA function *in vivo*.

Gel filtration of SrtA under native condition

To further provide evidence that full-length SrtA has a high tendency to dimerize we used size-exclusion chromatography to analyze the SrtA-His6-FLAG purified from *S. aureus* under native conditions. Purified SrtA-His6-FLAG protein, without being treated with EDC cross-linking, was loaded onto a gel filtration column and eluted under native conditions, yielding four distinctive peaks (Figure 4.4A). We collected the fractions belonging to these peaks and analyzed them using denaturing SDS-PAGE followed by western blot probed with anti-FLAG antibody (Figure 4.4B). The first and largest peak (peak 1) contained mainly the SrtA monomer. A small amount of the dimer was observed, even under denaturing conditions. The three smaller peaks (peak 2, 3 and 4) that eluted later were a mixture of the monomeric form of full-length SrtA and its degradation products. To confirm that peak 1 is the dimer peak we analyzed another sample of purified SrtA-His6-FLAG protein, which was first cross-linked *in vivo* with EDC and tested using the same gel filtration condition. In this case, the dimer band in peak 1 fraction became much more pronounced (Figure 4.4C, right lane). The data from this experiment indicated that the peak 1 fraction was the SrtA dimer. The ratio of SrtA

dimer to monomer under this non-denaturing condition was estimated based upon the integrated area of the dimer peak (peak 1) to that of the three monomer peaks (peak 2, 3 and 4) on the elution trace for native purified protein (Figure 4.4A). The ratio was calculated to be 2.94:1 at a protein concentration of 7.4 μM. Previous *in vitro* analytical ultracentrifugation experiments with the truncated form of SrtA showed a moderate K_d value of 55 μM (*100*), and the ratio of dimer to monomer was calculated to be approximately 0.23:1 under the same experimental condition and protein concentration. A more than 10-fold increase of dimer formation was observed when the *in vitro* truncated SrtA in solution was translated in its full-length form *in vivo*. These data suggested that the full-length SrtA exogenously expressed *in vivo* had a much lower K_d value under native conditions when compared to the *in vitro* truncated form,

We also tested to see if the dimerization could be attributed to the formation of a disulfide bond between the two molecules. Addition of 5 mM β-mercaptoethanol to protein purification and gel filtration steps did not change the chromatography profile, suggesting that the observed dimerization is not attributed to disulfide bonds. Full-length SrtA only contains one cysteine residue at its active site. It would be highly unlikely, due to steric hindrance, for this cysteine residue to form a disulfide with another cysteine residue located in the cavity of the active site.

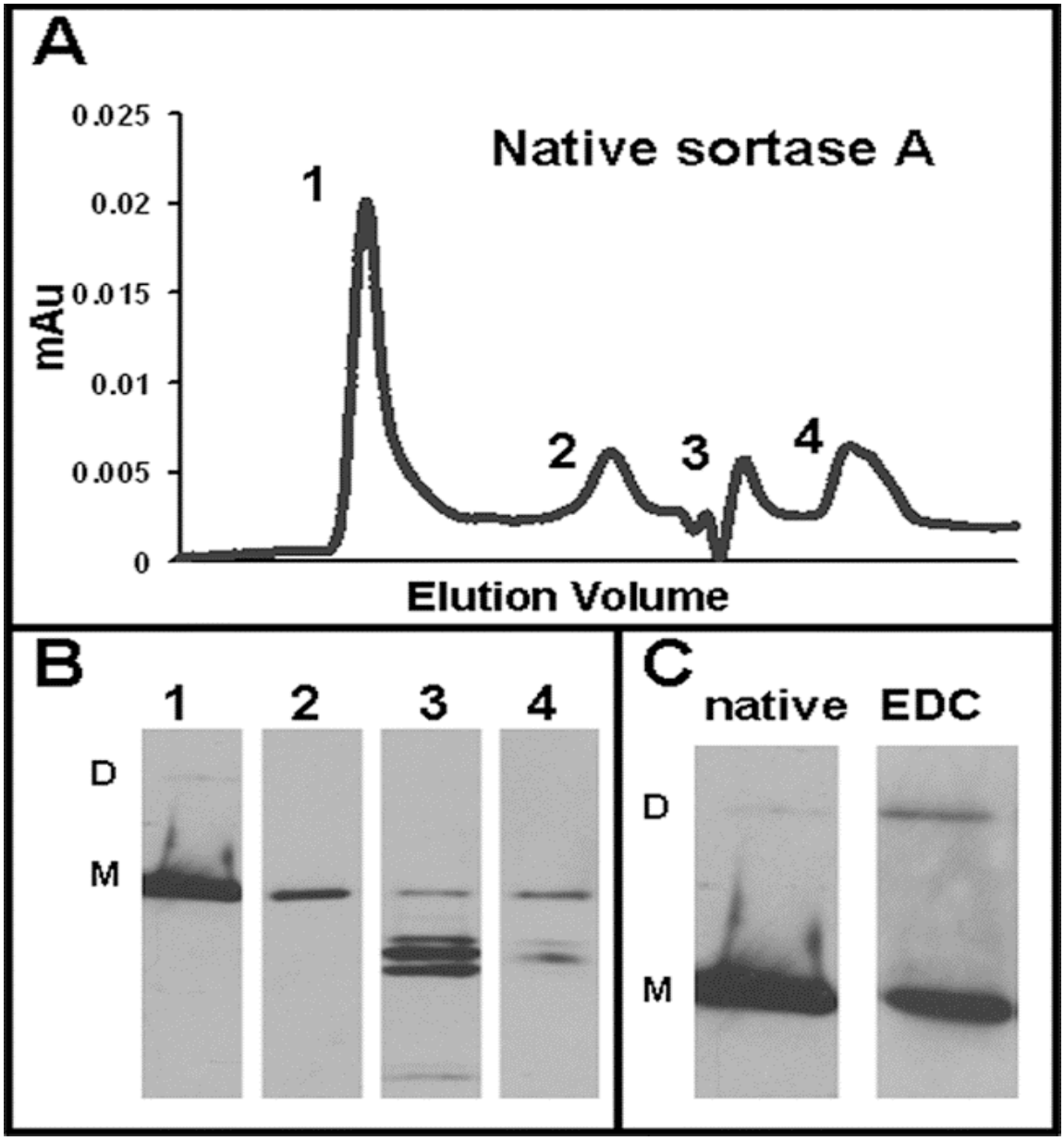

Figure 4.4: Size-exclusion chromatography of SrtA protein.

SrtA protein purified from cells without EDC treatment was applied on native condition size-exclusion chromatography to separate dimer and monomer fractions (Panel A). Peak fractions (labeled as 1 to 4 as fraction 1 to 4) were collected and analyzed by SDS-PAGE (Panel B). Peak 1 contains very faint dimer band even under denaturing electrophoresis condition. SrtA protein purified from EDC treated cells were also separated by FPLC. Peak 1 fraction was analyzed by SDS-PAGE and western blot (Panel C). The intensified dimer band confirmed the identity of peak 1 as dimer fraction.

Site-directed mutagenesis to disrupt SrtA dimerization *in vivo*

Next, we used site-directed mutagenesis to generate a full-length SrtA mutant incapable of forming a dimer *in vivo*, which is not crosslinked by EDC treatment. Previously, three mutant proteins in SrtA truncated form have been shown to exist as a monomer *in vitro* (*101*). Each of those mutant proteins contained a single point mutation of an amino acid residue on protein surface (N132A, K137A and Y143A respectively). To test if any of these single mutations could disrupt the dimerization of full-length SrtA *in vivo*, mutagenesis was introduced on the same site to generate three mutant proteins, N132A, K137A and Y143A. The SrtA knockout strain *srtA*$^-$ was used as the host to individually express three mutant proteins. Cell cultures were treated with EDC cross-linking reagent before harvesting and the cell lysates were analyzed by SDS-PAGE and visualized by western blot with anti-FLAG antibody (Figure 4.5). The 28 kDa SrtA monomer band was observed for wild type and all three mutant proteins. The intensity of the bands for mutant proteins was comparable to the monomer band of the wild type protein, confirming the regular expression level of these mutant proteins. However, the 56 kDa EDC cross-linked dimer band was also observed in the cell lysate of all three mutant protein expressing strains, which suggested that these three single residue mutant SrtA proteins formed a dimer *in vivo*. The same single-mutation that dramatically reduced the dimerization of SrtA catalytic domain when they were expressed in *E. coli* was unable to completely disrupt the dimerization of full-length SrtA expressed *in vivo*.

To further reduce dimerization tendency of full-length SrtA, we generated a mutant protein, which harbored all three mutations simultaneous. A significant reduction in dimerization was realized upon replacement of all three key hydrophilic residues (N132, K137, and Y143) with alanine residues. The SrtA knockout strain exogenously expressing $SrtA_{NKY}$ was treated with EDC prior to harvesting and the cell lysate was

analyzed using anti-FLAG antibody as described above. As shown in Figure 4.5, the 28 kDa SrtA monomer band was observed with a similar intensity comparable to the wild type protein and other mutant proteins. The 56 kDa dimer band was absent from this strain indicating the inability of $SrtA_{NKY}$ to form a dimer *in vivo*. Therefore the cumulative effect of all three mutations is strong enough to disrupt the dimerization whereas the effect of each single mutation is not. The disruption of full-length SrtA dimerization *in vivo* also suggested that the C-terminal His6/FLAG tag did not contribute to dimer formation on *S. aureus* cell membrane.

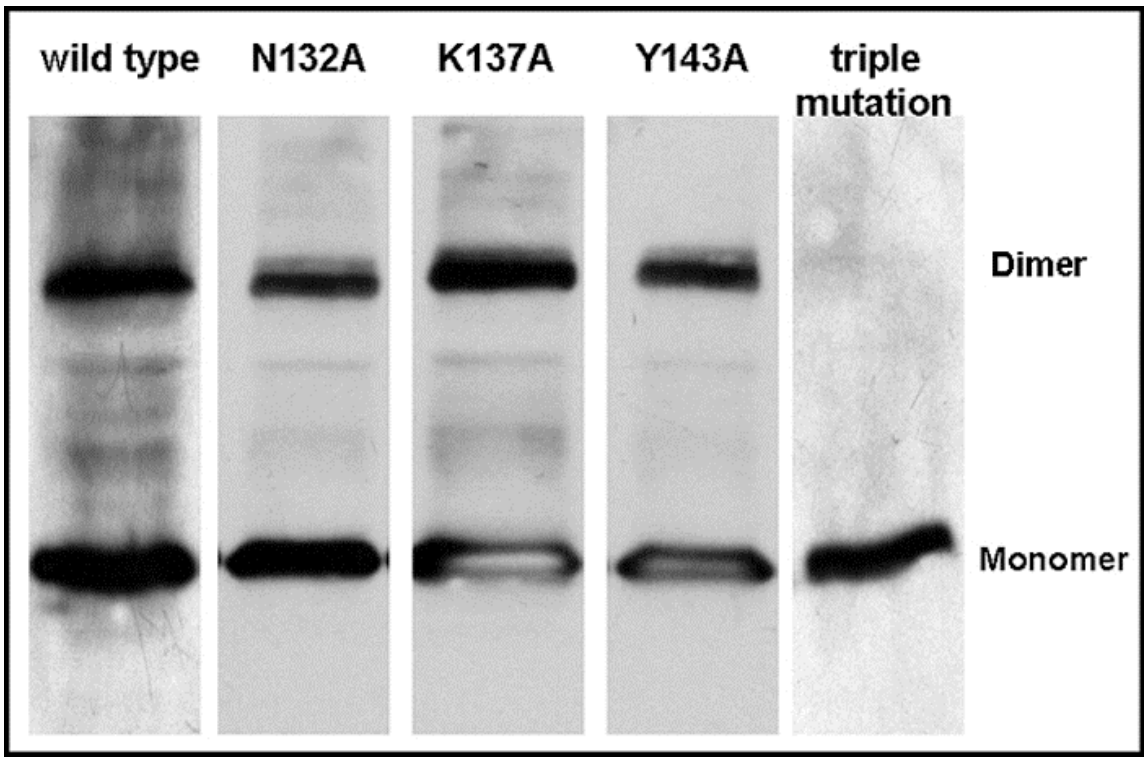

Figure 4.5: SDS-PAGE with full-length SrtA wild type and mutant proteins expressed from *S. aureus*.

S. aureus cells exogenously expressing SrtA wild type (lane 1) and mutant proteins (lane 2 for N132A, lane 3 for K137A, lane 4 for Y143A, lane 5 for $SrtA_{NKY}$) were treated with EDC cross-linking reagent before harvesting. Cell lysates were analyzed by SDS-PAGE. SrtA proteins were visualized by western blot using anit-FLAG M2 antibody.

To assess the folding behavior of the mutant protein $SrtA_{NKY}$, circular dichroism spectrum analyses were performed with wild type $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59,\ NKY}$ mutant. We chose the truncated form of SrtA since $SrtA_{\Delta N59}$ contains the catalytic domain with all three mutations and is much easier to purify *in vitro* (*44, 50*). We hypothesizes that if the catalytic domain retains its correct folding then the full-length SrtA should fold correctly as well. The recorded spectra in millidegrees of ellipticity (θ) were converted to mean residue ellipticity [θ] as indicatives of changes of secondary structure in proteins (Figure 4.6). The estimated contents of secondary structures of $SrtA_{\Delta N59,\ NKY}$ were derived from the Dichroweb server and listed in Table 1 (*165*). A NRMSD fit parameter of 0.051 was achieved for $SrtA_{\Delta N59,\ NKY}$, indicating the consistency of the estimated secondary structures and the experimental measurement. For comparison, content of secondary structures of wild type SrtA was calculated based on the published crystal structure of $SrtA_{\Delta N59}$ (1T2P from PDB) and also included in Table 4.1.

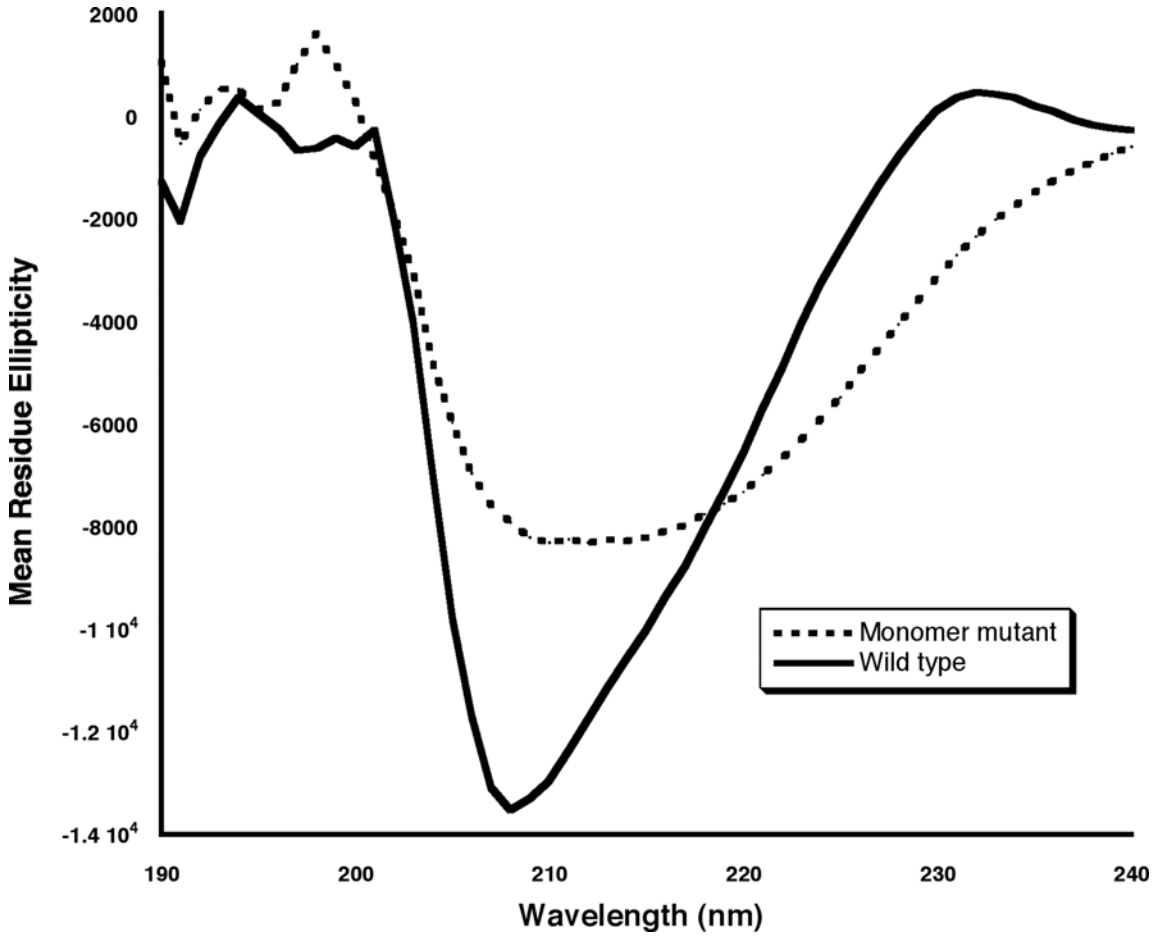

Figure 4.6: Circular Dichroism spectrum of $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59,\ NKY}$.

