Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Biologia Cellulare, Molecolare e Industriale

Ciclo XXV

Settore Concorsuale di afferenza: 05/E2- BIOLOGIA MOLECOLARE

Settore Scientifico disciplinare: BIO/11- BIOLOGIA MOLECOLARE

Characterization of the *Staphylococcus aureus* bone sialoprotein-binding protein SdrE and the serine protease EpiP

Presentata da

Prachi

Coordinatore Dottorato	Relatore	Co-Relatore
Prof. Vincenzo Scarlato	Prof. Vincenzo Scarlato	Dr. Fabio Bagnoli

Esame finale anno 2013

Alma Mater Studiorum – University of Bologna

DOCTOR OF PHILOSOPHY IN

Cellular, molecular and industrial biology

Cycle XXV

Scientific Discipline: BIO/11- MOLECULAR BIOLOGY

Characterization of the *Staphylococcus aureus* bone sialoprotein-binding protein SdrE and the serine protease EpiP

Submitted by

Prachi

Ph.D. Coordinator

Prof. Vincenzo Scarlato

Ph.D. supervisor

Dr. Fabio Bagnoli

Final examination year 2013

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name: Prachi

Signature:

ABSTRACT

CHARACTERIZATION OF THE *STAPHYLOCOCCUS AUREUS* BONE SIALOPROTEIN-BINDING PROTEIN SdrE AND THE SERINE PROTEASE EpiP

Prachi

Supervisor: Dr. Fabio Bagnoli
Coordinator: Prof. Vincenzo Scarlato

In the last decade, the reverse vaccinology approach shifted the paradigm of vaccine discovery from conventional culture-based methods to high-throughput genome-based approaches for the development of recombinant protein-based vaccines against pathogenic bacteria. The concept of reverse vaccinology was applied for the first time to serogroup B *Neisseria meningitidis* (MenB). In an attempt to develop a *Staphylococcus aureus* vaccine, we have applied a similar approach, mainly based on *in silico* screening and proteomics to select surface-exposed proteins, and identified SdrE, as one of the potential vaccine antigen against *S. aureus*. We characterized SdrE, a protein belonging to the serine-aspartate repeat (Sdr) protein family. We have investigated the biochemical properties as well as the vaccine potential of SdrE and its highly conserved CnaBE3 domain. We found the protein SdrE to be resistant to trypsin at 37°C. Mass spectrometry and N-terminal sequencing by Edman degradation of the resistant fragment revealed that it comprises a CnaBE3 domain of the protein. rCnaBE3 also showed partial trypsin resistant behavior. Furthermore, intact mass spectrometry of rCnaBE3 showed mass difference of 17 Da between theoretical and observed mass, suggesting the presence of isopeptide bond or some other post-translational modification. However, this observation needs further investigation. Furthermore, we found the last CnaB domain of other Sdr proteins i.e. CnaBE3 of SdrE to be highly conserved together with high sequence similarity with the last CnaB domain i.e. SdrC2 of SdrC and SdrD5 of SdrD respectively, even in strains that are phylogenetically distant. We found SdrE immunogenic against a clinical strain of *S. aureus* in murine abscess model. Moreover, we found the highly conserved CnaBE3 domain to be cross-reactive to other Sdr proteins and more interestingly mice immunized with the domain were protected from infection with a *S. aureus* strains lacking *sdrE*. Overall, our data suggested a possible role of CnaBE3 domain as vaccine candidate supported by the high homology, conservation, immunogenicity and

stability. In addition, Differential Scanning Fluorimetry (DSF) study reveals the effect of calcium in folding of SdrE. Indeed, the melting temperature (Tm) of the recombinant protein is increased with the increasing concentration of calcium (Ca^{+2}).

We also identified and characterized a protein, annotated as epidermin leader peptide processing serine protease (EpiP), as a novel *S. aureus* vaccine candidate. In collaboration with Northwestern University Chicago, USA, we determined the structure of the purified protein EpiP (rEpiP) at 2.05 Å (1 Å=0.1 nm) resolution by x-ray crystallography, revealing the fold of subtilisin-like protease in the protease domain. Moreover, we found the pro-domain non-covalently linked to the protease domain by a polypeptide chain acting as rubber. The crystal structure also showed that rEpiP was cleaved somewhere between residues 95 and 100 and we found that the cleavage occurs through an autocatalytic intra-molecular mechanism. In addition, the protein expressed by *S. aureus* cells also appeared to undergo a similar processing event. To determine if the protein acts as a serine protease, we mutated the hypothesized catalytic serine 393 residue to alanine, generating rEpiP-S393A. The crystal structure at a resolution of 1.95 Å of this mutant protein showed that the polypeptide chain was not cleaved and was not interacting stably with the active site. Indeed, rEpiP-S393A was shown to be impaired in its protease activity. Mice vaccinated with EpiP were protected from *S. aureus* infection. In addition, protective efficacy generated by the rEpiP and the non-cleaving mutant protein was comparable, implying that the two forms are interchangeable for vaccination purposes. This study revealed the first fundamental biochemical properties of this novel lantibiotic processing serine protease EpiP.

Key words: *Staphylococcus aureus*, reverse vaccinology, proteomics, serine-aspartate repeat protein, Differential Scanning Fluorimetry , EpiP.

ACKNOWLEDGEMENTS

My Ph.D. has been a new and long period of my life, that was 4 years spent in Italy. Completion of this doctoral dissertation was possible with the support of several people. I would like to express my sincere gratitude to all of them. It has been a great privilege to spend these memorable years in the Department of Protein Biochemistry, Novartis Vaccines Research Center, Siena, Italy and its members will always remain dear to me.

My first debt of gratitude must go to my supervisor Dr. Fabio Bagnoli. He patiently provided the vision, encouragement and advice necessary for me to proceed through the doctoral program and complete my dissertation. I want to thank him for his unflagging encouragement and serving as a role model to me as a supervisor. He has been a strong and supportive adviser to me throughout my doctoral career, he has always given me great freedom to pursue independent work.

Special thanks to my day to day mentor Dr. Silvana Savino for her support, guidance and helpful suggestions. Her guidance has served me well and I owe her my heartfelt appreciation.

I would like to express my deep gratitude and respect to Dr. Markus Hilleringmann, mentor for my first year of Ph.D., whose advices and insight was invaluable to me. For all I learned from him, and for providing the lab vision for the experiments. Besides my supervisor and day to day mentor, I would like to thank the rest of my Ph.D committee members: Dr. Nathalie Norais, Dr. Matthew Bottomley, and Dr. Paolo Ruggiero for their encouragement, insightful comments, hard questions and stimulating discussions. I also acknowledge to Massimiliano Biagini, Andrea Mannetti, Fabiana Falugi, Cecilia Brettoni, MariaRita Fontana and Elisa Viciani.

I would also like to thank Prof. Dr. Vincenzo Scarlato for his kindness, guidance and encouragement.

I deem it my privilege to extol my profound etiquette and sincere feeling of gratitude to Dr. Rino Rappuoli, Dr. Guido Grandi, Mikkel Nissum and Dr. Paolo Costantino for their support and contribution in different levels of my project.

I wish to thank my labmates in Protein Biochemistry Unit: Vincenzo Nardi-Dei, Marco Biancucci, Sara Marchi, Claudia Facciotti, Carlo Zambonelli, Paola Lo Surdo, Elena Cartocci, Mariangela Del Vecchio, Werner Pansegrau and Eva Grassi. I could not complete my work without invaluable friendly assistance.

I greatly value the friendship of Valentina Dimitrovska and I deeply appreciate her belief in me. Her support and care helped me overcome setbacks and stay focused on my doctoral program.

My special gratitude is extended to my precious family, to my parents, Dr. Santosh Kumar Dubey and Usha Dubey, my brothers Vivek and Vishal Dubey. Their love provided me inspiration and was my driving force. I owe them everything and wish I could show them just how much I love and appreciate them.

Apart from being my intellectual sounding board and partner, my husband Dr. Ravi. P. N Mishra has been my emotional base when living and working in Italy. I have written this dissertation with him by my side, bantering and exchanging ideas on a daily basis. His love and encouragement allowed me to finish this journey.

Finally, I would like to dedicate this work to my lost grandpa Late. Brij Bihari Dubey, who left us too soon. I hope that this work makes you proud.

Above all, this is a tribute to those who remains hidden but have always sprinkled blessing of success.

TABLE OF CONTENTS

LIST OF TABLES

Page No.

LIST OF FIGURES

Page No.

LIST OF ABBREVIATIONS

agr	Accessory gene regulator
Aur	Aureolysin
BEH	Bridged Ethyl Hybrid
BT	Bis-Tris
CDM	Chemically Defined Medium
$CaCl_2$	Calcium chloride
cDNA	Complementary DNA
CF	Cystic Fibrosis
CFU	Colony Forming Unit(s)
ClfA	Clumping Factor A
ClfB	Clumping Factor B
CO_2	Carbon Dioxide
Ct	Cycle threshold
DSF	Differential Scanning Fluorimetry
ECM	Extracellular Matrix
EDTA	Ethylenediaminetetraacetic acid
EpiP	Epidermin leader peptide processing serine protease P
ESI	Electrospray Ionisation
FnBPA	Fibronectin-binding protein protein A
FnBPB	Fibronectin-binding protein protein B
HMW	High-Molecular-Weight
HIV	Human immunodeficiency virus
IEX	Ion Exchange Chromatography
IgG	Immunoglobulin G
IMAC	Immobilized Metal Chelating Chromatography
I-PCR	Insert Polymerase Chain Reaction
IPTG	Isopropyl-β-D-ThioGalactopyranoside)
Kbp	Kilobase pairs
kDa	kilodalton(s)
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LDS	Lithium Dodecyl Sulfate
LPXTG protein	protein with an amino acid stretch LPXTG

MMPs	Mammalian matrix metalloproteinases
MRSA	Methicillin Resistant *Staphylococcus aureus*
MS	Mass spectrometry
MSCRAAMs	Microbial Surface Components Recognizing Adhesive Matrix Molecules
MSSA	Methicillin Sensitive *Staphylococcus aureus*
MWCO	Molecular Weight Cut-off
NCBI	National Center for Biotechnology Information
NCBInr	National Center for Biotechnology Information non-redundant
NH_3	Ammonia
OD	Optical Density
ORFs	Open Reading Frames
PAC	Prespotted AnchorChip
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PMF	Peptide Mass Fingerprinting
qRT-PCR	Quantitative real-time PCR
rpm	revolution per minute
sarA	Staphylococcal accessory regulator
scp	Staphylococcal cysteine protease
ScpA	Staphylococcal enzymes staphopain A
SspB	Staphylococcal enzymes staphopain B
SCV	Small Colony Variant
Sdr	Serine Aspartate Repeat
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SE-UPLC	Size Exclusion Ultra Performance Liquid Chromatography
ssp	Staphylococcal serine protease
Δ*spa* Newman	Protein A deficient Newman strain of *S. aureus*
TCA	Tri chloro acetic acid
TSA	Tryptic Soy Agar
TSB	Tryptic Soy Broth
V-PCR	Vector Polymerase Chain Reaction
WB	Western blot

CHAPTER 1

Characterization of the *Staphylococcus aureus* bone sialoprotein-binding protein SdrE

1. INTRODUCTION

1.1: Overview of *S. aureus*

Staphylococcus aureus is an important opportunistic Gram-positive human pathogen that causes a considerable burden of disease characterized by a spectrum of illnesses from mild skin infections to life-threatening diseases such as sepsis, pneumonia, endocarditis, and osteomyelitis (Lowy 1998) (Figure 1.1). *S. aureus* is particularly notorious for causing hospital-associated infections, which are often, complicated by the fact that many hospital strains are resistant to antibiotics, most notably methicillin (Lowy 2003). The burden of staphylococcal disease is increasing due to the ability of *S. aureus* to acquire resistance to various antibiotics including methicillin and vancomycin (Miyafusa, Caaveiro et al. 2012). In fact, methicillin-resistant *S. aureus* (MRSA) has been recognized as a major cause of infection in healthcare settings and community environments (Brumfitt and Hamilton-Miller 1989). For example, in 2005 it was estimated that invasive MRSA infections occurred at a rate of >30/100,000 US subjects, and caused over 18,000 deaths in the USA alone (Klevens, Morrison et al. 2007; Mariotti, Malito et al. 2013). The alarming increase in multi-antibiotic resistance of *S. aureus* together with the wide variety and severity of staphylococcal infections pose a threat to public health and challenge our ability to control the disease, in particular due to the lack of medical treatments alternative to antibiotics (Maskalyk 2002; Shorr 2007). Indeed, although several vaccine candidates have been proposed and some of them have been tested in clinical trials using both active and passive immunization modalities (Patti 2011), an effective vaccine is still missing.

Reasons behind the lack of efficacious vaccines include: 1) First and foremost, protective immunity against *S. aureus* is not completely understood, 2) *S. aureus* has multiple virulence factors, including hemolysins, toxins, and superantigens (Lowy 1998), 3) *S. aureus* infected patients present with a very broad range of diseases, which means that vaccine development must focus on preventing a wide spectrum of disease presentations, 4) *S. aureus* has a much more extensive array of pathogenicity factors that neutralize the host immune responses than the other bacterial pathogens, probably because it lives with us as normal flora (Lowy 1998).