Purified $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59,\ NKY}$ were scanned using circular dichroism spectroscopy. Each sample was analyzed at protein concentration of 39 μM and scanned from 190 nm to 250 nm. The recorded spectra measured in millidegrees of ellipticity (θ) was converted to mean residue ellipticity [θ] in $deg.cm^2/dmol.residue$ and then plotted against the scanning wavelength using the software KaleidaGraph (version 3.6). Wild type protein is shown as the solid line while $SrtA_{\Delta N59,\ NKY}$ is in the dotted line.

Table 4.1: Estimated secondary structure conformation of $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59, NKY}$ proteins.

Fraction Ratio	Wildtype[α]	Wildtype	$SrtA_{\Delta N59,NKY}$
NRMSD		0.030	0.051
Helix:	6%	3%	8%
Beta:	44%	45%	37%
Turn:	20%	21%	23%
Random:	30%	31%	31%
Total:	100%	100%	99%

Purified $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59, NKY}$ proteins were scanned using circular dichroism spectroscopy. Estimated secondary structure conformation was calculated by Dichroweb server. [α]As a reference, published X-ray crystal structure of $SrtA_{\Delta N59}$ 1T2P from PDB was used to calculate secondary structure content, which was listed in the first column.

$SrtA_{\Delta N59, NKY}$ contained 8% α-helix, 37% β-sheet, 23% β-turn, and 31% random loop, which was comparable to wild type $SrtA_{\Delta N59}$. In general, when compared to the wild type protein, $SrtA_{\Delta N59, NKY}$ had fewer β-sheet but more α-helix structures, and the β-turns and random structures were unchanged. Based upon this data, the global secondary structures of the protein remained largely unchanged, but the local structure at the dimerization interface appeared to be disturbed in such a way as to disrupt dimerization. The mechanism for disruption of SrtA dimerization by the triple mutation employed in these studies will remain unsettled until a crystal structure can be obtained.

The difficulty in disrupting SrtA dimerization *in vivo* suggested a high association affinity between the two SrtA molecules. *In vitro* experiments showed that a single mutation on any of the following three amino acid residues in the catalytic domain itself, N132, K137 and Y143, was enough to disrupt the dimerization by abolishing the hydrogen bonds, electrostatic interactions and hydrophobic interactions that hold two monomers together. However, when SrtA was expressed in its full-length form on the *S. aureus* membrane, single mutations on any of three residues can only lower the total

energy of dimerization and shift the equilibrium towards the monomer form. None of the three residues by itself can completely disrupt the homo-dimer. Only the $SrtA_{NKY}$ accumulated the effect of all three mutations and significantly reduced dimer formation. The similar phenomena that compared to the truncated form, full-length protein expressed *in vivo* forms a dimer with a higher affinity to activate its activity was also observed in other proteins, such as Outer Membrane PhosphoLipase A (OMPLA) in gram-negative bacteria *E. coli* (*102*).

The increased dimerization tendency of full-length SrtA expressed *in vivo* can be attributed to several possible explanations. Dimerization of the hydrophobic transmembrane domain would be favored due to a decrease in the free energy. Additionally, the transmembrane domain of *in vivo* expressed SrtA might interact specifically with other molecules within the membrane thereby forming a complex that could reinforce the intermolecular association between SrtA molecules. This type of association for a multi-enzyme complex has been proposed and experimentally supported in the peptidoglycan biosynthesis holoenzyme mechanism (*67, 186*). Another possibility is that the conformation of full-length SrtA expressed in *S. aureus* might be different from that expressed in *E. coli* since the catalytic domain might directly or indirectly interacting with the linker region, which was neglected when the *in vitro* conformation of SrtA was resolved. Our efforts to explore the mechanism are described in Chapter 5.

SrtA monomer confers more efficient *S. aureus* invasion into mammalian epithelial cells

A quantitative mammalian-cell invasion assay was applied to test the relationship between SrtA's *in vivo* activity and its oligomerization status on the cellular membrane (*187-189*). SrtA was responsible for the display of various LPXTG motif-bearing proteins that were identified as bacterial adhesion proteins (Spa, FnBP-A, FnBP-B, ClfA, ClfB,

Cna, SdrC, SdrD, SdrE) onto the surface of *S. aureus*. Among them, two fibronectin-binding proteins (FnBP-A and FnBP-B) have been demonstrated to be sufficient for *S. aureus* to enter non-professional phagocytes *in vitro*, such as various epithelial and endothelial cells (*190-192*). They bind to the extracellular matrix protein fibronectin (Fn) in the host environment or media, and recruit these proteins onto the bacterial surface. Fn attached on the cellular surface is then recognized by the principal host cell fibronectin receptor, the integrin $\alpha_5\beta_1$. This type of fibronectin bridge via FnBPs and integrin $\alpha_5\beta_1$ is the major route of *S. aureus* internalization. Thus, *S. aureus* strains with an inactivated SrtA protein (e.g. *SrtA*$^-$), can not properly display adhesion proteins and are therefore unable to invade mammalian cells. On the other hand, activity of SrtA is related to the efficiency of anchoring adhesion proteins and ultimately the total number of *S. aureus* cells that infect the host mammalian cells. However, it was labor and time intensive to accurately count the number of *S. aureus* bacteria isolated from infected hot cells at least in our hands (data not shown). To overcome this obstacle, a convenient and rapid approach to quantify the internalization of bacteria by non-professional phagocytes was adapted from a previously developed method (Illustrated in Figure 4.7A). The epithelial cells were infected with fluorescein labeled bacteria and the suspension of infected cells was applied to flow cytometry to detect cell-associated fluorescence. To quench the fluorescence signal from the extracellular bacteria attached on cell surface, trypan blue was added prior to FACS analysis. In this method, not only the percentage of infected cells can be determined, the strength of the fluorescence signal also provides the estimation of the relative amount of internalized bacteria in each sample (*193*). The uptake index (u.i.) is thus introduced to combine both read-outs and provides the correlated total amount of internalized bacteria by multiplying the percentage of fluorescent cells by the mean fluorescence intensity of the fluorescent population *(194)*.

In our assay, the u.i. was used to compare the activity of SrtA in different *S. aureus* strains. To validate such an invasion assay, *S. aureus* strain RN4220, which contains endogenous SrtA and SrtA knockout strain *srtA*$^-$ were labeled with carboxyfluorescein diacetate succinimidyl ester (CFSE) and used to infect the host HEK293H cells. CFSE is the chemical commonly used for *in vivo* cell tracing, which is originally colorless and nonfluorescent. It can passively diffuse into cells and be activated by cleavage of the acetate groups by intracellular esterases. The activated highly fluorescent carboxyfluorescein succinimidyl ester conjugates with intracellular amines and is well retained inside the cell. We repeated the invasion assay independently five times and reported the average results and standard deviation as shown in Figure 4.7. Internalized RN4220 cells were found in a large proportion, about 40%, of HEK293H cells, whereas intracellular *srtA*$^-$ cells were only detected in about 8% of HEK293H cells (Figure 4.7B). The calculated uptake index of RN4220 is 14.17, which is more than 10 fold higher than the 1.27 calculated for *srtA*$^-$ (Figure 4.7C). The significant difference of infections between RN4220 and *srtA*$^-$ clearly demonstrated the availability of SrtA on the bacterial surface. This data was consistent with the previous finding that SrtA plays a critical role in bacterial infection and is a potential drug target (*29*). It also approved that the uptake index (u.i.) is a reliable measurement of the bacterial internalization and an applicable indicator for SrtA *in vivo* activities for anchoring virulence factors.

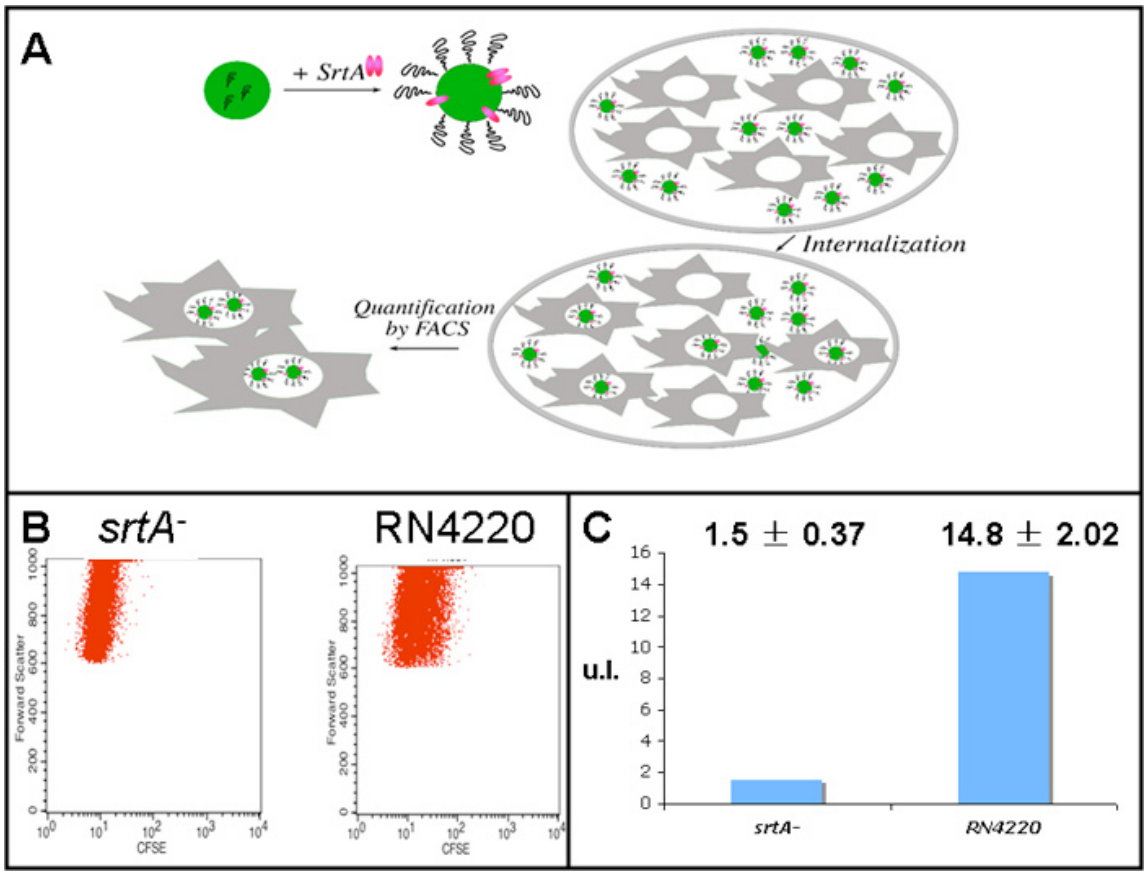

Figure 4.7: Quantification of *S. aureus* invasion in HEK293H by flow cytometry.

(A) Cartoon representation of invasion assays. SrtA expressed *in vivo* is responsible to display cell surface proteins on *S. aureus* cell wall to enable these cells to infect HEK293H. The activity of SrtA is related to the amount of internalized bacteria. Since *S. aureus* cells were labeled with CFSE prior to infection, the infected HEK293H cells can be detected by FACS based on internalized fluorescence signal. (B) HEK293H cells were infected for 60 min with CFSE-labeled *S. aureus* RN4220 and *srtA*$^-$ respectively, suspended in fixing solution with 0.2% trypan blue and analyzed by flow cytometry. The original dot plots of a representative experiment were shown. (C) HEK293H cells were infected and analyzed as in (B). The uptake index (u.i.) was calculated by multiplying the percentage of fluorescent cells by the mean fluorescence intensity of the fluorescent population. The bars represent the mean values of the u.i. for RN4220 and *srtA*$^-$ as observed in three independent experiments.

Using the same assay, the uptake indexes (u.i.) of *S. aureus* strains harboring dimeric SrtA and monomeric SrtA were also measured (Figure 4.8). To keep the genetic consistency among the strains throughout the assay, SrtA knock out strain *srtA*$^-$ was used as the host strain and served as the baseline for the uptake index measurement. As described previously, a plasmid encoding the full length SrtA with a C-terminal His6 and FLAG tag ($SrtA_{WT}$-His6-FLAG) was introduced into *srtA*$^-$ to afford *srtA*$^-$*+A* strain that majority of SrtA existed in a dimer-monomer equilibrium on the cell membrane. A plasmid (pT-P-$srtA_{NKY}$-His6-FLAG) encoding the full length SrtA with triple mutation and a C-terminal His6 and FLAG tag ($SrtA_{NKY}$-His6-FLAG) was introduced into *srtA*$^-$ to afford *srtA*$^-$*+*A_{NKY} strain that harbored monomeric SrtA on the membrane. The expression of these two SrtA proteins $SrtA_{WT}$-His6-FLAG and $SrtA_{NKY}$-His6-FLAG in *srtA*$^-$ was confirmed with western blot experiment probed with anti-FLAG antibody and the same expression level was observed for these two proteins (Figure 4.8C). EDC-mediated cross-linking confirmed that $SrtA_{WT}$-His6-FLAG existed in a dimer-monomer equilibrium and $SrtA_{NKY}$-His6-FLAG as monomer under the experimental condition. After the incubation with host mammalian cells, *srtA*$^-$*+A* cells were detected in about 20% of the infected mammalian cells (Figure 4.8A), with a calculated u.i. of around 5.2 (Figure 4.8B). *srtA*$^-$ *+*A_{NKY} cells were detected in about 45% of the infected cells with a u.i. of 14.6 (Figure 4.8A and B). The results clearly demonstrated that *srtA*$^-$ strain harboring monomeric SrtA on its membrane has significantly higher invasion activity than the same strain harboring wild type SrtA, which tends to form dimer. Since both $SrtA_{WT}$-His6-FLAG and $SrtA_{NKY}$-His6-FLAG had the similar expression level in *S. aureus* (data not shown), the significant difference of invasion activities between these two mutant strains resulted from the oligomerization status of SrtA *in vivo*. In the case of *srtA*$^-$*+A* strain, SrtA existed as a monomeric enzyme on its membrane and had the highest *in vivo* activity (u.i. of 14.6); on

the other hand, majority of SrtA existed as dimer on the membrane of *srtA*$^-$+*A* strain and the dimerization of SrtA reduced the enzyme's *in vivo* activity (u.i. of 5.2). Compared to of *srtA*$^-$ strain (u.i of 1.5), the rest small amount of SrtA monomers on *srtA*$^-$+*A* membrane contributed to the relatively low invasion activity for the strain. It is interesting to notice that the overall invasion activity of *S. aureus* was proportional to the amount of SrtA monomers *in vivo*.

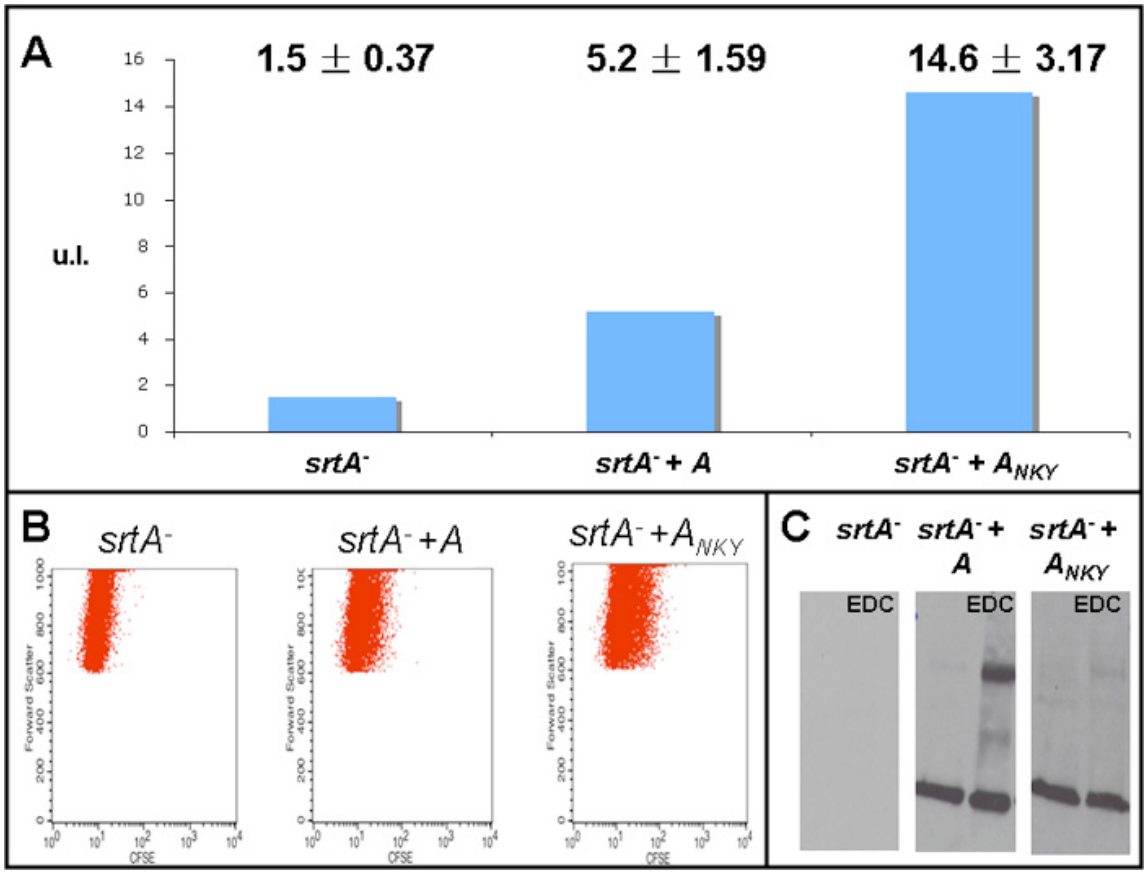

Figure 4.8: Comparison of *S. aureus* invasion in HEK293H between SrtA wildtype and monomer only mutant expressing strains.

(A) HEK293H cells were infected for 60 min with CFSE-labeled *S. aureus srtA*$^-$, *srtA*$^-$ + srtA, and *srtA*$^-$ + srtA$_{NKY}$ respectively, suspended in fixing solution with 0.2% trypan blue and analyzed by flow cytometry. The original dot plots of a representative experiment were shown. (B) HEK293H cells were infected and analyzed as in (A). The uptake index (u.i.) was calculated by multiplying the percentage of fluorescent cells by the mean

fluorescence intensity of the fluorescent population. The bars represent the mean values of the u.i. for *srtA*$^-$, *srtA*$^-$ + srtA, and *srtA*$^-$ + $srtA_{NKY}$ respectively as observed in three independent experiments. (C) Cell lysate of *srtA*$^-$, *srtA*$^-$ + srtA, and *srtA*$^-$ + $srtA_{NKY}$ were analyzed by SDS-PAGE and western blot to quantify the expression of SrtA *in vivo*. EDC was added to stabilize the dimerization between SrtA molecules.

This result demonstrated that *srtA*$^-$ strain harboring monomeric SrtA on its membrane has significantly higher invasion efficiency than the strain harboring wild type SrtA, which has majority of molecules in dimer form. In both strains of *srtA*$^-$ + *A* and *srtA*$^-$ + A_{NKY}, the expression of the reintroduced srtA genes was regulated by the same endogenous srtA promoter, and the sortase A concentration on the cellular membrane in *srtA*$^-$ + *A* and *srtA*$^-$ + A_{NKY} were about the same when detected with western blotting. In addition, wild type $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59,\ NKY}$ had the same *in vitro* enzymatic activity when both enzymes were analyzed at low concentration in monomer form. Thus, the significant difference of infectious activities between these two strains should result from the different oligomerization state of sortase A *in vivo*. In the case of *srtA*$^-$ + A_{NKY} strain, SrtA existed as a monomeric enzyme on its membrane and had the higher *in vivo* activity (u.i. of 14.6); on the other hand, majority of SrtA molecules existed as dimer on the membrane of *srtA*$^-$ + A strain and the dimerization of SrtA reduced the enzyme's *in vivo* activity from a u.i. 14.6 to a u.i. 5.2. Compared to *srtA*$^-$ strain (u.i of 1.5), the rest small amount of SrtA monomers on *srtA*$^-$ + A membrane contributed to the increased invasion efficiency of the strain from a u.i. 1.5 to a u.i. 5.2. The higher invasion efficiency of *srtA*$^-$ + A_{NKY} is attributed to the higher activity of monomeric sortase A and lower invasion efficiency of *srtA*$^-$ + *A* results from the lower activity of dimeric sortase A *in vivo*. Dimerization of SrtA greatly suppresses its enzymatic activity *in vivo*.

Although it is not a strictly fair comparison, it is interesting to observe that the expression of full-length sortase A in *srtA*$^-$ + *A* strain was unable to fully rescue its invasion efficiency as wild type RN4220. One possibility is that certain components

involving *S. aureus* invasion pathway were disrupted by such genetic modification on *S. aureus* genome. Such a hypothetical component might not directly regulate the function of SrtA, but could mediate various pathways in protein secretion, folding, and the cell wall biosynthesis in *S. aureus*. The other possibility lies in the fact that exogenously expressed SrtA *in vivo* contains C-terminal epitope, which might interfere with the function of SrtA on both wild type and mutant proteins.

The discovery that wild type SrtA, which can form dimer is less active than monomer only mutant would contribute to the current controversy in SrtA catalytic model. As discussed before in Chapter 1, current mechanism can only be reasonable when two active site residues Cys^{184} and His^{120} are in charged form. However, based on protein folding structures from crystal or NMR, under physiological pH, the chance of that both residues are charged is as low as 0.06% (*47, 63*). Therefore, 99.94% of the molecules are in their inactive form, which is biologically very inefficient. We hypothesized that the monomer form of SrtA *in vivo* can fold into a specific conformation so that the microenvironment at active site favors the protonation of His^{120} and deprotonation of Cys^{184}. However, this is just a speculation that has to be evaluated with a high-resolution structure of full-length SrtA monomer only mutant in its native condition.

CONCLUSION

Previously SrtA has been shown to exist as a dimer *in vitro* and in *E. coli (100)*. In this study, we have demonstrated that SrtA exists as a dimer on the membrane of *S. aureus* and the dimerization *in vivo* has a much higher affinity than that *in vitro*. An invasion assay was set up to compare the *in vivo* activities of SrtA dimer and monomer. The results of this assay demonstrated that SrtA mutant protein that only forms a monomer confers higher invasion efficiency for the strain, which suggested that dimerization of SrtA possibly suppresses the enzyme's activity *in vivo*. Current results

suggested that mechanism of SrtA function *in vivo* is far more complicated than the previous chemical mechanism (*47, 49, 51, 63-65, 86, 195-197*) and that its *in vivo* functions are regulated directly or indirectly by its self-association. Next, genomics and proteomics studies will be carried out to identify such regulator and/or regulatory pathway, which may be a new drug target for gram-positive bacteria infections.

METHODS AND MATERIALS

Construction of *srtA* knockout *S. aureus* strain

A SrtA knockout strain (*srtA*$^-$) served as the host for *in vivo* assays experiments and was generated by using the TargeTron™ Gene Knockout System (Sigma-Aldrich, St. Louis, MO). TargeTron™ Gene Knockout System uses site-specific group II intron insertions to permanently disrupt genes on the bacterial genomic DNA (*179*). A web-based algorithm was used to locate an optimal intron insertion site between the 93rd and 94th nucleotide in SrtA. This gave a score of 7.55 and E-value as 0.152. The primers designated as IBS (5'-AAAAAAGCTTATAATTATCCTTACATATCGATAATGTGC GCCCAGATAGGGTG-3'), EBS1d (5'-CAGATTGTACAAATGTGGTGATAACAGA TAAGTCGATAATTATAACTTACCTTTCTTTGT-3'), EBS2 (5'-TGAACGCAAGTT TCTAATTTCGGTTATATGTCGATAGAGGAAAGTGTCT-3') and EBS universal were used to perform the PCR reaction and amplify a 350 bp fragment from the template pNL9164. The amplified 350 bp fragment, which targets the intron to the SrtA sequence, was digested with Hind III and BsrG I (New England Biolabs, Beverly, MA) and inserted into pNL9164 to make pNL9164-srtA. The plasmid pNL9164-srtA was then transformed into *Staphylococcus aureus* strain RN4220 (general gift from Dr. Alan Lambowitz). These cells were grown in trypsin soy broth (TSB) at 32°C for one hour and diluted at a ratio of 1:5 into fresh media containing 10 μg/mL erythromycin (Sigma, St. Louis, MO).

Afterwards, the over-night cultures were diluted 1:100 into fresh media containing erythromycin and grown until an OD_{595} of 0.5 was reached. 10 μM cadmium chloride hemi (pentahydrate) (Sigma, St. Louis, MO) was added to induce intron production for 90 minutes. The cells were then selected on trypsin soy agar (TSA) plates containing erythromycin. Colony PCR was performed using EBS2 as the forward primer and a SrtA gene sequence downstream of the insertion site, srtA-R (5'-TTTGACTTCTGTAGCTACAAAGATT -3'), as the reverse primer. The colonies with a disrupted gene due to intron insertion had PCR products of 1.3 kb. The temperature sensitive plasmid was then eliminated from the cell by incubating the confirmed gene-disrupted mutant strain at 43°C in antibiotic-free media. As a result the generated strain contained a permanently disrupted SrtA gene. Southern blot was also performed to ensure that the intron insertion was unique to the SrtA gene. Restriction digestion enzymes Spe I, Nsi I, Nde I and Sph I (New England Biolabs, Beverly, MA) were used to fragment genomic DNA. Next, the sample was resolved by electrophoresis on a 0.8% agarose gel, and transferred onto a Hybond-N^+ nylon membrane (Amersham Bioscience, Piscataway, NJ) by capillary transfer. DIG high prime DNA labeling and detection starter kit (Roche, Indianapolis, IN) was then used to label the probe DNA and detect the hybridization. DIG labeled intron sequence was PCR amplified using DIG-dUTP. Only one nucleic acids band of 2.5 kb was detected by DIG antibody, which confirms that the intron sequence was only inserted into the SrtA gene.

Construction of SrtA wild type and mutant protein expression vectors

Primers designated as NcoI-SrtA-F (5'-CATGCCATGGGTCTATGGTTTGTGG TTCATAAG-3') and XhoI-SrtA-R (5'-CCGCTCGAGTTTGACTTCTGTAGCTACAA AGATT-3') were used to amplify the srtA sequence including its endogenous promoter from genomic DNA of *Staphylococcus aureus* stain RN4220 by the polymerase chain

reaction. The DNA fragments were then digested with Nco I and Xho I (New England Biolabs, Beverly, MA) and cloned into the pET28b expression vector (Novagen, La Jolla, CA) to generate the construct pET28-P-SrtA. Primers designated SphI-SrtA-F (5'-ACATGCATGCGTCTATGGTTTGTGGTTCATAAG-3') and KpnI-P2T-R (5'-CGAGGGTACCTTACTTGCGTCATCGTCTTTGTAGTCGTGGTGGTGGTGGTGGT G-3') were used to copy SrtA sequence along with its N-terminus promoter and C-teriminus His tag. The DNA fragment was digested with Sph I and Kpn I (New England Biolabs, Beverly, MA) and cloned into the pNL9164 shuttle vector (Sigma-Aldrich, St. Louis, MO)(*198, 199*) to generate the construct pT-SrtA-H-FLAG.

The following six primers designated as N132AF (5'-CCGAACTATCAATTTAC AGCTCTTAAAGCAGCCAAAAAA-3'), N132AR (5'-TTTTTTGGCTGCTTTAAGAG CTGTAAATTGATAGTTCGG-3'), K137AF (5'-CAAATCTTAAAGCAGCCGCAAA AGGTAGTATGGTGTAC-3'), K137AR (5'-GTACACCATACTACCTTTTGCGGCTG CTTTAAGATTTG-3'), Y143AF (5'-GCCAAAAAAGGTAGTATGGTGGCCTTTAAA GTTGGTAATG-3') and Y143AR (5'-CATTACCAACTTTAATGGCCACCATACTAC CTTTTTTGGC-3') were obtained from Invitrogen to generate the mutations N132A, K137A and Y143A respectively. Site-directed mutagenesis was performed as described in the Stratagene protocol (*170*) using pT-srtA-H-FLAG as the template plasmid. Triple mutation mutant $SrtA_{NKY}$ was generated by three-step site-directed mutagenesis.

The $SrtA_{\Delta N59}$ expression vector was generated as previously described (*100*). $SrtA_{\Delta N59,\ NKY}$ was then generated using the same protocol and primers by using $SrtA_{NKY}$ expression vector as the template.

Expression and purification of SrtA proteins

The pT-srtA-H-FLAG (wild type or mutant) construct was transformed into *S. aureus* SrtA knockout strain *srtA*$^-$. The pT-SrtA-H-FLAG transformed cells were grown

in Tryptic Soy Broth media supplemented with 10 μg/mL erythromycin (Sigma, St. Louis, MO) at 32 ºC over-night. The cells were harvested and treated with 1 μg/mL lysostaphin (Sigma, St. Louis, MO) at room temperature for 1 hour before lysed under native conditions using sonication. SrtA proteins were then purified from cell lysate using the QIAexpression kit according to the manufacturer's protocol (Qiagen, Valencia, CA) (*200*). The concentration of protein was determined using Bradford reagent (Pierce Biotechnology, Rockford, IL). *In vitro* expression and purification of $SrtA_{\Delta N59}$ and $SrtA_{\Delta N59, NKY}$ was the same as previously described (*100*).