1.2: *S. aureus*-associated diseases

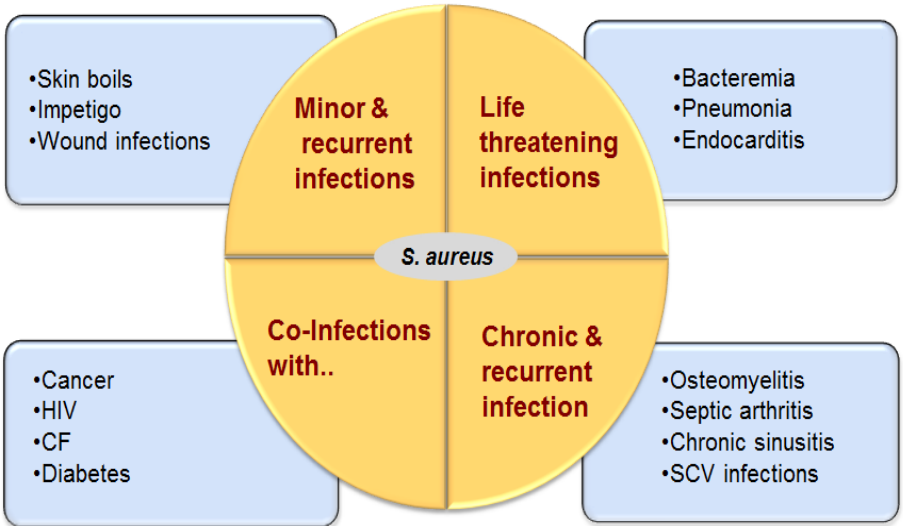

Figure 1.1: *S. aureus* causes a large array of different diseases. *S. aureus* causes a variety of cutaneous infections most common are skin boils, impetigo and wound infections. Moreover, *S. aureus* also causes life threatening disease like bacteremia which may be complicated by disease like endocarditis and pneumonia. Comorbid conditions are frequently seen in association with *S. aureus* bacteremia and that increase the risk of complications with chronic and recurrent infections like osteomyelitis, septic arthritis etc. and also include co-infections with diabetes, HIV infection, and cancer.

1.3: *S. aureus* pathogenesis

1.3.1: Cell-wall associated and secreted virulence factors

The armamentarium of virulence factors of *S. aureus* is extensive, with both structural and secreted products (proteins) playing a role in the pathogenesis of infection. Surface proteins are predominantly synthesized during the exponential growth phase, and the secreted proteins are synthesized during the stationary phase (Figure 1.2). These components and products have overlapping roles and can act either in concert or alone. Moreover, *S. aureus* is able to produce a wide range of toxins showing a deleterious effect on cell integrity and functions. Most of these factors (e.g., toxic shock syndrome toxin-1, exfoliatin toxins A and B, Panton-Valentine leukocidin, enterotoxins, and hemolysins) contribute to the virulence of clinical isolates in the context of acute infections (Francois, Scherl et al. 2007) (Figure 1.2).

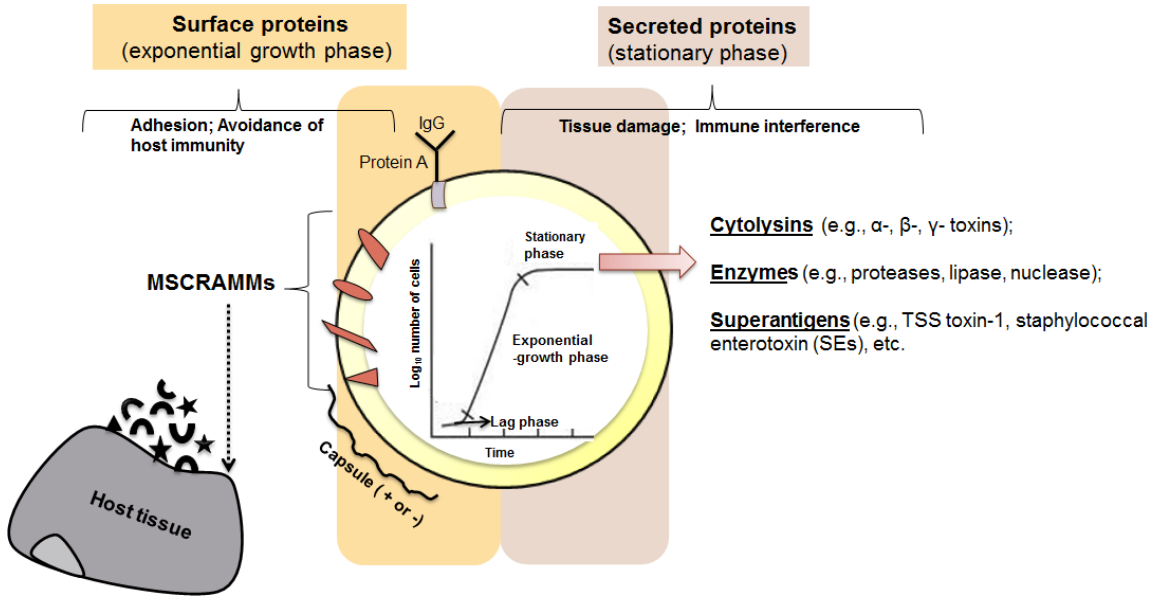

Figure 1.2: Pathogenic factors of *S. aureus* with surface and secreted products (proteins) both playing roles as virulence factors. *S. aureus* surface and secreted proteins during different phases of bacterial growth. The synthesis of many of these proteins is dependent on the growth phase, as shown by the graph. In lag phase, bacteria initiate an infection, then enter exponential phase where they multiply and synthesize surface proteins and essential proteins for growth, cell division and adhesion. Several specific cell-wall associated surface adhesins i.e. MSCRAMMs are expressed during exponential phase on the surface of *S. aureus*, which interact with a number of host proteins such as fibronectin, fibrinogen, collagen, vitronectin and laminin. Moreover, group of cytotoxins, enzymes and superantigens are secreted in the stationary phase that are known to be involved in tissue damage or have potent effects on cells of the immune system and inhibit host immune responses to *S. aureus.*

1.3.2: Microbial Surface Components Recognizing Adhesive Matrix Molecules (MSCRAMMs) protein family and serine-aspartate (Sdr) proteins

In addition to these excreted compounds, like other Gram-positive bacteria, *S. aureus* also employ an array of MSCRAMMs that are used to attach to host cells and other surfaces, and mediate pathogenesis. Many of these adhesins are anchored covalently to the cell-wall by the action of cysteine transpeptidase enzymes known as sortases (Marraffini, DeDent et al. 2006). Moreover, the role of MSCRAMM family, e.g., clumping factor ClfA and ClfB, fibronectin-binding protein protein A and B (FnBPA and B), and Sdr proteins, which allow *S. aureus* to

adhere to host tissue and thereby trigger colonization or infection has been reported (Rasmussen, Fowler et al. 2011). Furthermore, certain MSCRAMMs also mediate host cell internalization in order to escape host defense and antibacterial agents (Sinha, Francois et al. 2000; Edwards, Potts et al. 2010). A total of 21 MSCRAMMs have been identified but many host protein(s) to which these MSCRAMMs bind remain unknown (Roche, Massey et al. 2003). Importantly, Sdr group of virulence genes, in particular, SdrE have been shown to stimulate platelet activation and aggregation leading to thrombus formation (Miajlovic, Loughman et al. 2007). Moreover, it has been recently reported that *S. aureus* binds factor H (fH) via the surface protein SdrE and that fH remains functionally active when bound to recombinant SdrE (Sharp, Echague et al. 2012).

1.4: Structural organization of Sdr protein family

The Sdr proteins i.e. SdrC, SdrD and SdrE of *S. aureus* are members of the MSCRAMMs family that are encoded by the tandemly arrayed *sdrC*, *sdrD*, and *sdrE* genes, of approximately 2.8, 3.9, and 3.5 kbp, respectively, located in the *sdr* locus (Josefsson, O'Connell et al. 1998). These Sdr proteins are characterized by the presence of a R region containing various numbers of the serine-aspartate dipeptides encoded by DNA repeats in the 3′ region of the *sdr* genes. These SD repeats had earlier been found in the *S. aureus* fibrinogen-binding clumping factors, ClfA and ClfB. The putative Sdr proteins have both organizational and sequence similarity to ClfA and ClfB (Figure 1.3), wherein a signal peptide (S) is followed by an A domain (CnaA), which is similar in size among the different members of the Sdr protein family. However, they are not closely related, and show only 20 to 30% identical amino acid residues. The only conserved sequence in the A-region is the consensus motif TYTFTDYVD. The Sdr proteins differ from ClfA and ClfB by having two to five additional 110-113 residue repeated sequences (B-motifs in CnaB region) located between region A and the R-region. The Sdr proteins have two, three, or five additional 110- to 113 residue sequences (B repeat) that are tandemly repeated in SdrC, SdrE, and SdrD, respectively. Each B-repeat contains a consensus Ca^{2+}-binding EF-hand loop normally found in eukaryotic proteins. Indeed, the structural integrity of recombinant SdrD protein comprising the five B-repeats (D1-D5) was shown to be calcium (Ca^{2+}) dependent. The B-motifs in all Sdr proteins are followed by segments composed of the dipeptide serine-aspartate repeats (R-region). Moreover, the C-terminal end (region M) of the proteins is involved in anchoring the proteins to the bacterial cell wall.

It has been reported that the first B motif (CnaBC1) of SdrC, three central B motifs (CnaBD2, CnaBD3, and CnaBD4) of SdrD and the middle B motif (CnaBE2) of SdrE shows 65-85% residue identity. However, the first B motifs of SdrD (CnaBD1) and SdrE (CnaBE1) were reported to be quite different from each other, with only 42% residue identity, but with their respective neighbors 48 and 52% residue identity reported. Moreover, the B motifs lying adjacent to region R, i.e. CnaBC2, CnaBD5, and CnaBE3 in each protein are highly related with 95-96 % residue identity (Josefsson, O'Connell et al. 1998).

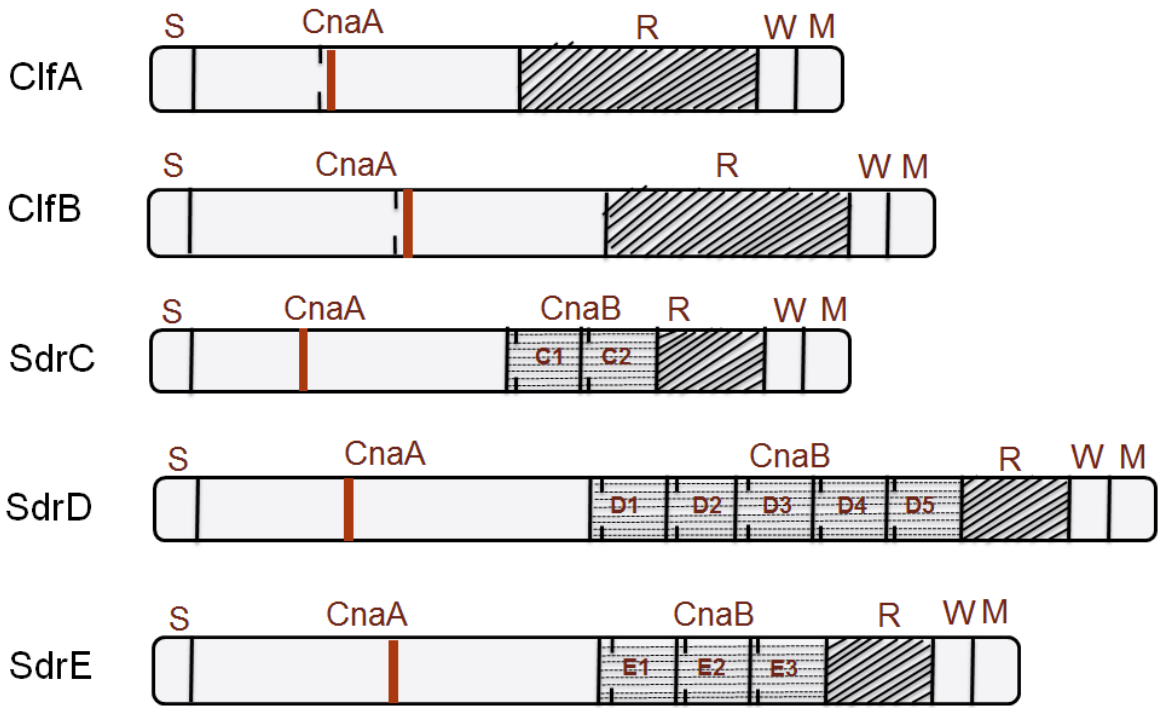

Figure 1.3: Structural organization of proteins in the SD-repeat gene family of *S. aureus* Newman. In CnaA region the thick brown line represents the TYTFTDWD motif. In the CnaA region of ClfA, ClfB and CnaB region of Sdr proteins, the thin broken black lines represent an EF-hand loop. The abbreviations are as follows: S: signal sequence; CnaA: putative ligand-binding A region; CnaB: B repeats; R: serine aspartate dipeptide repeats; W: short wall-spanning region; M: membrane-spanning segment. The LPXTG-motif occurs between domains W and M.