Cell fractionations

Cell fractionation was done according to published protocol (*201*). 5 mL staphylococcal over-night culture were recovered and the cells were digested with 5 μg/mL lysostaphin for 60 min at 37°C in Tris•HCl buffer (20 mM Tris, pH 7.1). The protoplasts were collected by centrifugation for 10 min at 13, 000 ×g., and the supernatant was removed. The protoplasts were lysed in 250 μL Tris•HCl EDTA buffer (20 mM Tris, 0.372% EDTA, pH 7.1) with a brief sonication. The cell pellet was removed by centrifugation at 8000 ×g for 5 min. The supernatant was transferred into a clean tube and centrifuged at 21, 000 ×g for 60 min. The supernatant was then saved as the cytosolic fraction. The pellet was washed with Tris•HCl buffer and centrifuged again at 21, 000 ×g for 60 min. The pellet was then resuspended in Tris•HCl buffer and saved as the membrane fraction. All experiments were performed at 4°C.

Polyacrylamide gel electrophoresis and western blot analysis

Proteins were analyzed on a 14% SDS polyacrylamide gel under denaturing conditions, and visualized by Coomassie Blue staining. For western blot analysis, proteins were first transferred onto a nitrocellulose membrane (Hybond-C Extra,

Amersham Bioscience, Piscataway, NJ) in a Tris-glycine buffered electrophoresis tank. The membranes were then probed with an anti-FLAG primary antibody (Qiagen, Valencia, CA) or an anti-sortase A primary antibody (ab13959, Abcam Inc., MA) and a horseradish peroxidase-conjugated goat anti-mouse or goat anti-rabbit secondary antibody (Bio-Rad, Hercules, CA). The Western Blot Detection Reagent (GE healthcare, UK) was applied to visualize the signals, which were detected by exposing the membrane to BioMax Light Film (Eastman Kodak Company, Rochester, NJ).

Chemical cross-linking in *S. aureus*

S. aureus cells were cultured as described above. 1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (EDC, Sigma, St. Louis, MO) was added to the media to a final concentration of 0.7 mg/mL. Cells were harvested 60 minutes later and analyzed as described above by SDS-PAGE and western blot.

Protein in-gel digestion and peptide fingerprint mapping by MS

The proposed dimer and monomer bands of SrtA were sliced from an SDS-polyacrylamide gel and subjected to in-gel trypsin digestion. Proteins were then extracted from the gel and analyzed by MALDI-TOF mass spectrometry and peptide fingerprint mapping analyses in the Analytical Instrumentation Core Facility at the University of Texas at Austin.

Size-exclusion chromatography

Purified SrtA protein was applied onto a Superdex 200 column (300 × 10 mm) that was first pre-equilibrated with phosphate-buffered saline (PBS, 138 mM NaCl, 2.7 mM KCl, pH 7.4). The sample was eluted with the same buffer at 0.75 ml/min at 4°C using a BioLogic DuoFlow system (Bio-Rad, Hercules, CA). Eluted protein was detected by monitoring at an absorbance of 280 nm. The same amount of purified, EDC stabilized,

SrtA dimer was analyzed by gel filtration column under the same running conditions. Eluted fraction peaks were collected and analyzed by SDS-PAGE and western blot using anti-FLAG antibody.

Circular dichroism spectrum

All the measurements were repeated three times, and each data point was recorded in triplet. Circular dichroism (CD) studies were conducted on a Jasco J-810 spectropolarimeter (Jasco International Co., Ltd. Tokyo, Japan), equipped with a Jasco PFD-425S temperature control system as described earlier(*101*). Briefly, $SrtA_{\Delta N59}$ wild type and $SrtA_{\Delta N59, NKY}$ were dialyzed into activity assay buffer (50 mM Tris-Cl, 150 mM NaCl, 5 mM $CaCl_2$, pH 7.5) and diluted to a concentration of 39 μM. The recorded spectra were analyzed using the Dichroweb server(*165*) to get an estimation of the composition of the secondary structures. The NRMSD parameter was used to evaluate the fitness of various methods (*175*). An NRMSD value less of than 0.1 indicates that the predicted spectra and the experimental spectra are in close agreement (*176*).

Epithelial cell culture

The human embryonic kidney cell line 293H (HEK293H; Invitrogen, Carlsbad, California) was grown in Dulbecco's Modified Eagle's Medium (DMEM; Sigma, St. Louis, MO)/10% Fetal Bovine Serum (FBS; Atlas Biologicals, Fort Collins, CO) at 37 °C, 5% CO2 prior to plating. One day prior to invasion assays, 1.5×10^6 cells were seeded in 6-well plates (Corning Incorporated, Corning, NY) yielding weakly confluent monolayers (3×10^6 cells/well) at the time of invasion.

Labeling of *S. aureus* cells

S. aureus (50 mL with OD_{600} 0.6) were washed three times with cold, sterile Dulbecco's Phosphate Buffered Saline (PBS; Sigma, St. Louis, MO) to reduce

aggregation, suspended in 10 mL 0.5 μM 5-(6)-carboxyfluorescein-succinimidylester (CFSE; Molecular Probes, Eugene, OR) in PBS, and incubated for 60 min in 37 °C shaker. CFSE-labelled bacteria were washed three times with PBS and protected from direct light exposure.

Invasion assays and flow cytometry

For invasion assays, 3×10^6 HEK293H cells in 6 cm^2 culture dishes were infected at a MOI of 20 CFSE-labelled bacteria/cell. The infection mixture was incubated in DMEM/10% FBS at 37°C for 60 min. Invasion was terminated by incubation with 300 μg/mL gentamicin (Sigma, St. Louis, MO) in DMEM/10% FBS for 30 min at 37°C. The infected cells were then washed three times with PBS and suspended by 1 min trypsin-EDTA solution treatment (Sigma, St. Louis, MO). Suspended cells were then resuspended in 1 ml fixing solution (paraformaldehyde solution 4% in PBS; USB Corporation, Cleveland, OH). Samples were kept on ice and dark until analysis. To eliminate signals from extracellular *S. aureus*, trypan blue solution (0.4%; Sigma, Taufkirchen, Germany) was added to a final concentration of 0.2% right before analysis. Samples were analyzed on a FACSCalibur Flow Cytometer (Becton Dickinson, Franklin Lakes, NJ) by gating 10, 000 HEK293H cells per sample based on forward and side scatter. Cell-associated fluorescence was measured in fluorescence channel 1 (FL1-H) to detect CFSE fluorescence. An estimate of the amount of internalized bacteria was calculated by multiply the percentage of CFSE-positive cells with the mean fluorescence intensity of these cells to obtain the uptake index (u.i.).

Chapter 5 Oligomerization of endogenous sortase family proteins in *Staphylococcus aureus*

INTRODUCTION

In gram-positive bacteria, most of the genomes can encode more than one sortase gene to anchor different classes of surface proteins based on their specific sorting signals (*41*). Those surface proteins anchored by different sortases have distinct biologic functions in a variety of processes such as virulence factors for host tissue invasion, sporulation, and iron acquisition (*35, 41*). In *S. aureus*, sortase B (SrtB) was identified in 2002, which shares about 45% primary sequence homology with SrtA (*202*). It is a membrane-bound cysteine transpeptidase, recognizing NPQTN sequence instead of LPXTG for SrtA (*202*). The gene of SrtB is part of an iron-regulated locus called iron-responsive surface determinants (isd), which contains the only SrtB substrate in *S. aureus*, IsdC. SrtB is required for iron acquisition from the living environment that is critical for the establishment of human infections (*90, 202, 203*). The crystal structure showed that the catalytic domain of SrtB folds into a very similar β-barrel structure as SrtA, in which some β strands are connected by short α-helices and loops (*195*). Similar to SrtA, the active site residues in SrtB are also located in a hydrophobic groove formed by two ß strands (ß7/ß8). A cysteine-histidine-asparagine triad (Cys^{223}-His^{130}-Arg^{225}), an equivalent of SrtA active sites (Cys^{184}-His^{120}-Arg^{197}) were proposed for SrtB catalytic mechanism and supported by various biochemical data (*195, 197*). One of the major differences between the structures of these two sortase family proteins is that there are two short helices at the N-terminus of SrtB catalytic domain, connecting the transmembrane fragment and the extracellular region, which is replaced with a long random loop structure in SrtA. It has been speculated that these two short helices may

provide the correct orientation of the active sites of SrtB, so that they face towards the bacterial extracellular surface and anchor the substrate proteins on maturely assembled peptidoglycan. For SrtA, the long loop structure makes the position of catalytic domain flexible and possibly facing the plasma membrane (*195*). The investigation of the molecular mechanism of SrtA and SrtB would help to explain how are the two enzymes anchoring their substrates at distinct locations on cell wall.

Various genetic and biochemical data from our lab as described in Chapter 2, 3, and 4 has demonstrated that the catalytic domain of SrtA selectively formed a homo-dimer *in vitro* and the dissociation constant of SrtA homo-dimer is 55 μM *(100, 101)*. Full length SrtA expressed in *S. aureus* formed a dimer with a higher affinity, which can only be disrupted by three key amino acids mutated simultaneously. The *S. aurues* strain with monomer only SrtA expressed *in vivo* was more likely to be internalized into mammalian cells compared to the strains with wide type SrtA. Although the exact mechanism involved in this significant difference is not truly revealed yet, it is obvious that the function of SrtA *in vivo* is closely related to the cell membrane associated oligomerization status. It will be interesting to know whether the dimer-monomer equilibrium is shared for other sortase family proteins, particularly SrtB. Therefore, a similar investigation about SrtB oligomerization status was performed. The data suggested that *in vitro* expressed SrtB catalytic domain $SrtB_{\Delta N30}$ and *in vivo* expressed full length SrtB do not form dimer, unlike $SrtA_{\Delta N59}$ and full length SrtA do. Domain swap experiments by generating mosaics of sortase transmembrane domain, linker region and catalytic domain from SrtA and SrtB, were performed. The tendency of heterogeneous sortases to form dimer *in vivo* was analyzed individually, in an attempt to identify the sequence critical for specific SrtA dimerization *in vivo*.

RESULTS AND DISCUSSION

Catalytic domain of SrtB expressed in *E. coli* is monomeric.

In order to obtain *S. aureus* SrtB protein for *in vitro* study, the truncated $SrtB_{\Delta N30}$ that retains the same transpeptidation activity as the full-length SrtB enzyme (*54, 202*) and is easy to purify was over-expressed and purified from *E. coli*. After purification under the native conditions, $SrtB_{\Delta N30}$ was analyzed by native polyacrylamide gel electrophoresis. When stained with Coomassie Blue, a single protein band was observed (Figure 1B). Western blot with an anti-Tetra-His antibody confirmed that this protein band contains $SrtB_{\Delta N30}$ (Data not shown). The protein was also analyzed by size-exclusion chromatography to determine the molecular weight of $SrtB_{\Delta N30}$ under native conditions. A distinctive peak was observed after the purified $SrtB_{\Delta N30}$ was applied to a Superdex 200 gel filtration column (Figure 1A). The calculated molecular weight of the peak is 24229 Da, which corresponds to the estimated size of 26780 Da for a monomer of $SrtB_{\Delta N30}$.

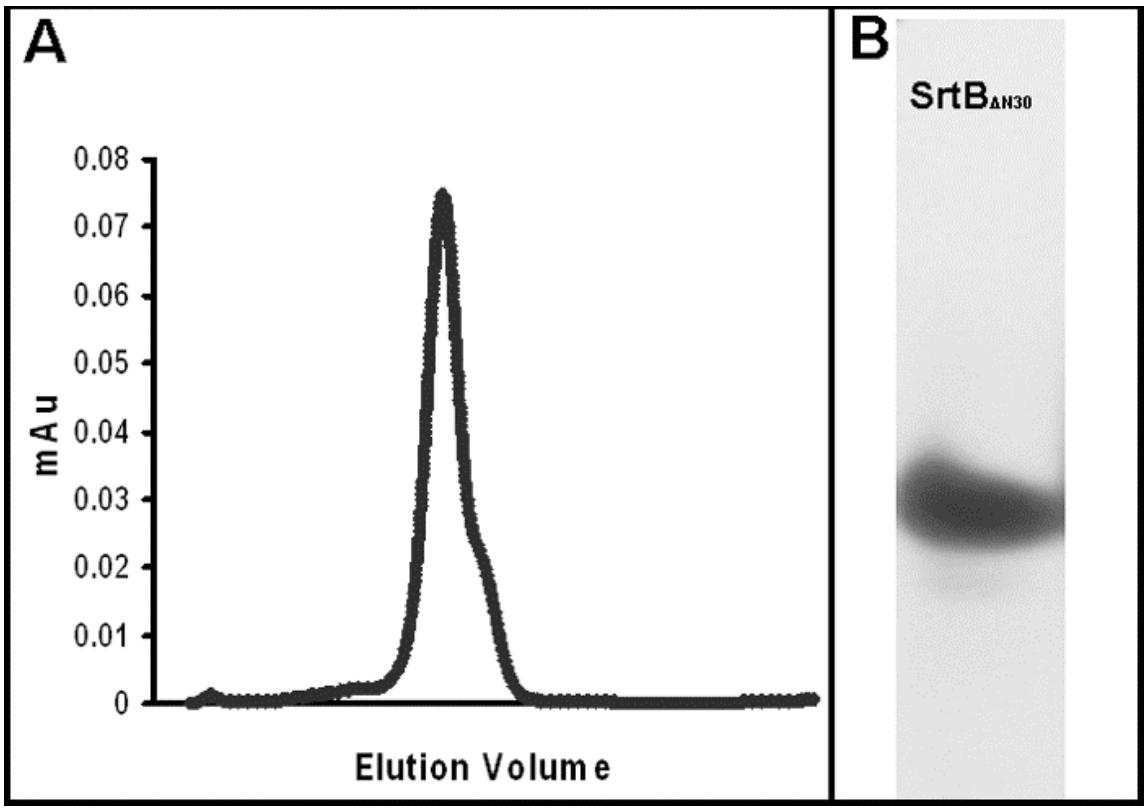

Figure 5.1: *In vitro* expressed $SrtB_{\Delta N30}$ exists as monomer.

$SrtB_{\Delta N30}$ was over-expressed in *E. coli* and purified for *in vitro* characterization. Size exclusion chromatography (A) and native PAGE analysis (B) indicated that SrtB protein is a monomeric protein *in vitro*. (A), 1.9 mg of purified protein was applied to a size exclusion gel filtration column. Estimation of molecular weight indicated a single peak of 24229 Da that corresponded to the size of $SrtB_{\Delta N30}$ monomer. (B), 53 μg of purified SrtB protein was resolved by PAGE in a native condition and then visualized by Coomassie Blue staining. A single monomer band was observed, which corresponded to the single peak eluted in size exclusion gel column.

Full-length membrane protein SrtB expressed *in vivo* is monomeric.

To study the oligomerization state of endogenous SrtB expressed in *S. aureus*, SrtB expressing strain, RN4220+B was generated using strain RN4220 as a host. The endogenous SrtB is not expressed constitutively but only under conditions when iron is limiting (*202*). Therefore, an over-expressing plasmid, pT-srtB-His-FLAG was created to

harbor the full-length srtB gene along with a cadmium-inducible promoter fused to its 5' prime end and a sequence encoding a His6/FLAG tag on its 3' prime end. The plasmid pT-srtB-FLAG was then transformed into RN4220 strain to over-express His6/FLAG tagged SrtB (SrtB-His6-FLAG) upon the addition of cadmium chloride. The expression of SrtB-His6-FLAG was probed with anti-FLAG antibody, after resolving cell lysate on SDS-PAGE followed by western blot, as shown in Figure 5.2, lane 1. We observed a single band of SrtB monomer at the size of around 30 kDa.

A chemical cross-linking experiment was performed to demonstrate that SrtB exists as a monomer *in vivo*. The same experiment was carried out before to demonstrate that full length SrtA protein expressed in *S. aureus* selectively formed homo-dimer. The chemical cross-linking reagent, 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), was used to stabilize the possible transient protein-protein interactions *in vivo* (*120*). We treated the cell culture of RN4220+B briefly with EDC before harvesting. Cell lysate was then resolved on SDS-PAGE and SrtB-His6-FLAG protein was visualized by western blot using anti-FLAG antibody (Figure 5.2, lane 2). In addition to the 30 kDa monomer band, no other band corresponding to the size of SrtB dimer was detected. The *in vivo* cross-linking experiment supported the notion that the full length SrtB has no tendency to form dimer in *S. aureus*.

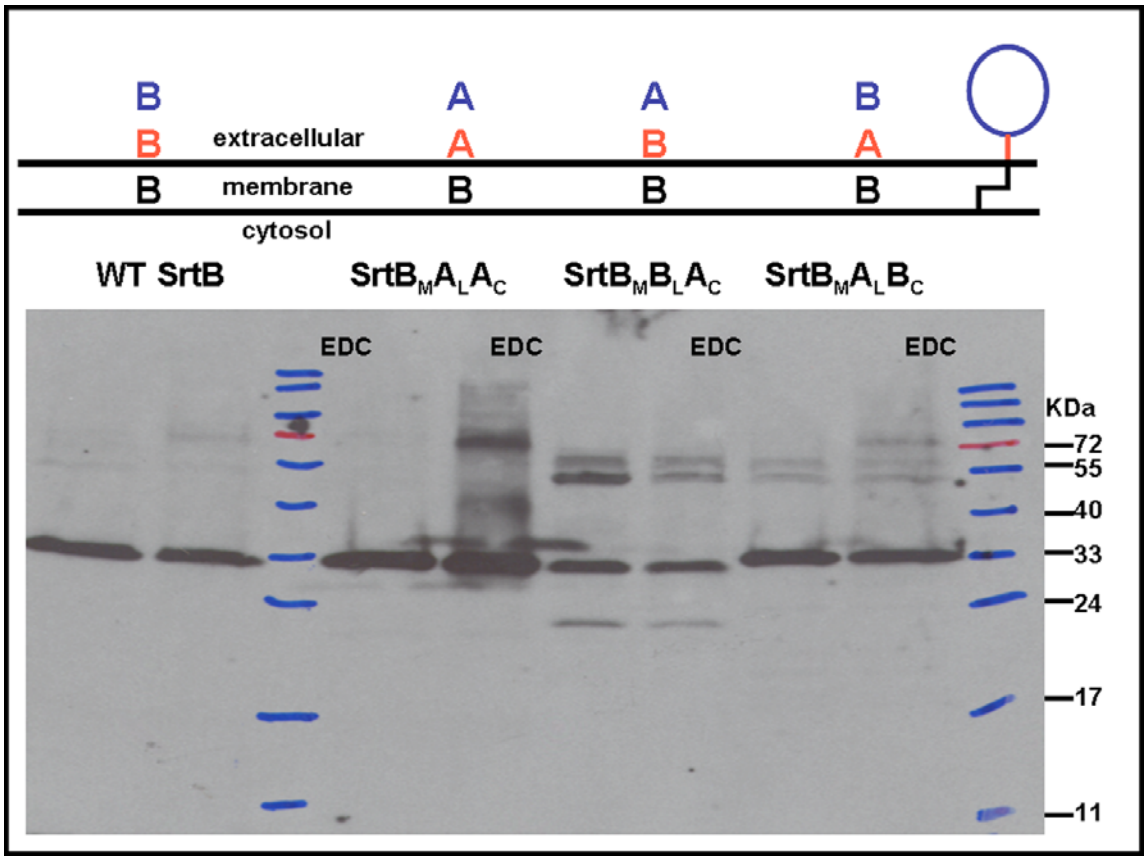

Figure 5.2: SDS-PAGE with full length SrtB and domain-swapped fusion proteins, $SrtB_MA_LA_C$, $SrtB_MB_LA_C$, and $SrtB_MA_LB_C$, expressed from *S. aureus*.

Full length SrtB and fusion proteins were fused with a His6/FLAG tag on C-terminus and over-expressed in *S. aureus*. Cell lysates were analyzed by SDS-PAGE and proteins were visualized by western blot with anti-FLAG antibody. Cells were treated with EDC cross-linking reagent (labeled with EDC) or not (no label) before harvesting. The cartoon on top of the gel illustrated the alignment of protein sequence in each sample. Each protein is composed of three fragments, transmembrane domain, linker region and catalytic domain.

Fusion proteins with domain-swap were generated for identification of structural determining factors for dimerization.

Fusion proteins were generated by domain-swapping between dimeric SrtA and monomeric SrtB proteins to seek for the structural determining regions on the sequence that contribute to and/or favor for dimerization. We first segmented both proteins in three

fragments respectively, transmembrane domain, catalytic domain and linker region connecting those two domains. The transmembrane domain of SrtA (A_M, amino acids 8-25) is predicted to form membrane-anchoring α-helix, which is equivalent to the transmembrane domain of SrtB (B_M, amino acids 6-25). The catalytic domains of SrtA (A_C, amino acids 74-206) and SrtB (B_C, amino acids 70-244) both fold into very similar sortase's unique 8-stranded β-barrel structure. The linker region is the fragment with the least similarity between SrtA and SrtB. SrtA linker region (A_L, amino acids 26-73) is predicted to be a long random loop and confirmed by X-ray chromatography and NMR structure (*50, 51*). SrtB linker region (B_L, amino acids 26-69) was shown in X-ray chromatography to form two short α-helices linked by a loop (*54*). The fusion proteins were mosaics of the three segments from different sortases and their *in vivo* expression was under the control of SrtA's endogenous promoter.

We first investigated the hypothesis that whether the transmembrane domains of sortase proteins determine the dimerization *in vivo*. SrtA transmembrane domain was fused with SrtB linker region and catalytic domain to form fusion protein $SrtA_MB_LB_C$. SrtB transmembrane domain was fused with SrtA linker and catalytic domain to form $SrtB_MA_LA_C$. Both proteins were expressed in *S. aureus* and treated with EDC cross-linking reagent. Western blot with anti-FLAG antibody was performed to visualize the FLAG-tagged proteins. As seen in Figure 5.2 and Figure 5.3, without EDC treatment, both proteins showed single monomer bands on denaturing SDS-PAGE at around 28 kDa (Figure 5.3, lane 3 for $SrtA_MB_LB_C$, Figure 5.2, lane 3 for $SrtB_MA_LA_C$). After EDC treatment, $SrtA_MB_LB_C$ protein had only monomer band (Figure 5.3, lane 4), while the $SrtB_MA_LA_C$ protein had an extra band of about 60 kDa, which corresponded to dimer (Figure 5.2, lane 4). The ratio of the EDC cross-linked dimer to the monomer was estimated by the density, which is about 1:4. Data from this swapping experiment

suggested that the transmembrane domain of SrtA is not sufficient by itself to bring together the protein sequences fused on top of it. The transmembrane domain of SrtB is not preventing the dimeraztion of proteins fused to it either.

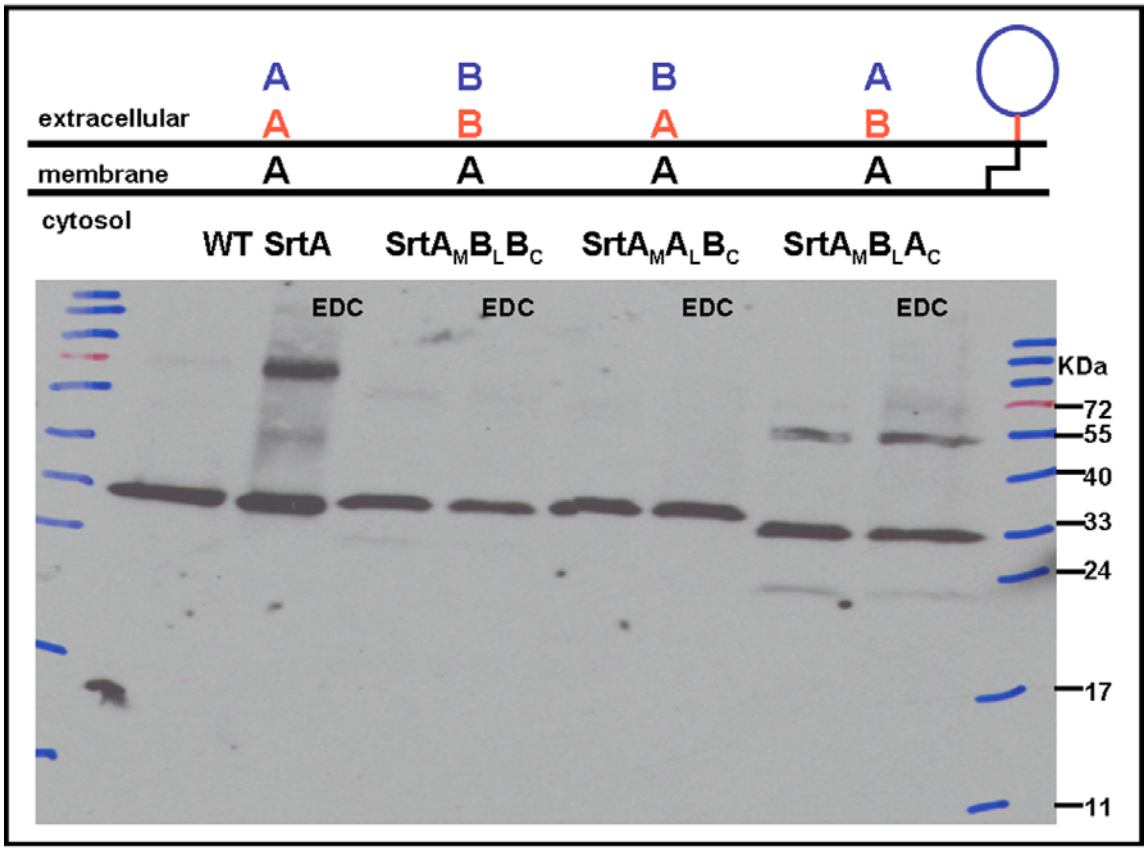

Figure 5.3: SDS-PAGE with full length SrtA and domain-swapped fusion proteins, $SrtA_MB_LB_C$, $SrtA_MA_LB_C$, and $SrtA_MB_LA_C$, expressed from *S. aureus*.

Full length SrtA and fusion proteins were fused with a His6/FLAG tag on C-terminus and over-expressed in *S. aureus*. Cell lysates were analyzed by SDS-PAGE and proteins were visualized by western blot with anti-FLAG antibody. Cells were treated with EDC cross-linking reagent (labeled with EDC) or not (no label) before harvesting. The cartoon on top of the gel illustrated the alignment of protein sequence in each sample. Each protein is composed of three fragments, transmembrane domain, linker region and catalytic domain.

We took further investigation on the linker region by generating another set of domain-swapped fusion proteins, $SrtA_MA_LB_C$ (SrtA transmembrane domain and linker

region with SrtB catalytic domain) and $SrtB_MB_LA_C$ (SrtB transmembrane domain and linker region with SrtA catalytic domain). With the same experiment condition, both proteins were expressed in monomer forms in *S. aureus* as shown as single bands on denaturing SDS-PAGE (Figure 5.3, lane 5 for $SrtA_MA_LB_C$, Figure 5.2, lane 5 for $SrtB_MB_LA_C$). With EDC treatment, neither of $SrtA_MA_LB_C$ and $SrtB_MB_LA_C$ sample had a strong dimer band indicating that no molecules were in the close proximity enough to form dimer. However, $SrtB_MB_LA_C$ sample had some bands showing at around 55 kDa, which did not match with the dimer form of the heterogeneous molecule. The identities of these bands were not confirmed yet. It is possible that the irregular folding of the fusion proteins stimulated the accumulation of certain accessory proteins, which lead to the formation of complex among them. This explanation was also supported by the fact that the expression level of this fusion protein was lower than the other proteins, coupled with significant band of degrading products. It is surprising that, although SrtA catalytic domain can form weak dimer *in vitro*, $SrtB_MB_LA_C$ did not yield a dimer band even after EDC treatment. These data, with both linker and transmembrane domains swapping between SrtA and SrtB, indicated that, the linker sequence between the catalytic and transmembrane domains of sortsase A confers the preference for dimer formation *in vivo*. The linker region of SrtB, which forms two helices somehow hinder the association between molecules anchored on top of it. The linker region of SrtA is a flexible loop that lacks secondary structure and therefore can not afford fixed topology to bring together the protein molecule fused to it.

To provide further evidence for the responsibilities assigned to individual segments of sortase proteins, we investigated the dimerization status of a third pair of fusion proteins. In this pair, only the linker region was switched between SrtA and SrtB. $SrtA_MB_LA_C$ and $SrtB_MA_LB_C$ were analyzed with the same experimental setup. $SrtA_MB_LA_C$

sample was accompanied with degradation and unknown bands that were not match with the size of dimer (Figure 5.3, lane 7 and 8). A very faint and neglectable dimer band was observed. Compared with wild type SrtA, the replacement of SrtA linker region with an ordered structure from SrtB liker region in this fusion protein substantially disrupted the majority of dimerization. The analysis of $SrtB_MA_LB_C$ sample proved the same hypothesis from an opposite perspective (Figure 5.2, lane 7 and 8). Upon the addition of EDC, aside from the unknown random bands also observed in no EDC treatment sample, a neglectable vague band migrated similar to the size of dimer was observed. The formation of this dimer band might be attributed to the variable structure of SrtA's long loop linker, which may randomly bring together the attached molecules non-specifically with low affinity.

CONCLUSION

In this chapter, we have demonstrated that SrtB, a sortase family protein that shares similar primary sequence, overall folding structure, and catalytic mechanism with SrtA, is not able to form dimer either *in vitro* or *in vivo*. It is not surprising since they are different enzymes with distinct substrates and cellular function. However, this phenomenon suggested that the dimerization of SrtA is not attributed by its transmembrane nature but specific for its *in vivo* activity and associated with function regulation, the characteristics of which is not conserved for other sortase family proteins.

Our domain-swapping experiment demonstrated the importance of SrtA linker region to significantly enhance dimerization formation of SrtA catalytic domain *in vivo*. This is an interesting discovery since the linker region of SrtA is composed of a flexible loop, which lacks secondary structure. It will be challenging for such region to be positioned together to bring two SrtA molecules together, the possibility of which has been proved not applicable based on our experiments. Instead of forming dimer directly,

we hypothesized that this region synergistically enhances dimer formation by interacting either directly or through other adaptor molecules with SrtA catalytic domain. Thus the enhancement of dimerization by SrtA linker region is only relevant to specific structure on SrtA catalytic domain. It is reasonable that the two fragments of a protein form some kind of contacts with each other to coordinate the overall structure and function. Based on this observation, we believed that a more comprehensive structure of full length, membrane protein SrtA would give us more accurate information about SrtA catalytic mechanism on molecular level.

MATERIALS AND METHODS

Construction of $SrtB_{\Delta N30}$ *in vitro* and $SrtB_{wt}$ *in vivo* Expression Vectors

Primers designated NcoI-srtB30-F (5'-CGATCCATGGGCGAAGACAAGCAA GAACGCGC-3'), NcoI-srtB-F (5'-CATGCCATGGGCAGAATGAAGCGATTTTTAA CTATTGTAC-3') and XhoI-srtB-R (5'-CCGCTCGAGTTGACTTACCTTAATTATTT TTGCGACAAC-3') were used to amplify a $SrtB_{\Delta N30}$ sequence (which express residues 31-244) and $SrtB_{wt}$ sequence (which express full length SrtB) from *Staphylococcus aureus* subsp. *aureus* (ATCC 700699D) genomic DNA by the polymerase chain reaction. The DNA fragments were cloned into the pET28b expression vector (Novagen, La Jolla, CA) to generate the construct pET28-SrtB30 and pET28-SrtB. Primers designated HindIII-srtB-F (5'-CCCAAGCTTAATATCGGAGGGTTTATTTTGAGAATGAAGCG ATTTTTAACTATTGTAC-3') and KpnI-FLAG-R (5'-GGGGATCCTTACTTGTCGTC ATCGTCTTTGTAGTCGTGGTGGTGGTGGTGGTGCTCGAG-3') were used to copy SrtB sequence with ribosome binding site on 5' end(*204*) and FLAG sequence on 3' end and cloned into the pNL9164 shuttle vector (Sigma-Aldrich, St. Louis, MO)(*198, 199*) to generate the construct pT-SrtB-FLAG.