1.5: Reverse vaccinology and vaccine candidate identification

After the publication of the first bacterial genome in the year 1995 (Fleishmann, Mor et al. 1995), it became clear that availability of the genomic sequence of pathogens was an invaluable source of information for vaccine research. In fact, only five years later, a new antigen identification approach, named reverse vaccinology, was applied to MenB (Pizza, Scarlato et al. 2000). The approach was termed reverse vaccinology because antigens were selected prior to experimental testing (Rappuoli 2000). The idea behind the method is to mine the pathogens genome with bioinformatic algorithms to identify Open Reading Frames (ORFs) coding for proteins predicted to be exposed on the surface of the pathogen or to be secreted in the extracellular milieu. The rationale of this selection relies on the assumption that surface and secreted factors are exposed to the host's immune system and therefore they are potential vaccine targets (Figure 1.4).

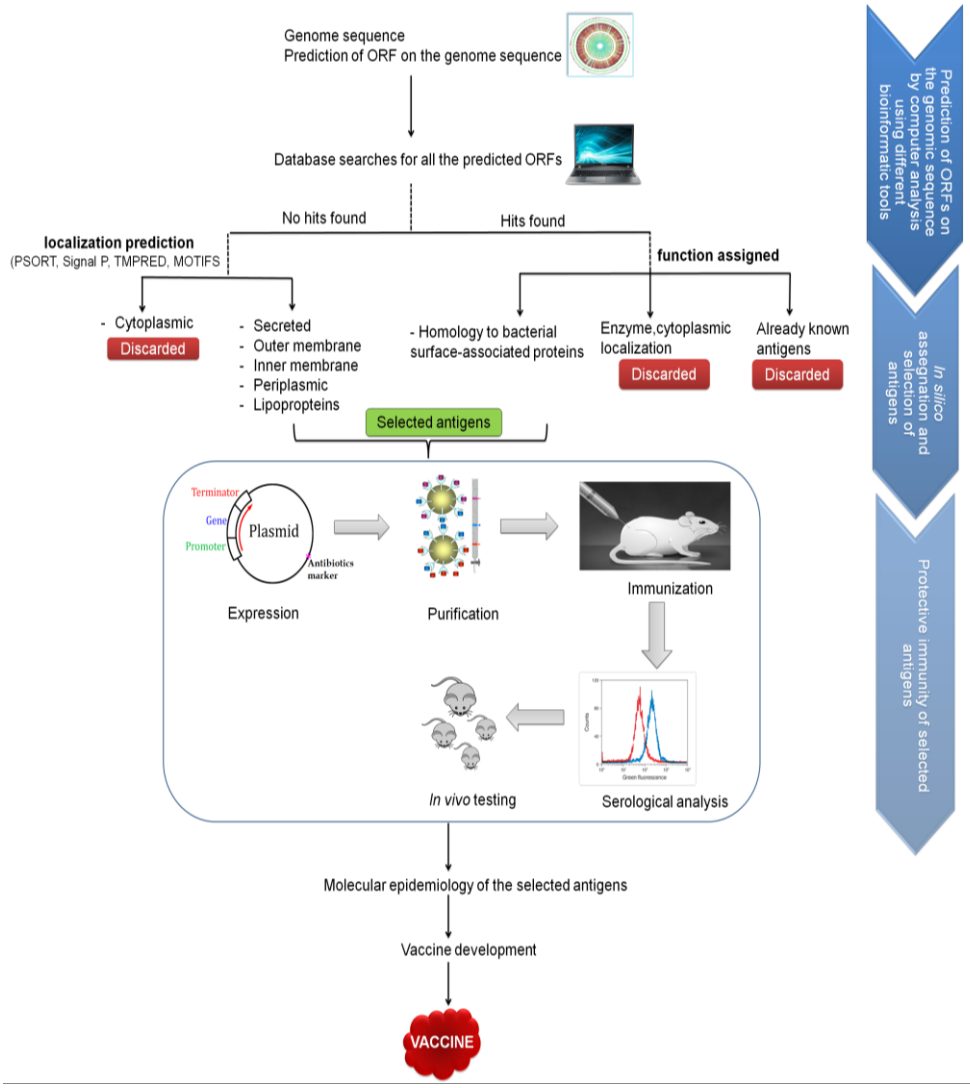

Figure 1.4: Flow chart summarizing the pathway of vaccine development starting from reverse vaccinology. Starting from the genome sequence, computer analysis of the whole genome identifies the genes coding for predicted antigens and eliminates antigens with homologies to human proteins. Then the identified antigens are screened for expression by the pathogen and for immunogenicity during infection. The selected antigens are further used to immunize animals and test whether immunization induces a protective response. Protective antigens are tested for their presence and conservation in a collection of strains representative of the species (molecular epidemiology). Finally, selected antigens are manufactured in large scale for clinical trials.

1.6: Proteomics approach of vaccine candidate identification

Furthermore, proteomics is added as further selection criteria for vaccine antigen. Proteomics is of critical importance for vaccine research because it investigates the actual protein expression in living cells and is not based on predictive assumptions as *in silico* analysis. New proteomics techniques can identify surface proteins and discriminate them from cytoplasmic ones. Such data play an essential role in vaccine candidate selection because they can indicate which antigens are more exposed on the bacterial surface and hence accessible to the immune system. An important point for the identification of vaccine candidates by proteomic techniques is the isolation of surface proteins without contamination from other cellular fractions. A new approach that allows fast and consistent identification of proteins that are expressed on the bacterial surface has been recently published (Rodriguez-Ortega, Norais et al. 2006). The technique, consisting of the surface digestion of live bacteria with different proteases and analysis by mass spectrometry, identifies the so-called "Surfome" (Figure 1.5). In addition to surface exposed antigens, secreted proteins and toxins are also viable vaccine candidates and in order to identify secreted factors (the ''secretome''), the same techniques mentioned above are applied for the analysis of bacterial culture supernatants (Ravipaty and Reilly 2010).

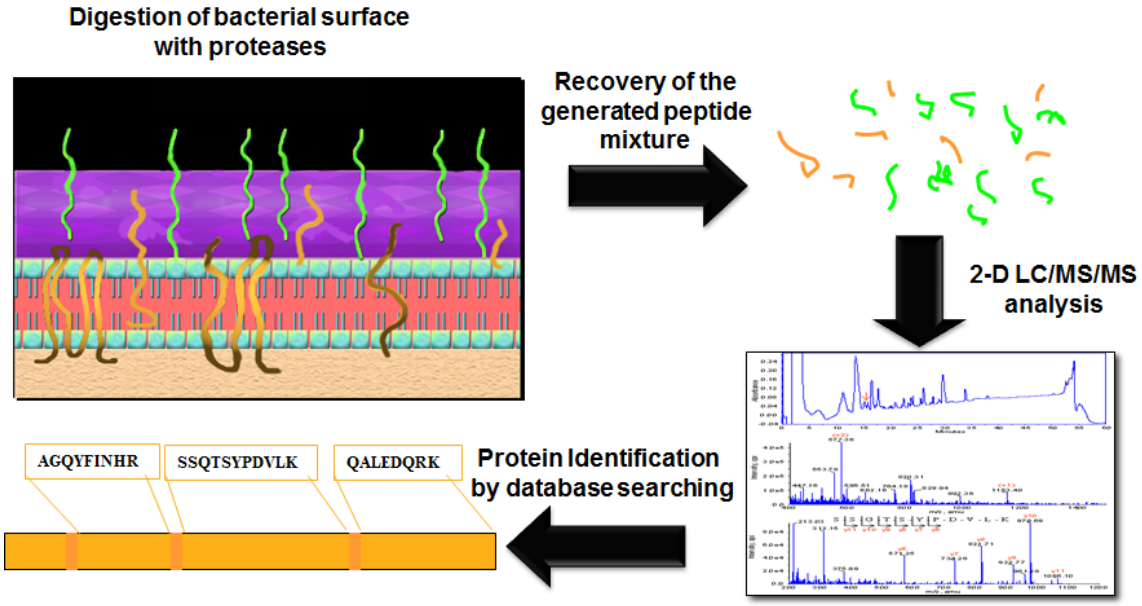

Figure 1.5: Representation of the proteomics strategy used to identify surface-exposed proteins. Peptides and polypeptides released into the supernatant by surface digestion are directly analyzed by LC-MS/MS. MS/MS spectra are then searched against a database containing protein sequence data against NCBInr database for protein identification.

In an attempt to develop an *S. aureus* vaccine, we have been applying a reverse vaccinology approach mainly based on *in silico* screening and proteomics to select surface-exposed proteins. The combination of these two different strategies led to the identification of several vaccine candidates. Among these antigens we found SdrE, a LPXTG-motif containing cell-wall anchored protein. Interestingly, we detected the protein only using harsh conditions suggesting the presence of some stability features (e.g., isopeptide bond) in the protein.

The presence of intermolecular isopeptide bond is well known fact in gram positive bacteria, which are formed between pilin subunits as covalently linked polymers through the action of cysteine transpeptidase enzymes called sortases. However, a new type of intramolecular crosslink has been recently discovered in gram positive bacteria (Kang, Coulibaly et al. 2007). These self-generated isopeptide bonds between side chains of lysine and asparagine/aspartic acid residues have emerged as a hallmark of surface proteins of Gram-positive bacteria after their discovery in the major pilin subunit RrgB of *Streptococcus pyogenes* (Kang, Coulibaly et al. 2007). Subsequently, these isopeptides were identified in proteins known to form, or associated with pili (Budzik, Marraffini et al. 2008; Kang and Baker 2009; Forsgren, Lamont et al. 2010; Izore, Contreras-Martel et al. 2010). All bacterial intramolecular isopeptides are found in β-sheet domains resembling the CnaA or CnaB folds of collagen binding protein Cna from *S. aureus* (Hagan, Bjornsson et al. 2010). Moreover, CnaA and CnaB domains are predicted to occur in thousands of bacterial surface proteins, and isopeptide bonds emerge as a very common post-translational modification underpinning Gram-positive pilus formation and protein stability. In bacterial pilus proteins, isopeptide bond formation depends on a catalytic glutamate or aspartate residue. It is thought to require location of the isopeptide triad [lysine (K), asparagine (N), catalytic carboxyl group] within the hydrophobic core (Kang, Coulibaly et al. 2007). Furthermore, since intramolecular isopeptide bonds have been found in most pilus subunits and surface proteins characterized to date (Budzik, Poor et al. 2009; Kang and Baker 2009; Izore, Contreras-Martel et al. 2010), these bonds may play a critical role in maintaining pilus and surface protein integrity in the face of severe mechanical and chemical stress while bound to host cells and thus may provide a functional mode of stabilization for cell surface proteins involved in host pathogenesis.

2: AIM OF THE STUDY

- To investigate the presence of stability feature in SdrE of *S. aureus*
- Determine the expression of SdrE in *S. aureus* clinical isolate
- Investigate the vaccine potential of the SdrE and highly conserved CnaBE3 domain of SdrE against *S. aureus* challenges
- Investigate the cross-reactive and cross-protective efficacy of CnaBE3 domain of SdrE
- Biochemical characterization of SdrE protein
- Investigate the structural organization of full-length SdrE protein

3: MATERIALS AND METHODS

3.1: Antigen identification by *in silico* analysis

In silico antigen identification was performed analyzing the *S. aureus* NCTC8325 genome by several bioinformatic algorithms as previously described (Rappuoli 2001; Bagnoli, Baudner et al. 2011; Palumbo, Fiaschi et al. 2012).

3.2: Surfome and secretome preparation of *Staphylococcus aureus*

Surfome preparation of *S. aureus* live cells was performed as previously described (Rodriguez-Ortega, Norais et al. 2006; Doro, Liberatori et al. 2009) with minor modifications. Briefly, *S. aureus* Newman strain was plated on Tryptic Soya Agar (TSA) and grown overnight. Bacterial colonies were then grown in Tryptic Soya Broth (TSB) at 37°C under agitation (250 rpm) in the presence of 5% CO_2 to optical density (λ=600 nm) of 0.4. Bacteria were harvested by centrifugation at 3,500 x g for 10 min, 4°C and washed twice with phosphate buffered saline (PBS). Bacterial pellets were re-suspended in 5 mM ammonium bicarbonate and incubated at 37°C with 20 µg of trypsin (Promega, Madison, USA). Bacterial cells were then spun down at 3,500 x g for 10 min at 4°C and the supernatant was analyzed by nanoLC-MS/MS followed by database search using Mascot (www.matrixscience.com).

Moreover, bacterial secretome preparation was also done. Secretome preparation was performed on the *S. aureus* Newman strain. Bacteria were grown as described above. After PBS washes, bacterial pellets were re-suspended and diluted in chemically defined medium (CDM) (Mickelson 1964) to optical density (λ=600 nm) of 0.05 and grown until a final optical density (λ=600 nm) of 0.4 is reached. Bacteria were removed by centrifugation at 3,500 x g for 10 min, 4°C and the supernatant was filtered through a 0.22 µm pore size filter (Millex, Millipore, Beford, U.S.A). Complete Protease Inhibitor Cocktail Tablets (Roche) was added. Proteins present in the supernatant were precipitated with 10% w/v trichloroacetic acid (TCA), 0.04% w/v sodium deoxycholate. Proteins were further re-suspended in PBS. Different aliquots were suspended on LDS buffer (Invitrogen), separated on 4-12% Bis-Tris (BT) polyacrylamide SDS-PAGE and stained with Coomassie Brilliant Blue. Protein bands of interest were subsequently analyzed after in-gel trypsin digestion by MALDI-TOF mass spectrometry, as previously described (Berlanda Scorza, Doro et al. 2008).

3.3: Protein identification by nanoLC-MS/MS

Peptides were separated by nanoLC on a NanoAcquity UPLC system (Waters) connected to a Q-ToF Premier Electro Spray Ionization (ESI) mass spectrometer equipped with a nanospray source (Waters). Samples were loaded onto a NanoAcquity 1.7 µm BEH130 C_{18} column (75

µm X 25 mm, Waters), through NanoAcquity 5 µm Symmetry® C18 trap column (180 µm X 20 mm, Waters). Peptides were eluted with a 120-min gradient of 2–40% of 98% acetonitrile, 0.1% formic acid solution at a flow rate of 250 nL/ min. The eluted peptides were subjected to an automated data-dependent acquisition, using the MassLynx software, version 4.1 (Waters), where a MS survey scan was used to automatically select multi-charged peptides over the m/z ratio range of 300–2,000 for further MS/MS fragmentation. Upto eight different components were subjected to MS/MS fragmentation at the same time. For all samples, a second nanoLC-MS/MS analysis was carried out for the selective fragmentation of mono-charged peptide species.

After data acquisition, individual MS/MS spectra were combined, smoothed and centroided using ProteinLynx, version 3.5 (Waters) to obtain the peak list file. The Mascot Daemon application (Matrixscience Ltd., London, UK) was used for the automatic submission of data files to an in-house licensed version of MASCOT, version 2.2.1, running on a local server. Protein identification was achieved by searching in a locally created database containing protein sequence data derived from the sequenced *S .aureus* strains. The MASCOT search parameters were set to (i) 1 as number of allowed missed cleavages, (ii) 0.3 Da as peptide tolerance, and (iii) 0.3 Da as MS/MS tolerance. Only significant hits were considered, as defined by the MASCOT scoring and probability system. The score thresholds for acceptance of peptide identification were ≥18 for trypsin digestion.

3.4: Cloning, expression and purification of SdrE

The *sdrE* gene was PCR amplified from *S. aureus* NCTC8325 strain and inserted into the pET-15b+ vector (Novagen). Gene coding for the protein was cloned as N-terminal 6X-histidine-tag fusion protein, expressed and purified. PCR primers were designed to amplify gene without predicted signal peptide coding sequences and C-terminal cell-wall anchoring LPXTG-motif. PCR fragments were cloned by using the Polymerase Incomplete Primer Extension (PIPE) method, developed by GNF (Genomics Institute of the Novartis Research Foundation, San Diego, CA, USA) (Klock and Lesley 2009). SdrE cloning operations were achieved by transforming HK100 competent cells with PCR products (I-PCR) immediately following amplification mixed with the V-PCR of Speed ET vector (N-term 6X-His tag) (Klock and Lesley 2009). The transformants were selected on Luria-Bertani (LB) plates supplemented with ampicillin (100 µg/ml) (Sigma). The plasmid pET-15b+-SdrE screened, were further confirmed by sequencing.

The plasmids containing the SdrE full-length sequence, thereof were transformed into expressing strain of *Escherichia coli* BL21 DE3 (Novagen) cells. Protein expression was induced by adding IPTG (isopropyl-β-D-thiogalactopyranoside) (Sigma) 1 mM final concentration to a bacterial culture (Luria-Bertani broth medium supplemented with 100 µg/ml ampicillin) at an optical density (λ=600 nm) of 0.4 to 0.5 and then growing the bacteria at 37°C for next 8 hours.

Protein purification was performed as follows; bacterial cells were harvested by centrifugation for 10 min in a Beckmann JA 81000 rotor at 8,000 x g at 4°C. Pellets (8 g) were lysed with lysostaphin (0.25 mg/ml) in 40 ml Bug Buster Reagent (Novagen) supplemented with Benzonase Nuclease (Novagen; 2.5-U/ml final concentration) and Protease Inhibitor Cocktail III (Calbiochem; 2.5 µl/ml lysate). Cells were then harvested by centrifugation (45 min at 30,000 x g, 4°C). The supernatants of these lysates were filtered through a 0.45 µm membrane. Soluble histidine fusion proteins present in the supernatants were initially purified by metal-chelating chromatography. In brief, the supernatants were applied to a 5 ml Ni^{2+}-charged His Trap chelating column (Armesham Biosciences) pre-equilibrated in 100 mM NaPPi, pH 8 buffer containing 15 mM imidazole and bound proteins were eluted with linear gradients of 0-250 mM imidazole in 100 mM NaPPi, pH 8, at a flow rate of 5 mL/min. Fractions were analyzed for protein content by determining their absorbance at 280 nm and those containing recombinant proteins were identified by SDS-PAGE, pooled and dialyzed overnight against 50 mM Tris/HCl pH 7.6. Dialyzed proteins were concentrated and further purified by ion-exchange chromatography by applying the samples to a 5 ml HiTrap Q column (Armesham Biosciences). The column was washed with 50 mM Tris-HCl, pH 7.6 buffer containing 50 mM NaCl and bound proteins were eluted by applying a step gradient from 0.2 to 1 M NaCl in increaments of 0.1 M, at a flow rate of 3 mL/min. Fractions were analyzed for protein content as described above. The pooled fractions containing SdrE were further purified by preparative gel filtration chromatography (HiLoad 26/60 Superdex 75) (GE Healthcare) in buffer 50 mM Tris-HCl, pH 7.6, with a flow rate of 1 mL/min. Protein concentration was estimated using BCA assay (Pierce). Protein purity was determined by RP-HPLC on a Vydac C4 4.6x150 mm column using acetonitrile gradient.

3.5: Preparation of *S. aureus* sub-cellular fractions

In order to evaluate the expression of SdrE in *S. aureus* in vitro, immuno-blot assay was performed. *S. aureus* Newman strain deficient for protein A (SpA) was used. Lysate fractions were obtained as follows. Bacteria were grown to stationary phase (overnight culture) in TSB

at 250 rpm, 37°C using aerated Erlenmeyer flasks. Overnight cultures were centrifuged at 4,000 rpm for 15 minute, 4°C. For preparing cell-wall fractions, pellets from 5 ml overnight culture was re-suspended in 500 µl TSM (50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 0.5 M sucrose) buffer. 50 µl of lysostaphin (5 µg/µl) was added to samples and incubated for 1 hour or until lysis at 37°C in a thermomix. DNase and protease inhibitors were then added. After lysis, sample was centrifuged at 4,000 rpm for 15 min, 4°C and supernatants containing the cell-wall fraction were used for immuno-blot. For protoplasts preparation, the pellet was re-suspended in 100 µl of SDS, then 1 µl of benzonase was added and incubated at 37°C for 15 min. Sample was boiled, the tube was centrifuged for 5 min at 3,500 x g and then the supernatant was collected and loaded on SDS-PAGE.

3.6: SDS-PAGE and immuno-blot assay

SDS-PAGE analysis was performed using Nu-Page 10% Tris-Acetate gradient gels (Invitrogen) according to the manufacturer's instructions. A Hi-Mark pre-stained high-molecular-weight protein standard (Invitrogen) was used. All protein samples were heat-denatured under reducing conditions. rSdrE and 20 µl of *S. aureus* lysate fractions (cell-wall and protoplasts) was placed into the wells of acrylamide gels and were subjected to SDS-PAGE. Gels were stained with colloidal Coomassie Brilliant Blue or processed for immuno-blotting by using standard protocols. Briefly, separated proteins were subsequently transferred onto a nitrocellulose membrane using i-blot system (Invitrogen). Membranes with the transferred proteins were incubated in PBS containing 1% Tween 20 (PBS-T) and 5% non-fat dry milk for 1 hour at room temperature followed by incubation with anti-SdrE mouse serum at 1/25000 dilution (diluted in PBS-T containing 5% milk) at room temperature. Following antisera incubation, the membranes were washed four times with PBS-T, 15 min each wash and incubated with horseradish peroxidase-conjugated goat anti-mouse antibody (Sigma) at a 1/30,000 dilution for 1 hour at room temperature. After being washed four times in PBS-T, the membrane was incubated with SuperSignal West Pico Chemiluminescent substrate (Pierce) for 3 min and exposed to X-ray film.

3.7: Enzymatic digestion of rSdrE

Protease resistant behavior of rSdrE was observed in two conditions i.e. in presence and in absence of calcium. In brief, rSdrE was dialyzed overnight against 50 mM Tris-HCl + 1 mM of $CaCl_2$, pH 7.6 and the other with 50 mM Tris-HCl, pH 7.6, with one change of buffer in both the cases. rSdrE (+/- Ca^{2+}) was further digested overnight at 37°C with sequencing grade

modified trypsin, using an enzyme/substrate ratio of 1/25 (wt/wt) in 50 mM ammonium bicarbonate, pH 8, containing 0.1% (wt/vol) rapigest (Waters). Overnight trypsin treated rSdrE (+/- Ca^{2+}) samples were further separated on 12% BT polyacrylamide SDS-PAGE and stained with coomassie brilliant blue.

3.8: Mass Spectrometry analyses

Coomassie Brilliant Blue stained overnight trypsin resistant bands of rSdrE were excised from the SDS-PAGE gel using a Pasteur pipette and destained overnight in 200 µl of 50% acetonitrile (J. T. Baker Inc.) and 50% of 50 mM ammonium bicarbonate (Fluka Chemie AG, Buchs, Switzerland) (50:50, v/v). The excised bands were then washed with 200 µl of acetonitrile. The acetonitrile was discarded, and the bands were allowed to air dry. Dried bands were digested overnight at 37°C in 12 µl of 0.012 µg/ml sequencing grade modified trypsin in 5 mM ammonium bicarbonate. After overnight digestion, 0.8 µl was directly spotted on a PAC (Prespotted AnchorChip 96, set for proteomics, Bruker Daltonics, Bremen, Germany) and air-dried. The air-dried spots were washed with 0.6 µl of a solution of 70% (vol/vol) ethanol, 0.1% (vol/vol) trifluoroacetic acid (TFA). Peptide mass spectra were recorded with a matrix-assisted laser desorption ionization–time of flight (MALDI-TOF)/TOF mass spectrometer (UltraFlex; Bruker Daltonics, Bremen, Germany). Ions generated by laser desorption at 337 nm (N_2 laser) were recorded at an acceleration of 25 kV in the reflector mode. About 200 single spectra were accumulated for improving the signal/noise ratio and analyzed by Flex Analysis (version 2.4; Bruker Daltonics). External calibration was performed using standard peptides pre-spotted on the target. Peptide identification was performed using BioTools and Sequence Editor 3.0 (Bruker Daltonics). Protein identification was carried from the generated peak list using the Mascot program (Mascot server version 2.2.01, Matrix Science). Mascot was run on a public database (National Center for Biotechnology Information non-redundant (NCBInr).

3.9: Size-Exclusion Ultra Performance Liquid Chromatography (SE-UPLC) for the purification of rSdrE major trypsin resistant fragment

To investigate which peptide the overnight trypsin resistant fragment of rSdrE comprises of, the resistant fragment was purified using SE-UPLC. In brief, overnight trypsin digested SdrE sample were purified using Bridged Ethyl Hybrid 200 (BEH200) column in UPLC (Waters) at flow rate of 0.5 mL/min with 20 mM phosphate, pH 8.0 buffer. Sample (60 µl) of overnight trypsin digested rSdrE were loaded on to the column, and peaks were assigned at A280.

Apparent molecular masses were determined from a standard curve of thyroglobulin (670 kDa), bovine-gamma-globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa) and cytochrome c (12.3 kDa). Based on the chromatogram different fractions were collected, subjected to 12% BT polyacrylamide SDS-PAGE and stained with coomassie brilliant blue and the fraction with the purified major resistant fragment were retained and further used for N-terminal sequencing by Edman Degradation.

3.10: Amino-terminal sequencing of SdrE trypsin resistant fragment by Edman degradation

To confirm the identity of the resistant fragment of rSdrE, the SE-UPLC purified resistant fragment was subjected to amino-terminal sequence analyses by Edman degradation (Lindenthal and Elsinghorst 1999). Amino-terminal sequence analyses were performed on an Agilent G1000A series protein sequencer following manufacturer's protocol.