Construction of fusion proteins *in vivo* expression vectors

Six fusion proteins were generated. $SrtA_MB_LB_C$ contains SrtA N-terminal residues 1-25 and SrtB C-terminal residues 26-244. $SrtB_MA_LA_C$ contains SrtB N-terminal residues 1-25 and SrtA C-terminal residues 26-206. $SrtA_MA_LB_C$ contains SrtA N-terminal residues 1-73 and SrtB C-terminal residues 70-244. $SrtB_MB_LA_C$ contains SrtB N-terminal residues 1-69 and SrtA C-terminal residues 74-206. $SrtA_MB_LA_C$ contains SrtA N-terminal residue 1-25, SrtB sequence 26-69, and SrtA C-terminal residues 74-206. $SrtB_MA_LB_C$ contains SrtB N-terminal residue 1-25, SrtA sequence 26-73, and SrtB C-terminal residues 70-244. Primers designated as Nco I srtA F (5'-CATGCCATGGGTCTATGGTTTGTGGTT CATAAG-3'), SrtABO1B (5'-CAATATATGTTTGAACAGCAAACAAATATGCTGC C-3'), SrtABO1C (5'-GGCAGCATATTTGTTTGCTGTTCAAACATATATTG-3'), SrtABO1D (5'-CCGCTCGAGACTTACCTTAATTATTTTTGCG-3'), SrtABO2B (5'-CCATCCAACAATGTCTGCCACTTTCGATTTATC-3'), SrtABO2C (5'-GATAAATC GAAAGTGGCAGACATTGTTGGATGG-3'), SrtBAO1B (5'- CGCTTCATTCTCATA CGTTAAGGCTCC-3'), SrtBAO1C (5'- GGAGCCTTAACGTATGAGAATGAAGCG-3'), SrtBAO1D (5'-GATAATTATCGATATGTGGTTTAATTTTGTAACCAAAAATG-3'), SrtBAO1E (5'- CATTTTTGGTTACAAAATTAAACCACATATCGATAATTATC-3'), Xho I srtA R (5'-CCGCTCGAGTTTGACTTCTGTAGCTACAAAGATT-3'), SrtBAO2D (5'-GAATTTCAATATAGCCTTTATTTATTTTTTCAAG-3'), SrtBAO2E (5'-CTTGAAAAAATAAATAAAGGCTATATTGAAATTC-3') were used to generate DNA fragments which express these fusion proteins listed above by PCR-driven overlap extension technique (*205*). The generated DNA fragments in pET28b(+) plasmid were then cloned into pNL9164 plasmid between Sph I and Kpn I sites to generate constructs pT-$SrtA_MB_LB_C$-FLAG, pT-$SrtB_MA_LA_C$-FLAG, pT-$SrtA_MA_LB_C$-FLAG, pT-$SrtB_MB_LA_C$-FLAG, pT-$SrtB_MA_LB_C$-FLAG, and pT-$SrtB_MA_LB_C$-FLAG.

Purification of $SrtB_{\Delta N30}$ protein from *E. coli*, polyacrylamide gel electrophoresis, western blot analysis and size-exclusion chromatography

The protocol was modified from previous methods described in Chapter 2.

Chemical cross-linking in *S. aureus*

RN4220 cells (general gift from Dr. Alan Lambowitz) transformed with the pT-SrtB-FLAG, pT-$SrtA_MB_LB_C$-FLAG, pT-$SrtB_MA_LA_C$-FLAG, pT-$SrtA_MA_LB_C$-FLAG, pT-$SrtB_MB_LA_C$-FLAG, pT-$SrtB_MA_LB_C$-FLAG, and pT-$SrtB_MA_LB_C$-FLAG constructs respectively were grown in 5 mL of Tryptic Soy Broth (TSB) media supplemented with 10 μg/mL erythromycin (Sigma, St. Louis, MO) at 32 ºC. When the OD_{600} reached 0.5. 10 μM cadmium chloride hemi (pentahydrate) (Sigma, St. Louis, MO) was added in pT-SrtB-FLAG to induce SrtB expression for another 3 hours before EDC treatment. Cultures with other constructs were grown till OD_{600} 0.8, 1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (EDC, Sigma, St. Louis, MO) was added to the media to a final concentration of 0.7 mg/mL. Cells were harvested 60 minutes later and treated with 1 μg/mL lysostaphin (Sigma, St. Louis, MO) at room temperature for 1 hour before lysed under native conditions using sonication. Cell lysate was analyzed by SDS-PAGE and western blot using anti-FLAG antibody (Qiagen, Valencia, CA).

References

1. Beers, M. H., Berkow, R., Bogin, R. M., Fletcher, A. J., Nace, B. A., Moy, D. E., and Merck Research Laboratories. (1999) The Merck manual of diagnosis and therapy, 17th ed., Merck, Whitehouse Station, NJ.

2. Lowy, F. D. (1998) Staphylococcus aureus infections, *N Engl J Med 339*, 520-532.

3. Banerjee, S. N., Emori, T. G., Culver, D. H., Gaynes, R. P., Jarvis, W. R., Horan, T., Edwards, J. R., Tolson, J., Henderson, T., and Martone, W. J. (1991) Secular trends in nosocomial primary bloodstream infections in the United States, 1980-1989. National Nosocomial Infections Surveillance System, *Am J Med 91*, 86S-89S.

4. Steinberg, J. P., Clark, C. C., and Hackman, B. O. (1996) Nosocomial and community-acquired Staphylococcus aureus bacteremias from 1980 to 1993: impact of intravascular devices and methicillin resistance, *Clin Infect Dis 23*, 255-259.

5. Noble, W. C., Valkenburg, H. A., and Wolters, C. H. (1967) Carriage of Staphylococcus aureus in random samples of a normal population, *J Hyg (Lond) 65*, 567-573.

6. Casewell, M. W., and Hill, R. L. (1986) The carrier state: methicillin-resistant Staphylococcus aureus, *J Antimicrob Chemother 18 Suppl A*, 1-12.

7. von Eiff, C., Becker, K., Machka, K., Stammer, H., and Peters, G. (2001) Nasal Carriage as a Source of Staphylococcus aureus Bacteremia, *The New England Journal of Medicine 344*, 6.

8. Wenzel, R. P., and Perl, T. M. (1995) The significance of nasal carriage of Staphylococcus aureus and the incidence of postoperative wound infection, *J Hosp Infect 31*, 13-24.

9. Hamill, R. J., Vann, J. M., and Proctor, R. A. (1986) Phagocytosis of Staphylococcus aureus by cultured bovine aortic endothelial cells: model for postadherence events in endovascular infections, *Infect Immun 54*, 833-836.

10. Lowy, F. D., Fant, J., Higgins, L. L., Ogawa, S. K., and Hatcher, V. B. (1988) Staphylococcus aureus--human endothelial cell interactions, *J Ultrastruct Mol Struct Res 98*, 137-146.

11. Hudson, M. C., Ramp, W. K., Nicholson, N. C., Williams, A. S., and Nousiainen, M. T. (1995) Internalization of Staphylococcus aureus by cultured osteoblasts, *Microb Pathog 19*, 409-419.

12. Almeida, R. A., Matthews, K. R., Cifrian, E., Guidry, A. J., and Oliver, S. P. (1996) Staphylococcus aureus invasion of bovine mammary epithelial cells, *J Dairy Sci 79*, 1021-1026.

13. Patti, J. M., and Hook, M. (1994) Microbial adhesins recognizing extracellular matrix macromolecules, *Curr Opin Cell Biol 6*, 752-758.

14. Ogawa, S. K., Yurberg, E. R., Hatcher, V. B., Levitt, M. A., and Lowy, F. D. (1985) Bacterial adherence to human endothelial cells in vitro, *Infect Immun 50*, 218-224.

15. Proctor, R. A., van Langevelde, P., Kristjansson, M., Maslow, J. N., and Arbeit, R. D. (1995) Persistent and relapsing infections associated with small-colony variants of Staphylococcus aureus, *Clin Infect Dis 20*, 95-102.

16. Wilkinson, B. (1997) *The staphylococci in human disease*, 1st ed., W.B. Saunders Company, New York.

17. Drake, T. A., and Pang, M. (1988) Staphylococcus aureus induces tissue factor expression in cultured human cardiac valve endothelium, *J Infect Dis 157*, 749-756.

18. Yao, L., Bengualid, V., Lowy, F. D., Gibbons, J. J., Hatcher, V. B., and Berman, J. W. (1995) Internalization of Staphylococcus aureus by endothelial cells induces cytokine gene expression, *Infect Immun 63*, 1835-1839.

19. Heumann, D., Barras, C., Severin, A., Glauser, M. P., and Tomasz, A. (1994) Gram-positive cell walls stimulate synthesis of tumor necrosis factor alpha and interleukin-6 by human monocytes, *Infect Immun 62*, 2715-2721.

20. Timmerman, C. P., Mattsson, E., Martinez-Martinez, L., De Graaf, L., Van Strijp, J. A., Verbrugh, H. A., Verhoef, J., and Fleer, A. (1993) Induction of release of tumor necrosis factor from human monocytes by staphylococci and staphylococcal peptidoglycans, *Infect Immun 61*, 4167-4172.

21. Bone, R. C. (1994) Gram-positive organisms and sepsis, *Arch Intern Med 154*, 26-34.

22. Bohach, G. A., Fast, D. J., Nelson, R. D., and Schlievert, P. M. (1990) Staphylococcal and streptococcal pyrogenic toxins involved in toxic shock syndrome and related illnesses, *Crit Rev Microbiol 17*, 251-272.

23. Marrack, P., and Kappler, J. (1990) The staphylococcal enterotoxins and their relatives, *Science 248*, 705-711.

24. Klein, E., Smith, D. L., and Laxminarayan, R. (2007) Hospitalizations and deaths caused by methicillin-resistant Staphylococcus aureus, United States, 1999-2005, *Emerg Infect Dis 13*, 1840-1846.

25. (2007) MRSA Deaths continue to rise in 2005, (Online, U. K. O. f. N. S., Ed.).

26. Appelbaum, P. C. (2006) MRSA--the tip of the iceberg, *Clin Microbiol Infect 12 Suppl 2*, 3-10.

27. Bonomo, R. A., and Tolmasky, M. E. (2007) *Enzyme-Mediated Resistance to Antibiotics: Mechanisms, Dissemination, and Prospects for Inhibition*, The American Society of Microbiology Press, Washington DC.

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

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

30. Schleifer, K. H., and Kandler, O. (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications, *Bacteriol Rev 36*, 407-477.

31. Giesbrecht, P., Kersten, T., Maidhof, H., and Wecke, J. (1998) Staphylococcal cell wall: morphogenesis and fatal variations in the presence of penicillin, *Microbiol Mol Biol Rev 62*, 1371-1414.

32. Navarre, W. W., and Schneewind, O. (1999) Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope, *Microbiol Mol Biol Rev 63*, 174-229.

33. Neuhaus, F. C., and Baddiley, J. (2003) A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria, *Microbiol Mol Biol Rev 67*, 686-723.

34. O'Riordan, K., and Lee, J. C. (2004) Staphylococcus aureus capsular polysaccharides, *Clin Microbiol Rev 17*, 218-234.

35. Mazmanian, S. K., Skaar, E. P., Gaspar, A. H., Humayun, M., Gornicki, P., Jelenska, J., Joachmiak, A., Missiakas, D. M., and Schneewind, O. (2003) Passage of heme-iron across the envelope of Staphylococcus aureus, *Science 299*, 906-909.

36. Bolken, T. C., Franke, C. A., Jones, K. F., Zeller, G. O., Jones, C. H., Dutton, E. K., and Hruby, D. E. (2001) Inactivation of the srtA gene in Streptococcus gordonii inhibits cell wall anchoring of surface proteins and decreases in vitro and in vivo adhesion, *Infect Immun 69*, 75-80.

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

38. Alksne, L. E., and Projan, S. J. (2000) Bacterial virulence as a target for antimicrobial chemotherapy, *Curr Opin Biotechnol 11*, 625-636.

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

40. Pallen, M. J., Lam, A. C., Antonio, M., and Dunbar, K. (2001) An embarrassment of sortases - a richness of substrates?, *Trends Microbiol 9*, 97-102.

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

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

43. Sjodahl, J. (1977) Repetitive sequences in protein A from Staphylococcus aureus. Arrangement of five regions within the protein, four being highly homologous and Fc-binding, *Eur J Biochem 73*, 343-351.

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

45. Perry, A. M., Ton-That, H., Mazmanian, S. K., and Schneewind, O. (2002) Anchoring of surface proteins to the cell wall of Staphylococcus aureus. III. Lipid II is an in vivo peptidoglycan substrate for sortase-catalyzed surface protein anchoring, *J Biol Chem 277*, 16241-16248.

46. Ruzin, A., Severin, A., Ritacco, F., Tabei, K., Singh, G., Bradford, P. A., Siegel, M. M., Projan, S. J., and Shlaes, D. M. (2002) Further evidence that a cell wall precursor [C(55)-MurNAc-(peptide)-GlcNAc] serves as an acceptor in a sorting reaction, *J Bacteriol 184*, 2141-2147.

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

48. Ton-That, H., Mazmanian, S. K., Faull, K. F., and Schneewind, O. (2000) Anchoring of surface proteins to the cell wall of Staphylococcus aureus. Sortase catalyzed in vitro transpeptidation reaction using LPXTG peptide and NH(2)-Gly(3) substrates, *J Biol Chem 275*, 9876-9881.

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

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

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

52. Race, P. R., Bentley, M. L., Melvin, J. A., Crow, A., Hughes, R. K., Smith, W. D., Sessions, R. B., Kehoe, M. A., McCafferty, D. G., and Banfield, M. J. (2009) Crystal structure of Streptococcus pyogenes sortase A: implications for sortase mechanism, *J Biol Chem 284*, 6924-6933.

53. Zhang, R., Wu, R., Joachimiak, G., Mazmanian, S. K., Missiakas, D. M., Gornicki, P., Schneewind, O., and Joachimiak, A. (2004) Structures of sortase B from Staphylococcus aureus and Bacillus anthracis reveal catalytic amino acid triad in the active site, *Structure 12*, 1147-1156.

54. Zong, Y., Mazmanian, S. K., Schneewind, O., and Narayana, S. V. (2004) The structure of sortase B, a cysteine transpeptidase that tethers surface protein to the Staphylococcus aureus cell wall, *Structure 12*, 105-112.

55. Maresso, A. W., Wu, R., Kern, J. W., Zhang, R., Janik, D., Missiakas, D. M., Duban, M. E., Joachimiak, A., and Schneewind, O. (2007) Activation of inhibitors by sortase triggers irreversible modification of the active site, *J Biol Chem 282*, 23129-23139.

56. Manzano, C., Contreras-Martel, C., El Mortaji, L., Izore, T., Fenel, D., Vernet, T., Schoehn, G., Di Guilmi, A. M., and Dessen, A. (2008) Sortase-mediated pilus fiber biogenesis in Streptococcus pneumoniae, *Structure 16*, 1838-1848.