3.11: Cloning, expression and purification of CnaBE3 domain of SdrE

Based on the MALDI-TOF mass spectrometry and N- terminal sequencing result of rSdrE resistant fragment, CnaBE3 domain of SdrE was cloned as C-terminal 6X-histidine tag using the Polymerase Incomplete Primer Extension (PIPE) method (Klock and Lesley 2009). The *E. coli* strain expressing recombinant CnaBE3 was propagated with the EnPresso Tablet Cultivation Set (BioSilta) (Panula-Perala, Siurkus et al. 2008; Zhou, Szeker et al. 2013). CnaBE3 was purified in single step using metal-chelating chromatography. In brief, the supernatants were applied to a 5 ml Ni^{2+}-charged His Trap chelating column and bound proteins were eluted with 200 ml linear gradients of 0-250 mM imidazole in 300 mM NaH_2PO_4 + 50 mM NaCl, pH 8.0, at a flow rate of 5 mL/min. Fractions were analyzed for protein content by determining their absorbance at 280 nm and those containing recombinant proteins were identified by SDS-PAGE. Protein concentration was estimated using BCA assay (Pierce).

3.12: Enzymatic digestion of rCnaBE3

rCnaBE3 was digested with sequencing grade modified trypsin, following the same protocol as used for rSdrE. rCnaBE3 samples overnight incubated with trypsin was further separated on 12% BT polyacrylamide SDS-PAGE and and stained with Coomassie Brilliant Blue.

3.13: Amino-terminal sequencing of rCnaBE3 trypsin resistant fragment by Edman degradation

Trypsin digestion of rCnaBE3 resulted in the appearance of one major resistant fragment at molecular weight of 12 kDa. In order to further confirm the identity of the overnight trypsin resistant fragment of rCnaBE3, the resistant fragments of rCnaBE3 were subjected to amino-terminal sequence analyses. Amino-terminal sequence analyses were performed on an Agilent G1000A series protein sequencer following manufacturer's protocol as described for SdrE in 3.10.

3.14: Intact mass spectrometry of rCnaBE3

To look for the mass difference between theoretical and experimental mass of rCnaBE3, intact mass spectrometry of rCnaB3 was done. rCnaBE3 were diluted in 0.1% formic acid. The acidified protein solutions were loaded onto a Protein Microtrap cartridge (from 60 to 100 pmols), desalted for 2 min with 0.1% formic acid at a flow rate of 200 mL/min and eluted directly into the mass spectrometer using a step gradient of acetonitrile (55% acetonitrile, 0.1% formic acid). Spectra were acquired in positive mode on a SynaptG2 HDMS mass spectrometer equipped with a Z-spray ESI source. The quadrupole profile was optimized to ensure the best transmission of all ions generated during the ionization process.

3.15: Cloning and site directed mutagenesis of five asparagines at different positions in CnaBE3 domain of SdrE

We designed six point mutation construct in the CnaBE3 domain of SdrE, where asparagine present at six different position in the CnaBE3 domain of SdrE were substituted with alanine and are named as CnaBE3 mutant 1 (N23A), CnaBE3 mutant 2 (N49A), CnaBE3 mutant 3 (N63A), CnaBE3 mutant 4 (N70A), CnaBE3 mutant 5 (N92A) and CnaBE3 mutant 6 (N119A) (Figure 1.6) and all six asparagine point mutation construct with 6X- histidine tag at C-terminal. However, till date we are having only five CnaBE3 mutants (1 to 5) ready for further analysis. The primers used for five point mutation construct are listed in Table 1.1. The Stratagene QuikChange™ site- directed mutagenesis kit was used to construct all the five mutants according to the procedure outlined in the manufacture's technical manual. Transformants were screened for recombinant plasmids, and then mutation constructs were identified and further confirmed by DNA sequencing.

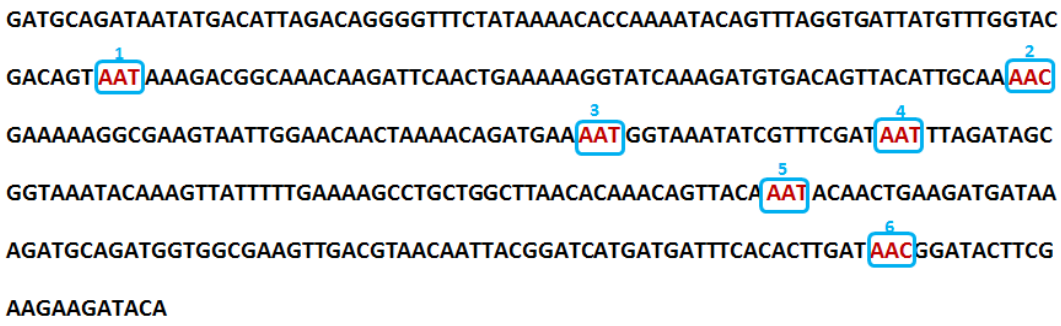

Figure 1.6: Nucleotide sequence of CnaBE3 domain of SdrE with asparagine point mutation site. Nucleotide sequence of CnaBE3 domain of SdrE. Blue boxes in the sequence represent the different site of point mutation where asparagine is substituted with alanine (1-6).

Table 1.1: PCR primers designed to amplify corresponding gene

Gene	Primer	Nucleotide sequence
sdrE	sdrEF	5'-CTGTACTTCCAGGGCGCTGAAAACACTAGTACAGAAAATGCAAAACAAG
	sdrER	5'-AATTAAGTCGCGTTATGCTTTTGCTTTATTGTGATGGTCTTTAGTAG
cnaBE3	cnaBE3F	5'-GAAGGAGATATACATATGGATGCAGATAATATGACATTAGAC
	cnaBE3R	5'-GTGGTGGTGGTGGTGTGTATCTTCTTCGAAGTATCCGTT
cnaBE3 mutant 1 (N23A)	cnaBE3 (N23A)F	5'-GACAGTGCTAAAGACGGCAAACAA
	cnaBE3 (N23A)R	5'-GTCTTTAGCACTGTCGTACCAAAC
cnaBE3 mutant 2 (N49A)	cnaBE3 (N49A)F	5'-TTGCAAGCGGAAAAAGGCGAAGTA
	cnaBE3 (N49A)R	5'-TTTTTCCGCTTGCAATGTAACTGT
cnaBE3 mutant 3 (N63A)	cnaBE3 (N63A)F	5'-GATGAAGCGGGTAAATATCGTTTC
	cnaBE3 (N63A)R	5'-TTTACCCGCTTCATCTGTTTTAGT

cnaBE3 mutant 4 (N70A)	cnaBE3 (N70A)F	5'-TTCGATGCGTTAGATAGCGGTAAA
	cnaBE3 (N70A)R	5'-ATCTAACGCATCGAAACGATATTT
cnaBE3 mutant 5 (N92A)	cnaBE3 (N92A)F	5'-GTTACAGCGACAACTGAAGATGAT
	cnaBE3 (N92A)R	5'-AGTTGTCGCTGTAACTGTTTGTGT

3.16: Expression and purification of five asparagine point mutation construct in CnaBE3 domain of SdrE (CnaBE3 mutant 1-5)

All five CnaBE3 mutation construct were transformed into *E. coli* BL21 (DE3) expression cells. The *E. coli* strain expressing recombinant five CnaBE3 asparagine mutant construct were propagated with the EnPresso Tablet Cultivation Set (BioSilta) (Panula-Perala, Siurkus et al. 2008; Zhou, Szeker et al. 2013). Five (1-5) CnaBE3 point mutation constructs were purified using single step by metal-chelating chromatography according to procedures described above for the purification of wild type CnaBE3 protein.

3.17: Enzymatic digestion of five point mutation CnaBE3 recombinant proteins

All five point mutation purified recombinant proteins i.e. CnaBE3 mutant 1 (N23A), CnaBE3 mutant 2 (N49A), CnaBE3 mutant 3 (N63A), CnaBE3 mutant 4 (N70A), CnaBE3 mutant 5 (N92A) were digested with sequencing grade modified trypsin, using an enzyme/substrate ratio of 1/25 (wt/wt) in 50 mM ammonium bicarbonate, pH 8, containing 0.1% (wt/vol) rapigest (Waters) overnight at 37°C.

3.18: Active immunization of mice and generation of polyclonal antibodies

Polyclonal mouse antibodies have been generated and used for western blot (WB) studies. Five-week-old CD1 mice were immunized intraperitoneally with a prime-booster injection of 20 µg purified recombinant SdrE and CnaBE3 separately adsorbed to aluminum hydroxide adjuvant (alum, 2 mg/ml) in 14' day interval (Mishra, Mariotti et al. 2012). Control mice received equal amounts of PBS and alum adjuvant. Animals were bled immediately prior to

the first immunization and 23 days thereafter, and sera were examined for IgG antibodies directed against purified SdrE and CnaBE3 using the Luminex technology.

3.19: Sub-cellular fraction preparation in ten different *S. aureus* strains and immuno-blot assay using anti-CnaBE3 mouse sera

Bacterial cell-wall were obtained as described previously (Sitkiewicz, Babiak et al. 2011). Ten *S. aureus* strains namely NCTC8325, Newman, MSSA476, MW2, N315, Mu50, Mu3, USA300-FPR3757, MRSA252 and TW20 were grown to mid-log phase in TSB supplemented with 5 mM $CaCl_2$ to an optical density (λ=600 nm) of 0.6. Bacterial cell-wall were harvested by centrifugation at 4,000 rpm for 15 minute, 4°C. Cells were then washed in PBS once and resuspended in 100 µl lysis buffer (50 mM Tris-HCl, 20 mM $MgCl_2$, pH 7.5) supplemented with 30% (w/v) raffinose and 40 µl/ml EDTA-free protease inhibitors cocktail. 10 µl of lysostaphin (200 µg/ml) was added to samples and incubated at 37°C for 1 hour. After lysis, samples were boiled for 10 minute with LDS sample buffer and sample reducing agent (Life Technologies) and separated on 3-8% Tris-Acetate gel. Electrophoretically separated protein samples were transferred to nitrocellulose membranes with i-Blot transfer system. Membranes were blocked for 2 hours at room temperature in PBS-T with 10% skimmed milk (Bio-Rad), followed by incubation with anti-CnaBE3 mouse serum at 1/1,000 dilution (diluted in PBS-T containing 5% milk) at room temperature. Following antisera incubation, the membranes were washed four times with PBS-T, 15 min each wash and then incubated with horseradish peroxidase-conjugated goat anti-mouse antibody at a 1/5,000 dilution in 10% skimmed milk for 1 hour at room temperature. After being washed four times in PBS-T, the membrane was incubated with SuperSignal West Pico Chemiluminescent substrate for 3 min and exposed to X-ray film.

3.20: Immunogenicity assay

SdrE and CnaBE3 antibody titers present in sera of immunized mice were measured by Luminex technology (Luminex® 200 TM). The protein was covalently conjugated to the free carboxyl groups of microspheres using a N-hydroxysulfosuccinimide-enhanced carbodiimide-mediated conjugation chemistry. Antigen specific antibodies were revealed by phycoerythrin-labelled secondary antibodies. The assay read-out is a measure of fluorescence intensity expressed as arbitrary Relative Luminex Units (RLU/mL).

3.21: Murine abscess model for protective efficacy of SdrE and CnaBE3

In order to look for the vaccine efficacy of SdrE and CnaBE3, immunization with proteins followed by *S. aureus* challenge was done. Immunized animals were challenged on day 24 by intravenous injection of a sub-lethal dose of *S. aureus* Newman strain. TSB cultures of *S. aureus* were centrifuged, washed twice and diluted in PBS before challenge. Further dilutions were needed for the desired inoculum, which was experimentally verified by agar plating and colony counting. Mice were infected with approximately 2 to 6 $x10^7$ CFU of *S. aureus* Newman strain. On day 28, mice were euthanized and kidneys were removed and homogenized in 1% Triton X-100, aliquots diluted and plated on agar media for triplicate determination of colony forming units (CFU). Data were analyzed by Mann Whitney U test.

3.22: Murine abscess model for cross-protective efficacy of CnaBE3 of SdrE

In order to look for the cross-protective efficacy of the highly conserved CnaBE3 domain of SdrE, mouse were immunized with CnaBE3. CnaBE3 immunized animals were challenged on day 24 by intravenous injection of a sub-lethal dose (~ 2 to 6 $x10^7$ CFU/ml) of *S. aureus* NCTC8325 strain, which is naturally devoid for *sdrE* gene. On day 28, mice were euthanized, kidneys were removed and bacterial load was measured. At least three independent experiments were carried out under the same conditions to assess reproducibility. Data were analyzed by the Mann-Whitney U-test.

3.23: Differential scanning fluorimetry (DSF)

DSF experiments were performed using thin wall PCR plates (Axigen). rSdrE was used at a final concentration of 10 μM; both $CaCl_2$ (dissolved in water) and EDTA were used up to 1 mM. The SYPRO orange dye 5000X (Invitrogen) was used at the final concentration of 5X in each well. The reaction mixtures were 40 μl in 50 mM Tris-HCl buffer, pH 7.6. Fluorescence intensities were monitored using an Mx3005 RT-PCR instrument (Stratagene) using the FAM (492 nm) and ROX (610nm) filters for excitation and emission, respectively. Samples were heated from 25°C to 95°C at scan rate of 1°C/min. Tm values were extrapolated by fitting the raw data to a Boltzmann model using GraphPad Prism 5.0 software. Each experiment was performed atleast in triplicate.