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

58. Suree, N., Yi, S. W., Thieu, W., Marohn, M., Damoiseaux, R., Chan, A., Jung, M. E., and Clubb, R. T. (2009) Discovery and structure-activity relationship analysis of Staphylococcus aureus sortase A inhibitors, *Bioorg Med Chem 17*, 7174-7185.

59. Naik, M. T., Suree, N., Ilangovan, U., Liew, C. K., Thieu, W., Campbell, D. O., Clemens, J. J., Jung, M. E., and Clubb, R. T. (2005) Staphylococcus aureus sortase a transpeptidase: Calcium promotes sorting signal binding by altering the mobility and structure of an active loop, *J Biol Chem*.

60. Bentley, M. L., Gaweska, H., Kielec, J. M., and McCafferty, D. G. (2007) Engineering the substrate specificity of Staphylococcus aureus Sortase A. The beta6/beta7 loop from SrtB confers NPQTN recognition to SrtA, *J Biol Chem 282*, 6571-6581.

61. Bentley, M. L., Lamb, E. C., and McCafferty, D. G. (2008) Mutagenesis studies of substrate recognition and catalysis in the sortase A transpeptidase from Staphylococcus aureus, *J Biol Chem 283*, 14762-14771.

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

63. Connolly, K. M., Smith, B. T., Pilpa, R., Ilangovan, U., Jung, M. E., and Clubb, R. T. (2003) Sortase from Staphylococcus aureus does not contain a thiolate-imidazolium ion pair in its active site, *J Biol Chem 278*, 34061-34065.

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

65. Marraffini, L. A., Ton-That, H., Zong, Y., Narayana, S. V., and Schneewind, O. (2004) Anchoring of surface proteins to the cell wall of Staphylococcus aureus. A conserved arginine residue is required for efficient catalysis of sortase A, *J Biol Chem 279*, 37763-37770.

66. Frankel, B. A., Tong, Y., Bentley, M. L., Fitzgerald, M. C., and McCafferty, D. G. (2007) Mutational analysis of active site residues in the Staphylococcus aureus transpeptidase SrtA, *Biochemistry 46*, 7269-7278.

67. Holtje, J. V. (1996) A hypothetical holoenzyme involved in the replication of the murein sacculus of Escherichia coli, *Microbiology 142 (Pt 8)*, 1911-1918.

68. Holtje, J. V. (1993) *Bacterial Growth and Lysis-Metabolism and Structure of the Bacterial Sacculus*, Plenum Press, New York.

69. Pinho, M. G., and Errington, J. (2005) Recruitment of penicillin-binding protein PBP2 to the division site of Staphylococcus aureus is dependent on its transpeptidation substrates, *Mol Microbiol 55*, 799-807.

70. Kline, K. A., Kau, A. L., Chen, S. L., Lim, A., Pinkner, J. S., Rosch, J., Nallapareddy, S. R., Murray, B. E., Henriques-Normark, B., Beatty, W., Caparon, M. G., and Hultgren, S. J. (2009) Mechanism for sortase localization and the role of sortase localization in efficient pilus assembly in Enterococcus faecalis, *J Bacteriol 191*, 3237-3247.

71. Raz, A., and Fischetti, V. A. (2008) Sortase A localizes to distinct foci on the Streptococcus pyogenes membrane, *Proc Natl Acad Sci U S A 105*, 18549-18554.

72. DeDent, A., Bae, T., Missiakas, D. M., and Schneewind, O. (2008) Signal peptides direct surface proteins to two distinct envelope locations of Staphylococcus aureus, *EMBO J 27*, 2656-2668.

73. Park, B. S., Kim, J. G., Kim, M. R., Lee, S. E., Takeoka, G. R., Oh, K. B., and Kim, J. H. (2005) Curcuma longa L. constituents inhibit sortase A and Staphylococcus aureus cell adhesion to fibronectin, *J Agric Food Chem 53*, 9005-9009.

74. Oh, K. B., Oh, M. N., Kim, J. G., Shin, D. S., and Shin, J. (2005) Inhibition of sortase-mediated Staphylococcus aureus adhesion to fibronectin via fibronectin-binding protein by sortase inhibitors, *Appl Microbiol Biotechnol*, 1-5.

75. Oh, K. B., Mar, W., Kim, S., Kim, J. Y., Oh, M. N., Kim, J. G., Shin, D., Sim, C. J., and Shin, J. (2005) Bis(indole) alkaloids as sortase A inhibitors from the sponge Spongosorites sp, *Bioorg Med Chem Lett 15*, 4927-4931.

76. Kim, S. H., Shin, D. S., Oh, M. N., Chung, S. C., Lee, J. S., and Oh, K. B. (2004) Inhibition of the bacterial surface protein anchoring transpeptidase sortase by isoquinoline alkaloids, *Biosci Biotechnol Biochem 68*, 421-424.

77. Kim, S. H., Shin, D. S., Oh, M. N., Chung, S. C., Lee, J. S., Chang, I. M., and Oh, K. B. (2003) Inhibition of sortase, a bacterial surface protein anchoring transpeptidase, by beta-sitosterol-3-O-glucopyranoside from Fritillaria verticillata, *Biosci Biotechnol Biochem 67*, 2477-2479.

78. Kim, S. W., Chang, I. M., and Oh, K. B. (2002) Inhibition of the bacterial surface protein anchoring transpeptidase sortase by medicinal plants, *Biosci Biotechnol Biochem 66*, 2751-2754.

79. Oh, K. B., Oh, M. N., Kim, J. G., Shin, D. S., and Shin, J. (2006) Inhibition of sortase-mediated Staphylococcus aureus adhesion to fibronectin via fibronectin-binding protein by sortase inhibitors, *Appl Microbiol Biotechnol 70*, 102-106.

80. Kang, S. S., Kim, J. G., Lee, T. H., and Oh, K. B. (2006) Flavonols inhibit sortases and sortase-mediated Staphylococcus aureus clumping to fibrinogen, *Biol Pharm Bull 29*, 1751-1755.

81. Jang, K. H., Chung, S. C., Shin, J., Lee, S. H., Kim, T. I., Lee, H. S., and Oh, K. B. (2007) Aaptamines as sortase A inhibitors from the tropical sponge Aaptos aaptos, *Bioorg Med Chem Lett 17*, 5366-5369.

82. Scott, C. J., McDowell, A., Martin, S. L., Lynas, J. F., Vandenbroeck, K., and Walker, B. (2002) Irreversible inhibition of the bacterial cysteine protease-transpeptidase sortase (SrtA) by substrate-derived affinity labels, *Biochem J 366*, 953-958.

83. Frankel, B. A., Bentley, M., Kruger, R. G., and McCafferty, D. G. (2004) Vinyl sulfones: inhibitors of SrtA, a transpeptidase required for cell wall protein anchoring and virulence in Staphylococcus aureus, *J Am Chem Soc 126*, 3404-3405.

84. Supuran, C. T., Scozzafava, A., and Clare, B. W. (2002) Bacterial protease inhibitors, *Med Res Rev 22*, 329-372.

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

86. Kruger, R. G., Barkallah, S., Frankel, B. A., and McCafferty, D. G. (2004) Inhibition of the Staphylococcus aureus sortase transpeptidase SrtA by phosphinic peptidomimetics, *Bioorg Med Chem 12*, 3723-3729.

87. Liew, C. K., Smith, B. T., Pilpa, R., Suree, N., Ilangovan, U., Connolly, K. M., Jung, M. E., and Clubb, R. T. (2004) Localization and mutagenesis of the sorting signal binding site on sortase A from Staphylococcus aureus, *FEBS Lett 571*, 221-226.

88. Jung, M. E., Clemens, J. J., Suree, N., Liew, C. K., Pilpa, R., Campbell, D. O., and Clubb, R. T. (2005) Synthesis of (2R,3S) 3-amino-4-mercapto-2-butanol, a threonine analogue for covalent inhibition of sortases, *Bioorg Med Chem Lett 15*, 5076-5079.

89. Oh, K. B., Kim, S. H., Lee, J., Cho, W. J., Lee, T., and Kim, S. (2004) Discovery of diarylacrylonitriles as a novel series of small molecule sortase A inhibitors, *J Med Chem 47*, 2418-2421.

90. Weiss, W. J., Lenoy, E., Murphy, T., Tardio, L., Burgio, P., Projan, S. J., Schneewind, O., and Alksne, L. (2004) Effect of srtA and srtB gene expression on the virulence of Staphylococcus aureus in animal models of infection, *J Antimicrob Chemother 53*, 480-486.

91. Jonsson, I. M., Mazmanian, S. K., Schneewind, O., Verdrengh, M., Bremell, T., and Tarkowski, A. (2002) On the role of Staphylococcus aureus sortase and sortase-catalyzed surface protein anchoring in murine septic arthritis, *J Infect Dis 185*, 1417-1424.

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

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

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

95. Barnett, T. C., and Scott, J. R. (2002) Differential recognition of surface proteins in Streptococcus pyogenes by two sortase gene homologs, *J Bacteriol 184*, 2181-2191.

96. Lee, S. F., and Boran, T. L. (2003) Roles of sortase in surface expression of the major protein adhesin P1, saliva-induced aggregation and adherence, and cariogenicity of Streptococcus mutans, *Infect Immun 71*, 676-681.

97. Osaki, M., Takamatsu, D., Shimoji, Y., and Sekizaki, T. (2002) Characterization of Streptococcus suis genes encoding proteins homologous to sortase of gram-positive bacteria, *J Bacteriol 184*, 971-982.

98. Yamaguchi, M., Terao, Y., Ogawa, T., Takahashi, T., Hamada, S., and Kawabata, S. (2006) Role of Streptococcus sanguinis sortase A in bacterial colonization, *Microbes Infect 8*, 2791-2796.

99. Lalioui, L., Pellegrini, E., Dramsi, S., Baptista, M., Bourgeois, N., Doucet-Populaire, F., Rusniok, C., Zouine, M., Glaser, P., Kunst, F., Poyart, C., and Trieu-Cuot, P. (2005) The SrtA Sortase of Streptococcus agalactiae is required for cell wall anchoring of proteins containing the LPXTG motif, for adhesion to epithelial cells, and for colonization of the mouse intestine, *Infect Immun 73*, 3342-3350.

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

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

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

103. Buisson, M., Rivail, L., Hernandez, J. F., Jamin, M., Martinez, J., Ruigrok, R. W., and Burmeister, W. P. (2006) Kinetics, inhibition and oligomerization of Epstein-Barr virus protease, *FEBS Lett 580*, 6570-6578.

104. Shoemaker, B. A., and Panchenko, A. R. (2007) Deciphering protein-protein interactions. Part I. Experimental techniques and databases, *PLoS Comput Biol 3*, e42.

105. Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D. S., Zhang, L. V., Wong, S. L., Franklin, G., Li, S., Albala, J. S., Lim, J., Fraughton, C., Llamosas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R. S., Vandenhaute, J., Zoghbi, H. Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M. E.,

Hill, D. E., Roth, F. P., and Vidal, M. (2005) Towards a proteome-scale map of the human protein-protein interaction network, *Nature 437*, 1173-1178.

106. Carey, J. (1991) Gel retardation, *Methods Enzymol 208*, 103-117.

107. Wilton, R., Myatt, E. A., and Stevens, F. J. (2004) Analysis of protein-protein interactions by simulation of small-zone gel filtration chromatography, *Methods Mol Biol 261*, 137-154.

108. Herberg, F. W., and Taylor, S. S. (1993) Physiological inhibitors of the catalytic subunit of cAMP-dependent protein kinase: effect of MgATP on protein-protein interactions, *Biochemistry 32*, 14015-14022.

109. Phizicky, E. M., and Fields, S. (1995) Protein-protein interactions: methods for detection and analysis, *Microbiol Rev 59*, 94-123.

110. Stevens, F. J. (1986) Analysis of protein-protein interaction by simulation of small-zone size-exclusion chromatography: application to an antibody-antigen association, *Biochemistry 25*, 981-993.

111. Stevens, F. J., Carperos, W. E., Monafo, W. J., and Greenspan, N. S. (1988) Size-exclusion HPLC analysis of epitopes, *J Immunol Methods 108*, 271-278.

112. Myatt, E. A., Stevens, F. J., and Benjamin, C. (1994) Solution-phase binding of monoclonal antibodies to bee venom phospholipase A2, *J Immunol Methods 177*, 35-42.

113. Hummel, J. P., and Dreyer, W. J. (1962) Measurement of protein-binding phenomena by gel filtration, *Biochim Biophys Acta 63*, 530-532.

114. Yong, H., Thomas, G. A., and Peticolas, W. L. (1993) Metabolite-modulated complex formation between alpha-glycerophosphate dehydrogenase and lactate dehydrogenase, *Biochemistry 32*, 11124-11131.

115. Gegner, J. A., and Dahlquist, F. W. (1991) Signal transduction in bacteria: CheW forms a reversible complex with the protein kinase CheA, *Proc Natl Acad Sci U S A 88*, 750-754.

116. Soltes, L. (2004) The Hummel-Dreyer method: impact in pharmacology, *Biomed Chromatogr 18*, 259-271.

117. Lebowitz, J., Lewis, M. S., and Schuck, P. (2002) Modern analytical ultracentrifugation in protein science: a tutorial review, *Protein Sci 11*, 2067-2079.

118. Stafford, W. F. (1997) Sedimentation velocity spins a new weave for an old fabric, *Curr Opin Biotechnol 8*, 14-24.

119. Liu, J., Andya, J. D., and Shire, S. J. (2006) A critical review of analytical ultracentrifugation and field flow fractionation methods for measuring protein aggregation, *AAPS J 8*, E580-589.

120. Staros, J. V., Wright, R. W., and Swingle, D. M. (1986) Enhancement by N-hydroxysulfosuccinimide of water-soluble carbodiimide-mediated coupling reactions, *Anal Biochem 156*, 220-222.

121. Wood, C. L., and O'Dorisio, M. S. (1985) Covalent cross-linking of vasoactive intestinal polypeptide to its receptors on intact human lymphoblasts, *J Biol Chem 260*, 1243-1247.

122. Sorrentino, S., Yakovlev, G. I., and Libonati, M. (1982) Dimerization of deoxyribonuclease I, lysozyme and papain. Effects of ionic strength on enzymic activity, *Eur J Biochem 124*, 183-189.

123. Nisbet, A. D., Saundry, R. H., Moir, A. J., Fothergill, L. A., and Fothergill, J. E. (1981) The complete amino-acid sequence of hen ovalbumin, *Eur J Biochem 115*, 335-345.

124. Mathews, C. K., Sjoberg, B. M., and Reichard, P. (1987) Ribonucleotide reductase of Escherichia coli. Cross-linking agents as probes of quaternary and quinary structure, *Eur J Biochem 166*, 279-285.

125. Moshnikova, A. B., Afanasyev, V. N., Proussakova, O. V., Chernyshov, S., Gogvadze, V., and Beletsky, I. P. (2006) Cytotoxic activity of 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide is underlain by DNA interchain cross-linking, *Cell Mol Life Sci 63*, 229-234.

126. Darke, P. L., Jordan, S. P., Hall, D. L., Zugay, J. A., Shafer, J. A., and Kuo, L. C. (1994) Dissociation and association of the HIV-1 protease dimer subunits: equilibria and rates, *Biochemistry 33*, 98-105.

127. Renatus, M., Stennicke, H. R., Scott, F. L., Liddington, R. C., and Salvesen, G. S. (2001) Dimer formation drives the activation of the cell death protease caspase 9, *Proc Natl Acad Sci U S A 98*, 14250-14255.