3.24: Design of SdrE construct for structural study

The design of optimized protein expression construct is frequently a crucial point to increase the probability of success of crystallization. In order to do that, flexible parts of the protein like long loops or disordered regions are reduced or removed because they can interfere

negatively with crystal formation. More compact or stable domains, without flexible regions, are known to enhance the crystallization of "tricky proteins".

In addition, in the design of protein expression constructs suitable for crystallization, it is normal practice to study the related protein structures (if any) present in the Protein Data Bank (PDB), in order to understand which regions of homologous proteins have shown a propensity to crystallize and reveal well-ordered, folded structures. PDB search using full-length SdrE sequence identified the structure 1R17 (SdrG bound to Fibrinogen), 331 residues with 50% SEQID from residue 227-553. An analysis of the SdrE protein sequence was performed using several publicly available software suites e.g., Medor: http://www.vazymolo.org/MeDor/index.html, which were developed in order to identify regions of secondary structure, order and disorder (Ward, McGuffin et al. 2004; Lieutaud, Canard et al. 2008). By Medor analysis of the full-length SdrE amino acid sequence, it is likely to be secondary structure and hydrophobic clusters up to residues Phenylalanine (F-935) to aspartic acid (D-941). The Disopred server (http://bioinf.cs.ucl.ac.uk/disopred/) suggests that SdrE is disordered starting from the first residue until the residue valine (V-275), approximately 275 amino acid of the SdrE sequence. From the secondary structure analysis of SdrE, (performed using the bioinformatics servers available at http://toolkit.tuebingen.mpg.de and http://groups.csail.mit.edu/cb/paircoil2/paircoil2.html) no major COILS predicted. Phobius server predicted signal peptide from residue 1-52 (underlined) in SdrE sequence (Figure 1.7). The SDSD repeats in the sequence are highly likely to be unstructured.

Based on the different online bioinformatic algorithms, two construct of SdrE named as long (T_{212}NPK- DTSD$_{942}$) and short version (T_{251}APT- DTSD$_{942}$) of SdrE were designed, cloned in pET-15b+ vector by PIPE cloning method and checked by sequencing. Both the plasmids are sent to Northwestern University, Chicago for crystallization.

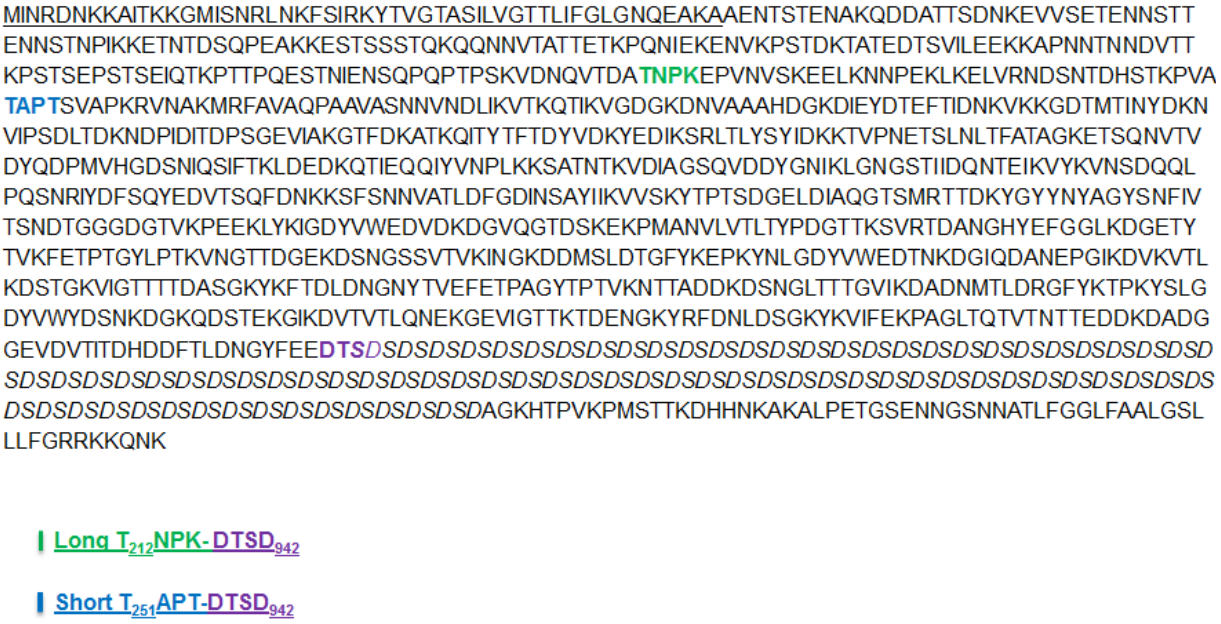

Figure 1.7: Full-length SdrE amino acid sequence indicating the construct designed for structural study. Based on different bioinformatic algorithms two construct namely long and short version of SdrE were designed. Two construct designed for SdrE starts with different N-terminal residues (green and blue), but end to same C-terminal residue (in purple). In green, is the starting residues at N-terminal for the long version of SdrE and in blue is the starting residues at N-terminal for short verion of SdrE.

4: RESULTS

4.1: *In silico* prediction and mass spectrometry identification of SdrE

SdrE was identified as a putative vaccine antigen during analysis of the *S. aureus* NCTC8325 genome by several bioinformatic algorithms as described in material and method. The *S. aureus* SdrE protein was predicted as cell-wall anchored protein due to the presence of a leader peptide and cell wall anchoring LPXTG-motif (Figure 1.3). *sdrE* gene is approximately 3.5 kbp and is close to *sdrC* and *sdrD* in sdr locus. Although *sdrE* is not highly conserved among different staphylococcal strains, is present in epidemiologically relevant strains of *S. aureus* (data not shown). In parallel with the *in silico* analysis, we performed a mass spectrometry (MS)-based study of the staphylococcal surface and secreted proteins. The surfome and secretome of *S. aureus* Newman were analyzed by nanoLC-MS/MS. SdrE was unexpectedly identified only in the secretome (data not shown). We assume that this is due to the different sample treatment in analyzing the secretome as compared to the surfome. Indeed, the secretome is treated with harsher conditions. Therefore, this observation suggested that the protein is highly stable and protease resistant. On the other hand, by immuno-blot analysis we found the protein to be expressed in the cell-wall of the epidemiologically relevant *S. aureus* Newman strain (Figure 1.8).

Moreover, it is known that in most cases the protein with the pilin subunits and LPXTG-type cell- wall sorting sequences are presumably targeted by sortases (Guttilla, Gaspar et al. 2009) and the sortase-mediated isopeptide bond cross-linking stabilizes and strengthens Gram-positive bacterial pili, that are the unique examples of covalent biological polymers. Pili are long and thin, typically 1–5 μm in length but only ~30–60 nm in width (no more than one molecule wide), but their strength and integrity are maintained by a remarkable sequence of covalent cross-links, both between the individual pilin subunits and within them. In both cases, these cross-links take the form of isopeptide bond either intermolecular or intramolecular isopeptide bond. The intermolecular isopeptide bond, as mentioned earlier is catalyzed by the action of the sortase enzymes, which join the C-terminal carboxylate of one subunit to a specific lysine side chain on the next (Hendrickx, Budzik et al. 2011). The internal cross-links, however, arise through autocatalytic, intramolecular reactions that occur spontaneously in the pilin subunits (Kang and Baker 2009; Kang and Baker 2011). In order to see if rSdrE forms a high molecular mass ladder and are involved in formation of intemolecular isopeptide bond, like some of the surface proteins with CnaA and CnaB domain involved in the formation of intermolecular isopeptide bond, immuno-blot analysis using anti-SdrE mouse serum was performed.

4.2: Expression of SdrE in cell-wall fraction of *S. aureus* Δ*spa* Newman strain by immuno-blot

Immuno-blot assay confirms the expression of SdrE in the cell-wall fraction of *S. aureus* whereas no ladder-like polymeric high-molecular-weight (HMW) was observed under the condition tested (Figure 1.8).

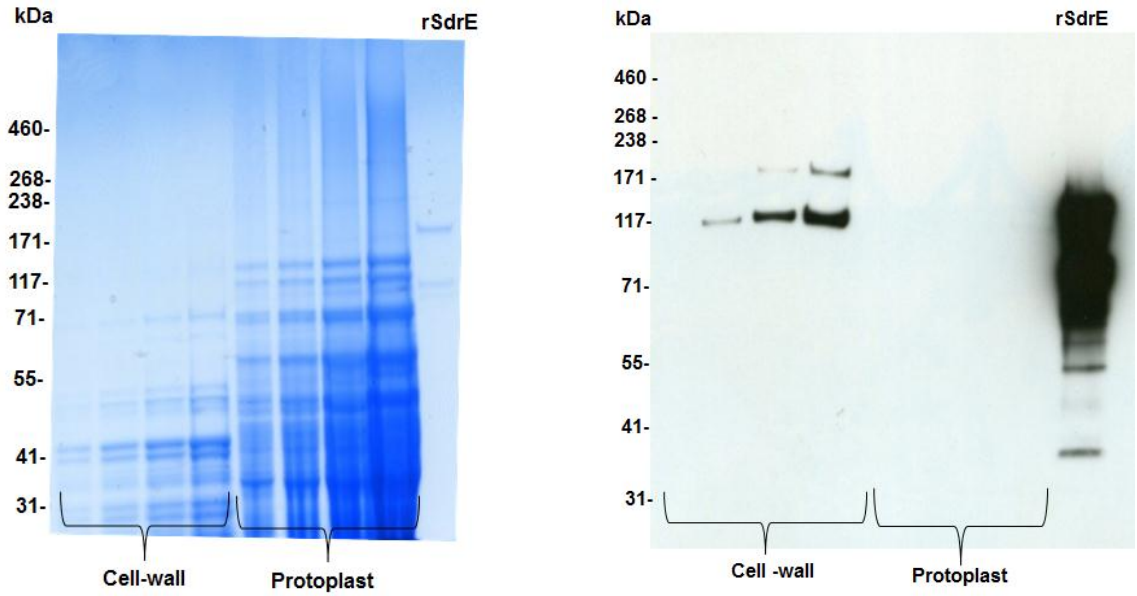

Figure 1.8: SdrE expression in *S. aureus* Δ*spa* Newman strain. (A) SDS-PAGE of cell-wall (Lane 1 to 4, with different loading volume), protoplasts fraction (Lane 5 to 8, with different loading volume) and rSdrE (Lane 9). The sizes of molecular mass standards (in kilodalton) are indicated on the left. Proteins from each fraction were separated on 10% Tris-Acetae SDS-PAGE. (B) Analysis of cell-wall fraction and protoplast of *S. aureus* strain Δ*spa* Newman by immuno- blot using anti-SdrE mouse serum. Loading volume in each lane is similar to SDS-PAGE. In the cell-wall preparation, band with a MW compatible with the rSdrE protein were visible. No immunoreactivity was detected in the lane where the protoplasts preparation was loaded. However, no ladder-like polymeric high-molecular-weight (HMW) seen, unlike some of the surface proteins with CnaA and CnaB domain and involved in the formation of intermolecular isopeptide bond in gram positive bacteria.

4.3: rSdrE shows resistance to proteolytic cleavage

The surface proteins with CnaA and CnaB domains in most of the Gram-positive bacteria have been shown to resist proteolytic cleavage, a feature associated with the presence of intramolecular isopeptide bonds (Kang and Baker 2009). We investigated whether SdrE, a surface protein containing CnaA and CnaB domains were also protease resistant and if this feature is associated with the presence of intramolecular isopeptide bond. It has been reported that because of the presence of EF-hand loop in SdrD protein, a protein belonging to Sdr protein family, shows high calcium affinity, which in turn helps in maintaining the structural stability of the protein (Josefsson, O'Connell et al. 1998). Therefore, trypsin proteolytic-cleavage reactions were performed on the purified full-length SdrE in presence and absence of 1 mM of calcium. We observed that under both the conditions rSdrE were significantly resistant to enzymatic digestion compared to other recombinant *S. aureus* proteins, which under the same reaction conditions, were completely digested. In fact, the SDS-PAGE pattern revealed one major polypeptide fragment with an apparent molecular mass of 37 kDa in both the cases (Figure 1.9). Furthermore, analysis of the overnight trypsin resistant fragment of rSdrE by PMF reveals the peptides coverage mainly from the CnaBE3 domain of the rSdrE protein (Figure 1.10). In order to further investigate what the major trypsin resistant fragment was comprised of, we performed N-terminal sequencing by Edman degradation. Moreover, N-terminal sequencing identified T_{793} of TPKYSLGDYV as an N-terminal start of resistant fragment (Figure 1.11), and all the subsequent released residues agreed with the downstream sequence located in the CnaBE3 domain of SdrE.

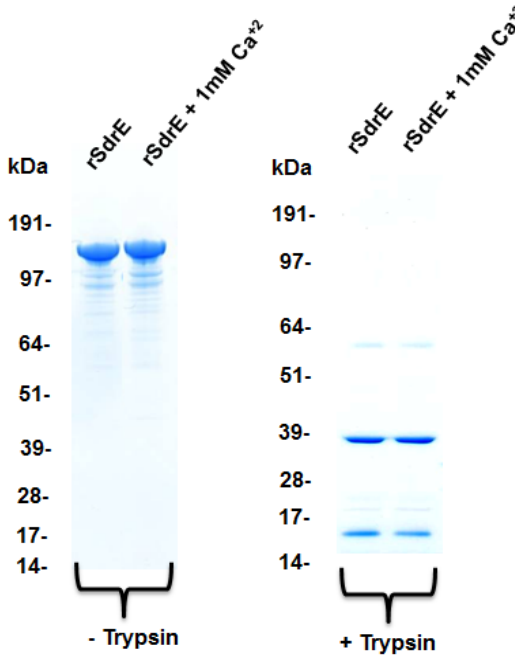

Figure 1.9: rSdrE shows resistance to proteolysis in absence and presence of calcium. SDS-PAGE analysis of overnight trypsin incubated rSdrE in absence and presence of 1 mM calcium. The molecular weight markers are shown on the left-hand side. In both the cases, the SDS-PAGE pattern revealed one major polypeptide fragment with an apparent molecular mass of approximately 37 kDa.

MGSSHHHHHHENLYFQGAENTSTENAKQDDATTSDNKEVVSETENNSTTENNSTNPIKKETNTDSQPEAKKESTSSSTQKQQNNVTA
TTETKPQNIEKENVKPSTDKTATEDTSVILEEKKAPNNTNNDVTTKPSTSEPSTSEIQTKPTTPQESTNIENSQPQPTPSKVDNQVTDAT
NPKEPVNVSKEELKNNPEKLKELVRNDSNTDHSTKPVATAPTSVAPKRVNAKMRFAVAQPAAVASNNVNDLIKVTKQTIKVGDGK
DNVAAAHDGKDIEYDTEFTIDNKVKKGDTMTINYDKNVIPSDLTDKNDPIDITDPSGEVIAKGTFDKATKQITYTFTDYVDKYEDIKS
RLTLYSYIDKKTVPNETSLNLTFATAGKETSQNVTVDYQDPMVHGDSNIQSIFTKLDEDKQTIEQQIYVNPLKKSATNTKVDIAGSQV
DDYGNIKLGNGSTIIDQNTEIKVYKVNSDQQLPQSNRIYDFSQYEDVTSQFDNKKSFSNNVATLDFGDINSAYIIKVVSKYTPTSDGEL
DIAQGTSMRTTDKYGYYNYAGYSNFIVTSNDTGGGDGTVKPEEKLYKIGDYVWEDVDKDGVQGTDSKEKPMANVLVTLTYPDGTT
KSVRTDANGHYEFGGLKDGETYTVKFETPTGYLPTKVNGTTDGEKDSNGSSVTVKINGKDDMSLDTGFYKEPKYNLGDYVWEDTN
KDGIQDANEPGIKDVKVTLKDSTGKVIGTTTTDASGKYKFTDLDNGNYTVEFETPAGYTPTVKNTTADDKDSNGLTTTGVIKDADN
MTLDRGFYK**TPKY**SLGDYVWYDSNKDGKQDSTEKGIKDVTVTLQNEKGEVIGTTKTDENGKYRFDNLDSGKYKVIFEKPAGLTQT
VTNTTEDDKDADGGEVDVTITDHDDFTLDNGYFEEDTSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS
DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS
DSDSDSDSDSDSDSDSDSDSDSDAGKHTPVKPMSTTKDHHNKAKA

Figure 1.10: **rSdrE amino acid sequence with peptides identified in the major resistant fragment of SdrE by PMF and Edman degradation.** Underlined sequences correspond to the peptides identified by PMF. Dark bold black sequence indicates amino acid residues identified by Edman degradation. The residues identified by Edman degradation lies in the CnaBE3 domain of SdrE.

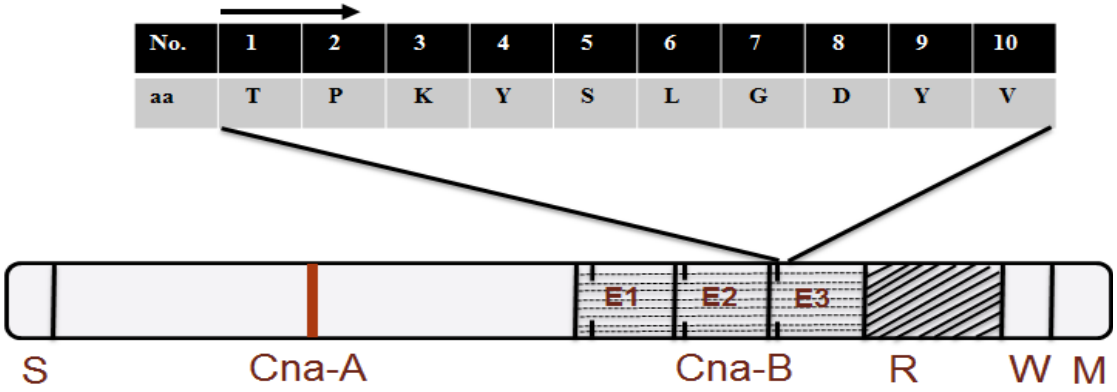

Figure 1.11: Schematic representation of the SdrE structural organization with the peptides identified from the major resistant fragment of SdrE by Edman degradation. Peptide sequence identified by Edman degradation shown here in the box suggests the purified overnight trypsin resistant fragment starts from the CnaBE3 domain of SdrE.

4.4: rCnaBE3 shows partial resistance to proteolytic cleavage

Based on PMF and Edman degradation analysis of the SE-UPLC purified overnight trypsin resistant fragment of rSdrE, we asked whether CnaBE3 domain of SdrE also shows trypsin resistant behavior. In order to look for the CnaBE3 behavior toward trypsin incubation, CnaBE3 domain were cloned, expressed and purified. Furthermore, trypsin proteolytic cleavage reactions were performed on the purified recombinant CnaBE3 domain of SdrE, which resulted in appearance of one major fragment with an apparent molecular mass of 12 kDa (Figure 1.12).

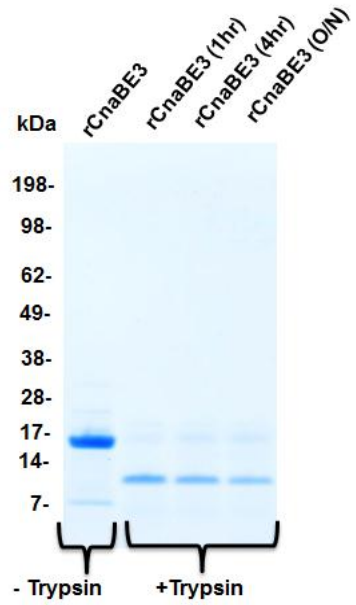

Figure 1.12: Proteolytic stability of rCnaBE3. Trypsin incubation of rCnaBE3 for different time interval (1hr, 4hr, and overnight), separated on 12% BT polyacrylamide. SDS-PAGE analaysis showed the partial resistance behavior of rCnaBE3 with the appearance of 12 kDa band after trypsin incubation. The molecular weight markers are shown on the left-hand side.

4.5: Intact mass spectrometry of rCnaBE3 shows loss of 17 Da

The result of rCnaBE3 trypsin incubation revealed that rCnaBE3 is partially resistant to trypsin. In order to investigate whether the resistance is due to the presence of intramolecular isopeptide bond, we did intact mass spectrometry of rCnaBE3. Indeed, the intact mass of rCnaBE3 domain revealed the loss of 17 Da which suggests the presence of an intramolecular isopeptide bond or some other post-translational modification (Figure 1.13). The theoretical (calculated) mass of rCnaBE3 is 15129.08 Da and interestingly we observed an average experimental (observed) mass of 14975.08 Da (Table 1.2), a measurement that is in agreement with the formation of one intramolecular isopeptide bond, resulting in the mass difference of 17 Da between the theoretical and experimental mass due to the loss of NH_3 group. Indeed, it is known from the literature that the NH_3 is eliminated when ε group of lysine bonds to the carboxyamide group of asparagine and result in formation of intramolecular isopeptide bond in gram positive cell surface proteins with CnaA and CnaB domains (Kang, Middleditch et al. 2009). This data indicated that the CnaBE3 domain shows the loss of 17 Da suggesting the presence of an isopeptide bond. However, low intensity of

the peptide with 17 Da loss and scarce reproducibility of the data do not allow us to definitively conclude that CnaBE3 contains an isopeptide bond. In particular, this could also be due to some other post-translational modification.

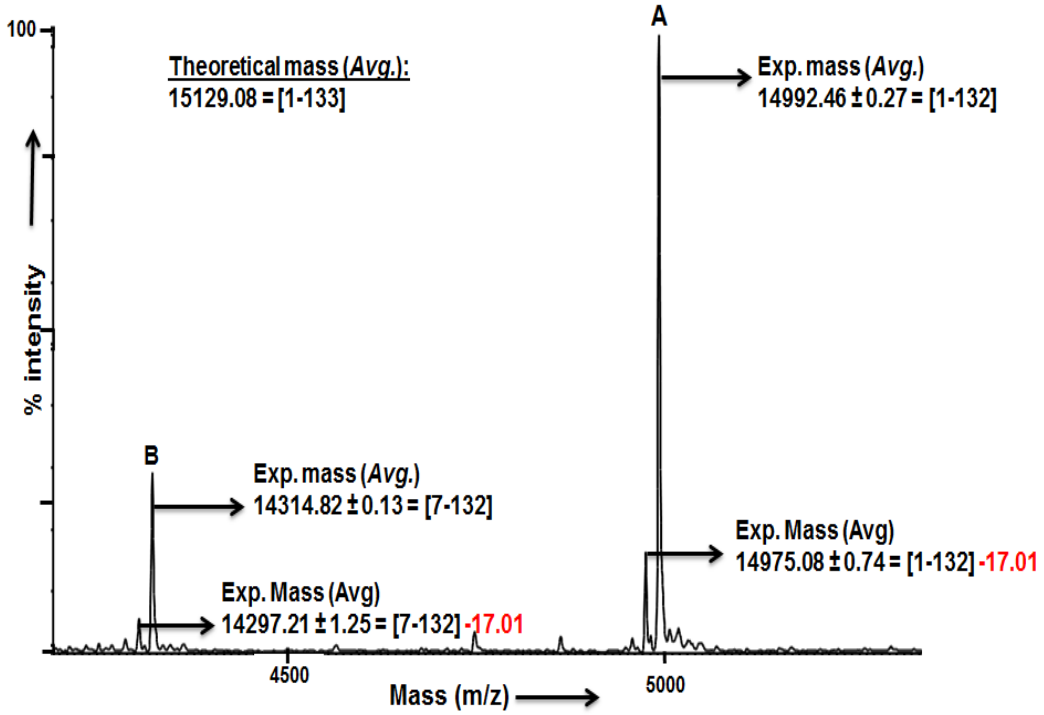

Figure 1.13: Exact mass determination of rCnaBE3 of SdrE. Two species of rCnaBE3 were identified by intact mass spectrometry of rCnaBE3 indicated here A and B, with different CnaBE3 amino acid sequence coverage. And each of the two species showed a mass difference of 17 Da between the theoretical and experimental mass.

Protein	$M_{average}$ (Da)		$M_{calculated} - M_{observed}$ (Da)	Intramolecular isopeptide bond predicted
	Calculated	**Observed**		
rCnaBE3	15129.08	14992.46	17	1

Table 1.2:Intact mass measurement shows mMass difference of 17Da between theoretical and experimental mass of rCnaBE3 of SdrE. Experimental mass observed by intact mass measurement of the rCnaBE3 shows mass difference of 17 Da between theoretical and experimental mass of rCnaBE3, indicating the presence of one intramolecular isopeptide bond.

4.6: Asparagine residue present in rCnaBE3 are not involved in isopeptide bond formation

Based on the fact that most of the surface proteins containing CnaA and CnaB domains form intramolecular isopeptide bond and the bonds are formed between lysine-asparagine/ lysine-aspartic acid residues in presence of glutamic acid, acting as a catalyst, we decided to further investigate if the loss is due to isopeptide bond and if so, explore the residues involved in the formation of intramolecular isopeptide bond in rCnaBE3. In order to do so, we designed six asparagine point mutation construct in the CnaBE3 domain of SdrE, where asparagines at different sites were substituted with alanine, but was able to generate only five construct till date. Therefore, trypsin proteolytic-cleavage reactions were performed for the five asparagine mutant of rCnaBE3, similar to that of the purified full-length rSdrE and wild-type rCnaBE3. We found that all the five mutants were partially resistant to enzymatic digestion similar to wild-type rCnaBE3. In fact, the SDS-PAGE pattern revealed one major polypeptide fragment with an apparent molecular mass of 12 kDa (Figure 1.14) in all the cases. This data indicates that the five asparagines mutated in CnaBE3 domain are not involved in the formation of intramolecular isopeptide bond.