128. Fersht, A. (1999) *Structure and mechansim in protein science*, W.H.Freeman and Company.

129. Sambrook, J., Fritsch, E. F., Maniatis T. (1989) *Molecular Cloning: A Laboratory Manual.*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

130. Shih, C. L., Chen, M. J., Linse, K., and Wang, K. (1997) Molecular contacts between nebulin and actin: cross-linking of nebulin modules to the N-terminus of actin, *Biochemistry 36*, 1814-1825.

131. Demeler, B. (2005) UltraScan Software, in *www.ultrascan.uthscsa.edu* 6.2 ed., The University of Texas Health Science Center at San Antonio.

132. Nooren, I. M., and Thornton, J. M. (2003) Structural characterisation and functional significance of transient protein-protein interactions, *J Mol Biol 325*, 991-1018.

133. Bahadur, R. P., Chakrabarti, P., Rodier, F., and Janin, J. (2004) A dissection of specific and non-specific protein-protein interfaces, *J Mol Biol 336*, 943-955.

134. Johnson, J. E. (1996) Functional implications of protein-protein interactions in icosahedral viruses, *Proc Natl Acad Sci U S A 93*, 27-33.

135. Chothia, C., and Janin, J. (1975) Principles of protein-protein recognition, *Nature 256*, 705-708.

136. Morgan, R. S., Miller, S. L., and McAdon, J. M. (1979) The symmetry of self-complementary surfaces, *J Mol Biol 127*, 31-38.

137. Connolly, M. L. (1986) Shape complementarity at the hemoglobin alpha 1 beta 1 subunit interface, *Biopolymers 25*, 1229-1247.

138. Lawrence, M. C., and Colman, P. M. (1993) Shape complementarity at protein/protein interfaces, *J Mol Biol 234*, 946-950.

139. Reichmann, D., Rahat, O., Cohen, M., Neuvirth, H., and Schreiber, G. (2007) The molecular architecture of protein-protein binding sites, *Curr Opin Struct Biol 17*, 67-76.

140. Xu, D., Lin, S. L., and Nussinov, R. (1997) Protein binding versus protein folding: the role of hydrophilic bridges in protein associations, *J Mol Biol 265*, 68-84.

141. Kundrotas, P. J., and Alexov, E. (2006) Electrostatic properties of protein-protein complexes, *Biophys J 91*, 1724-1736.

142. Uetz, P., and Vollert, C. S. (2006) Protein-protein interactions, in *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine* (Ganten, D., and Ruckpaul, K., Eds.) 1st edition ed., pp 1548-1552, Springer Verlag.

143. Reichmann, D., Rahat, O., Albeck, S., Meged, R., Dym, O., and Schreiber, G. (2005) The modular architecture of protein-protein binding interfaces, *Proc Natl Acad Sci U S A 102*, 57-62.

144. Keskin, O., Ma, B., Rogale, K., Gunasekaran, K., and Nussinov, R. (2005) Protein-protein interactions: organization, cooperativity and mapping in a bottom-up Systems Biology approach, *Phys Biol 2*, S24-35.

145. Brinda, K. V., and Vishveshwara, S. (2005) Oligomeric protein structure networks: insights into protein-protein interactions, *BMC Bioinformatics 6*, 296.

146. Keeble, A. H., and Kleanthous, C. (2005) The kinetic basis for dual recognition in colicin endonuclease-immunity protein complexes, *J Mol Biol 352*, 656-671.

147. Tobi, D., and Bahar, I. (2005) Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state, *Proc Natl Acad Sci U S A 102*, 18908-18913.

148. Grunberg, R., Nilges, M., and Leckner, J. (2006) Flexibility and conformational entropy in protein-protein binding, *Structure 14*, 683-693.

149. Pearce, K. H., Jr., Cunningham, B. C., Fuh, G., Teeri, T., and Wells, J. A. (1999) Growth hormone binding affinity for its receptor surpasses the requirements for cellular activity, *Biochemistry 38*, 81-89.

150. Brown, R. J., Adams, J. J., Pelekanos, R. A., Wan, Y., McKinstry, W. J., Palethorpe, K., Seeber, R. M., Monks, T. A., Eidne, K. A., Parker, M. W., and Waters, M. J. (2005) Model for growth hormone receptor activation based on subunit rotation within a receptor dimer, *Nat Struct Mol Biol 12*, 814-821.

151. Horn, J. R., Kraybill, B., Petro, E. J., Coales, S. J., Morrow, J. A., Hamuro, Y., and Kossiakoff, A. A. (2006) The role of protein dynamics in increasing binding affinity for an engineered protein-protein interaction established by H/D exchange mass spectrometry, *Biochemistry 45*, 8488-8498.

152. Cunningham, B. C., and Wells, J. A. (1989) High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis, *Science 244*, 1081-1085.

153. Clackson, T., and Wells, J. A. (1995) A hot spot of binding energy in a hormone-receptor interface, *Science 267*, 383-386.

154. Bogan, A. A., and Thorn, K. S. (1998) Anatomy of hot spots in protein interfaces, *J Mol Biol 280*, 1-9.

155. Burdine, L., Gillette, T. G., Lin, H. J., and Kodadek, T. (2004) Periodate-triggered cross-linking of DOPA-containing peptide-protein complexes, *J Am Chem Soc 126*, 11442-11443.

156. Burzio, L. A., and Waite, J. H. (2000) Cross-linking in adhesive quinoproteins: studies with model decapeptides, *Biochemistry 39*, 11147-11153.

157. Liu, B., Burdine, L., and Kodadek, T. (2006) Chemistry of periodate-mediated cross-linking of 3,4-dihydroxylphenylalanine-containing molecules to proteins, *J Am Chem Soc 128*, 15228-15235.

158. Xie, J., and Schultz, P. G. (2005) Adding amino acids to the genetic repertoire, *Curr Opin Chem Biol 9*, 548-554.

159. Alfonta, L., Zhang, Z., Uryu, S., Loo, J. A., and Schultz, P. G. (2003) Site-specific incorporation of a redox-active amino acid into proteins, *J Am Chem Soc 125*, 14662-14663.

160. Zhang, Z., Smith, B. A., Wang, L., Brock, A., Cho, C., and Schultz, P. G. (2003) A new strategy for the site-specific modification of proteins in vivo, *Biochemistry 42*, 6735-6746.

161. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (2002) *Current protocols in molecular biology*, Vol. 2, 5 ed., John Wiley and Sons, Inc.

162. Putnam, C. (2006) Protein calculator v3.3.

163. Fasman, G. D. (1996) *Circular dichroism and the conformational analysis of biomolecules*, Plenum Press, New York.

164. Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set, *Anal Biochem 287*, 252-260.

165. Lobley, A., Whitmore, L., and Wallace, B. A. (2002) DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra, *Bioinformatics 18*, 211-212.

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

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

168. Paddock, M. L., Weber, K. H., Chang, C., and Okamura, M. Y. (2005) Interactions between cytochrome c2 and the photosynthetic reaction center from Rhodobacter sphaeroides: the cation-pi interaction, *Biochemistry 44*, 9619-9625.

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

170. Mao, H. (2004) A self-cleavable sortase fusion for one-step purification of free recombinant proteins, *Protein Expr Purif 37*, 253-263.

171. Kegler-Ebo, D. M., Docktor, C. M., and DiMaio, D. (1994) Codon cassette mutagenesis: a general method to insert or replace individual codons by using universal mutagenic cassettes, *Nucleic Acids Res 22*, 1593-1599.

172. Sreerama, N., and Woody, R. W. (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism, *Anal Biochem 209*, 32-44.

173. Provencher, S. W., and Glockner, J. (1981) Estimation of globular protein secondary structure from circular dichroism, *Biochemistry 20*, 33-37.

174. van Stokkum, I. H., Spoelder, H. J., Bloemendal, M., van Grondelle, R., and Groen, F. C. (1990) Estimation of protein secondary structure and error analysis from circular dichroism spectra, *Anal Biochem 191*, 110-118.

175. Mao, D., Wachter, E., and Wallace, B. A. (1982) Folding of the mitochondrial proton adenosinetriphosphatase proteolipid channel in phospholipid vesicles, *Biochemistry 21*, 4960-4968.

176. Brahms, S., and Brahms, J. (1980) Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism, *J Mol Biol 138*, 149-178.

177. Thibodeaux, G. N., Cowmeadow, R., Umeda, A., and Zhang, Z. (2009) A tetracycline repressor-based mammalian two-hybrid system to detect protein-protein interactions in vivo, *Anal Biochem 386*, 129-131.

178. Umeda, A., Thibodeaux, G. N., Zhu, J., Lee, Y., and Zhang, Z. (2009) Site-specific protein cross-linking with genetically incorporated 3,4-dihydroxy-L-phenylalanine, *Chembiochem In press*.

179. Yao, J., Zhong, J., Fang, Y., Geisinger, E., Novick, R. P., and Lambowitz, A. M. (2006) Use of targetrons to disrupt essential and nonessential genes in Staphylococcus aureus reveals temperature sensitivity of Ll.LtrB group II intron splicing, *Rna 12*, 1271-1281.

180. Sjoquist, J., Meloun, B., and Hjelm, H. (1972) Protein A isolated from Staphylococcus aureus after digestion with lysostaphin, *Eur J Biochem 29*, 572-578.

181. Wolfe, M. S., and Kopan, R. (2004) Intramembrane proteolysis: theme and variations, *Science 305*, 1119-1123.

182. Sakoh, M., Ito, K., and Akiyama, Y. (2005) Proteolytic activity of HtpX, a membrane-bound and stress-controlled protease from Escherichia coli, *J Biol Chem 280*, 33305-33310.

183. van Roosmalen, M. L., Geukens, N., Jongbloed, J. D., Tjalsma, H., Dubois, J. Y., Bron, S., van Dijl, J. M., and Anne, J. (2004) Type I signal peptidases of Gram-positive bacteria, *Biochim Biophys Acta 1694*, 279-297.

184. DeDent, A. C., McAdow, M., and Schneewind, O. (2007) Distribution of protein A on the surface of Staphylococcus aureus, *J Bacteriol 189*, 4473-4484.

185. Hu, P., Bian, Z., Fan, M., Huang, M., and Zhang, P. (2008) Sec translocase and sortase A are colocalised in a locus in the cytoplasmic membrane of Streptococcus mutans, *Arch Oral Biol 53*, 150-154.

186. Vollmer, W., and Bertsche, U. (2008) Murein (peptidoglycan) structure, architecture and biosynthesis in Escherichia coli, *Biochim Biophys Acta 1778*, 1714-1734.

187. Jevon, M., Guo, C., Ma, B., Mordan, N., Nair, S. P., Harris, M., Henderson, B., Bentley, G., and Meghji, S. (1999) Mechanisms of internalization of Staphylococcus aureus by cultured human osteoblasts, *Infect Immun 67*, 2677-2681.

188. Ahmed, S., Meghji, S., Williams, R. J., Henderson, B., Brock, J. H., and Nair, S. P. (2001) Staphylococcus aureus fibronectin binding proteins are essential for internalization by osteoblasts but do not account for differences in intracellular levels of bacteria, *Infect Immun 69*, 2872-2877.

189. Alexander, E. H., and Hudson, M. C. (2001) Factors influencing the internalization of Staphylococcus aureus and impacts on the course of infections in humans, *Appl Microbiol Biotechnol 56*, 361-366.

190. Sinha, B., Francois, P., Que, Y. A., Hussain, M., Heilmann, C., Moreillon, P., Lew, D., Krause, K. H., Peters, G., and Herrmann, M. (2000) Heterologously expressed Staphylococcus aureus fibronectin-binding proteins are sufficient for invasion of host cells, *Infect Immun 68*, 6871-6878.

191. Sinha, B., Francois, P. P., Nusse, O., Foti, M., Hartford, O. M., Vaudaux, P., Foster, T. J., Lew, D. P., Herrmann, M., and Krause, K. H. (1999) Fibronectin-binding protein acts as Staphylococcus aureus invasin via fibronectin bridging to integrin alpha5beta1, *Cell Microbiol 1*, 101-117.

192. Dziewanowska, K., Patti, J. M., Deobald, C. F., Bayles, K. W., Trumble, W. R., and Bohach, G. A. (1999) Fibronectin binding protein and host cell tyrosine kinase are required for internalization of Staphylococcus aureus by epithelial cells, *Infect Immun 67*, 4673-4678.

193. Pils, S., Schmitter, T., Neske, F., and Hauck, C. R. (2006) Quantification of bacterial invasion into adherent cells by flow cytometry, *Journal of Microbiological Methods 65*, 301-310.

194. Pils, S., Schmitter, T., Neske, F., and Hauck, C. R. (2006) Quantification of bacterial invasion into adherent cells by flow cytometry, *J Microbiol Methods 65*, 301-310.

195. Zong, Y., Mazmanian, S. K., Schneewind, O., and Narayana, S. V. (2004) The structure of sortase B, a cysteine transpeptidase that tethers surface protein to the Staphylococcus aureus cell wall, *Structure (Camb) 12*, 105-112.

196. Dessen, A. (2004) A new catalytic dyad regulates anchoring of molecules to the Gram-positive cell wall by sortases, *Structure (Camb) 12*, 6-7.

197. Zhang, R., Wu, R., Joachimiak, G., Mazmanian, S. K., Missiakas, D. M., Gornicki, P., Schneewind, O., and Joachimiak, A. (2004) Structures of sortase B from Staphylococcus aureus and Bacillus anthracis reveal catalytic amino acid triad in the active site, *Structure (Camb) 12*, 1147-1156.

198. Grkovic, S., Brown, M. H., Hardie, K. M., Firth, N., and Skurray, R. A. (2003) Stable low-copy-number Staphylococcus aureus shuttle vectors, *Microbiology 149*, 785-794.

199. Charpentier, E., Anton, A. I., Barry, P., Alfonso, B., Fang, Y., and Novick, R. P. (2004) Novel cassette-based shuttle vector system for gram-positive bacteria, *Appl Environ Microbiol 70*, 6076-6085.

200. Qiagen. The QIAexpressionist: A handbook for high-level expression and purification of 6XHis-tagged proteins.

201. Pourkomaillan, B. (1998) Possible link between a 35 kDa membrane protein and osmolyte transport in Staphylococcus aureus, *Lett Appl Microbiol 26*, 149-152.

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

203. Jonsson, I. M., Mazmanian, S. K., Schneewind, O., Bremell, T., and Tarkowski, A. (2003) The role of Staphylococcus aureus sortase A and sortase B in murine arthritis, *Microbes Infect 5*, 775-780.

204. McLaughlin, J. R., Murray, C. L., and Rabinowitz, J. C. (1981) Unique features in the ribosome binding site sequence of the gram-positive Staphylococcus aureus beta-lactamase gene, *J Biol Chem 256*, 11283-11291.

205. Heckman, K. L., and Pease, L. R. (2007) Gene splicing and mutagenesis by PCR-driven overlap extension, *Nat Protoc 2*, 924-932.

Vita

Jie Zhu was born on August 29, 1980 as the daughter of Mr. Xinsheng Zhu and Mrs. Suyi Cui at Qinjiang in China. She was raised in Nanjing, China and attended Qingliangshan Elementary School, Jinling Middle School, and Jinling High School. In 1998 she entered Nanjing University, China and started research about plant physiology in Dr. Yonghua Yang's lab since fall, 2000. She received the degree of Bachelor of Science from Nanjing University in June 2002. During the following two years she studied at The University of Kentucky, at Lexington in Dr. Balazs Siminsky's lab and was awarded the M.Sc. degree for plant physiology. In 2004 she moved to The University of Texas, at Austin to pursue her Ph.D. degree in Cellular and Molecular Biology and worked in Dr. Zhiwen Jonathan Zhang's lab.

Permanent address: 15 lakeshore Road, D2 31 World Village, Nanjing, Jiangsu 211000, P. R. China

This dissertation was typed by Jie Zhu.