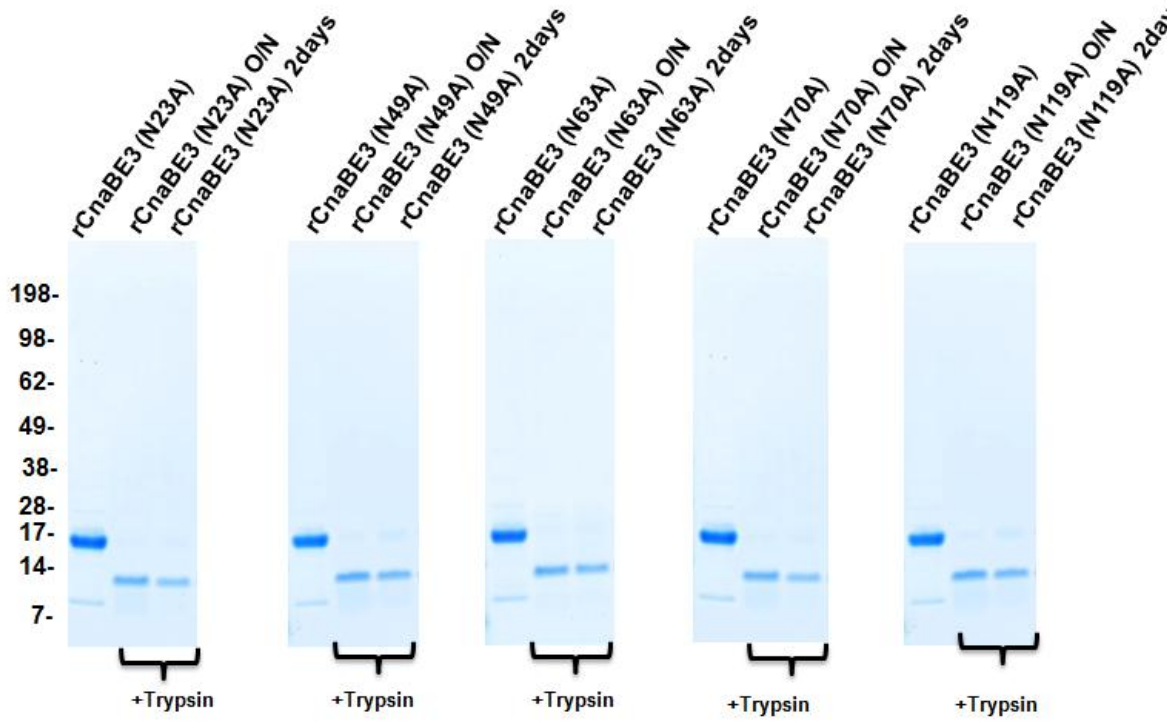

Figure 1.14: Proteolytic stability of five different asparagine point mutation constructs of rCnaBE3. Trypsin incubation of all five rCnaBE3 point mutation constructs at various time points separated on 12% BT SDS-PAGE. The molecular weight markers are shown on the left-hand side. In all digestions, the SDS-PAGE pattern revealed one major polypeptide fragment with an apparent molecular mass of 12 kDa.

4.7: CnaBE3 domain is highly conserved among epidemiologically relevant *S. aureus* strains

Apart from the partial trypsin resistant behavior of CnaBE3 domain of SdrE, taking into account exclusively the amino acid sequence of the last CnaB domain lying adjacent to region R in each Sdr protein i.e C3, D5 and E3 repeats of SdrC, SdrD and SdrE respectively (Figure 1.3) of the different *S. aureus* strains, which is reported to show 95-97% residue identity (Josefsson, McCrea et al. 1998), we compared the amino acid sequence of the three Sdr proteins and their respective CnaB domains of Newman strain to the amino acid sequences of Sdr proteins and CnaB domains of a representative panel of the NCBI published *S. aureus* strains. The comparison revealed that Sdr full-length proteins and their last CnaB domain shows a very high sequence homology and conservation to Newman Sdr proteins and last CnaB domain, even when belonging to phylogenetically distant strains (Figure 1.15). In

particular, considering exclusively the last CnaB domain of Sdr proteins in *S. aureus*, the residue identity percentage was never smaller than 97%, well above the percentages obtained comparing full-length proteins in phylogenetically distant *S. aureus* strains (Table 1.3). Remarkably even the strains lacking *sdrD* or *sdrE* gene, such as MRSA252 and NCTC8325 appeared to have a very high sequence identity to the CnaBD5 and CnaBE3 sequence of SdrD and SdrE respectively in Newman strain due to the high sequence similarity of last CnaB domains i.e. C2, D5 and E3 among each other. This finding prompted us to investigate whether the antibody raised against the highly conserved domain CnaBE3 of SdrE recognizes other Sdr proteins because of high homology in clinically relevant strains.

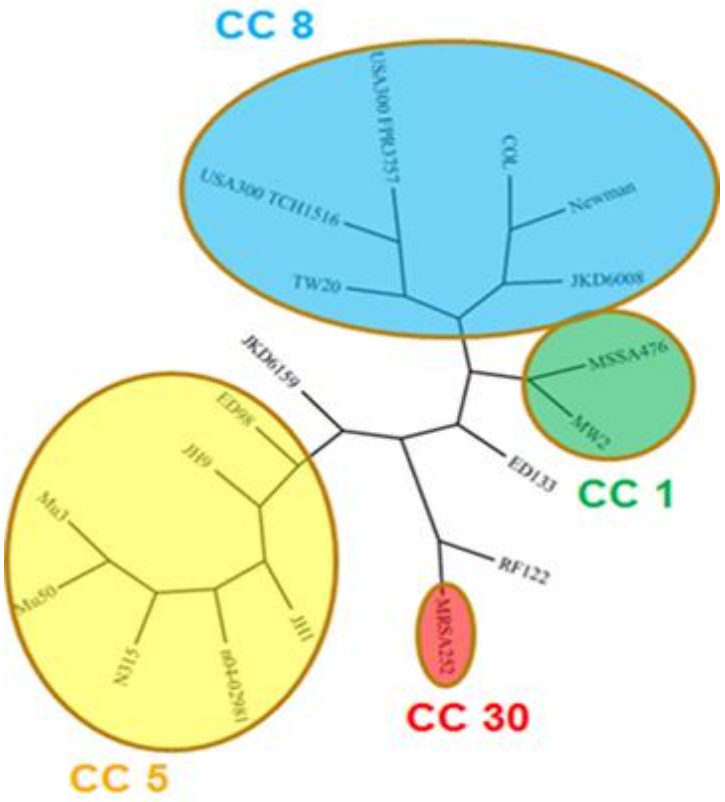

Figure 1.15: CnaBE3 domain is highly conserved among phylogenetically distant *S. aureus* strains. Phylogenetically distant *S. aureus* strains belonging to different clonal complexes (CC 1, CC 5, CC 8 and CC 30) with conserved CnaBE3 domain of SdrE.

Strains	Full-length Sdr proteins			Last CnaB domain of each Sdr proteins		
Newman	**SdrC**	**SdrD**	**SdrE**	**CnaBC2**	**CnaBD5**	**CnaBE3**
Mu50	95	91	95	100	98	97
Mu3	95	91	97	100	98	97
N315	95	91	95	100	98	98
MRSA252-erma 16	89		85	98	95	97
Col	100	99	99	100	100	99
MSSA476	96	95	97	100	98	99
USA300-FPR3757	99	95	98	100	100	99
MW2	96	96	97	100	97	99
NCTC8325	95	97		100	98	97

Present (≥ 95% sequence identity)
Present but variable (> 75 < 95% sequence identity)
Absent (≤ 75% sequence identity)

Table 1.3: Amino acid sequence homology and conservation of full-length Sdr proteins (SdrC, SdrD and SdrE) and last CnaB domain (i.e C2, D5 and E3) of Newman strain to Sdr proteins and the last CnaB domain of nine clinically relevant *S. aureus* strains. In the left panel is the comparison of sequence homology and conservation of *S. aureus* Newman strain full-length Sdr proteins. In particular, for SdrC the sequence homology of SdrC of Newman strain to SdrC of Mu50, Mu3, N315, Col, MSSA476, USA300-FPR3757 and MW2 strains shows high level (green) of sequence similarity (≥95%), whereas in strain MRSA-emra16 the sequence identity is 89%. For SdrD, sequence homology of all the nine strains with respect to Newman strain SdrD are ≥95% (green) except for strain MRSA-emra16 strain which lacks *sdrD* gene (red). Similarly for SdrE, sequence identity is ≥95% (green) for all strains except NCTC8325, which lacks *sdrE* gene (red).

In the right panel is the comparison of sequence homology and conservation of *S. aureus* Newman strain last CnaB domain of the three Sdr proteins to the last CnaB domain of the nine *S. aureus* strains. Table indicates >95% sequence homology for each of the strains, even in the strain lacking *sdrD* and *sdrE* gene.

4.8: Anti-CnaBE3 domain antibodies recognize all three Sdr full-length proteins in the analyzed panel of phylogenetically different *S. aureus* strains

Having described the high sequence homology and conservation of CnaBE3 domain throughout a panel representative of the phylogenetically distant *S. aureus* strains, we investigated whether polyclonal antibodies raised against the CnaBE3 domain were able to detect all three Sdr proteins in the same panel of *S. aureus* strains. The rCnaBE3 was injected in mice and anti-CnaBE3 polyclonal antibodies were recovered. Cell-wall protein from NCTC8325, Newman, MSSA476, MW2, N315, Mu50, Mu3, USA300-FPR3757, MRSA252, and TW20 strains were prepared and analyzed using anti-CnaBE3 mouse serum by immuno-blot assay. As shown in Figure 1.16, the anti-CnaBE3 polyclonal antibodies recognized the three Sdr proteins in the cell-wall fraction from different strains at their respective apparent molecular weight. As expected, neither SdrD protein in the *sdrD* negative strain MRSA252, nor SdrE protein in the *sdrE* negative strain NCTC8325 was detected. These data demonstrated that anti-CnaBE3 domain antibodies cross-reacts with other Sdr proteins i.e. SdrC and SdrD, even when they belong to evolutionarily distant strains and therefore underline the cross-reactive potential of CnaBE3 domain of SdrE to other Sdr proteins. Furthermore, this result prompted us to consider CnaBE3 domain as a possible vaccine candidate. Therefore, we decided to analyze the immunogenicity and the protective efficacy of the CnaBE3 domain whether small highly conserved structural domain of SdrE is sufficient for providing protection against *S. aureus* infection.

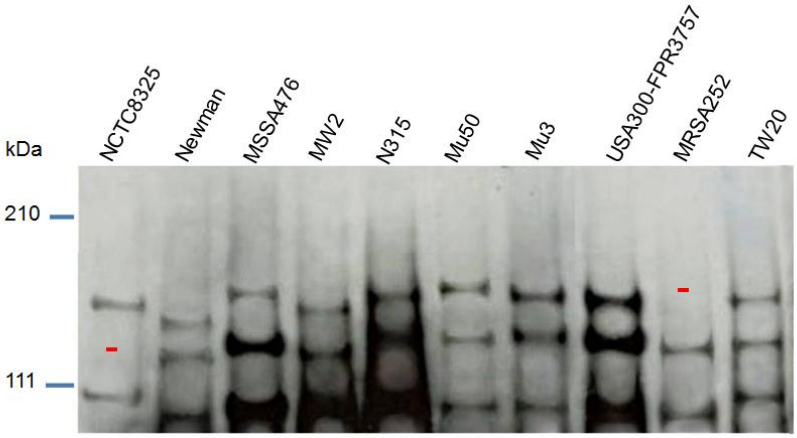

Figure 1.16: Immuno-blot assay showing anti-CnaBE3 antibodies recognize all three Sdr proteins in clinically relevant *S. aureus* strains. The cell-wall fraction of *S. aureus* strains NCTC8325, Newman, MSSA476, MW2, N315, Mu50, Mu3, USA300-FPR3757, MRSA252 and TW20 showed the presence of an immunoreactive band at a molecular weight comparable with the three Sdr proteins i.e. SdrC, SdrD and SdrE respectively. Indeed, no immunoreactive

bands at a molecular weight compatabile with the SdrD and SdrE were detected in the strains lacking *sdrD* and *sdrE* gene shown here by red bar.

4.9: SdrE vaccination protects mice against the challenge with *S. aureus* clinical isolates

Having demonstrated that SdrE is surface exposed and are involved in *S. aureus* pathogenesis and also CnaBE3 of SdrE showing high sequence identity with CnaBC2 and CnaBD5 of SdrC and SdrD, we then asked whether SdrE and CnaBE3 immunization could confer protection in mouse model. Protective efficacy of SdrE and CnaBE3 were tested in murine abscess model. In the murine abscess model, CD1 mice immunized with alum (sham), SdrE, and CnaBE3 separately, were infected intraperitonealy with sub-lethal inocula (approximately 2–6 × 10^7 CFU) of *S. aureus* Newman strain. Four days after infection, kidneys of sham, SdrE and CnaBE3 immunized were harvested and bacterial load was measured. SdrE vaccination induced 1.5 $\log_{10}$ CFU/ml reduction of bacterial burden in the kidneys of mice, whereas CnaBE3 vaccination reduced the bacterial burden in kidney by 1 $\log_{10}$ CFU/ml. Moreover, no significant difference in the reduction of bacterial count seen in the kidney when compared between the SdrE and CnaBE3 vaccinated mice (Figure 1.17).

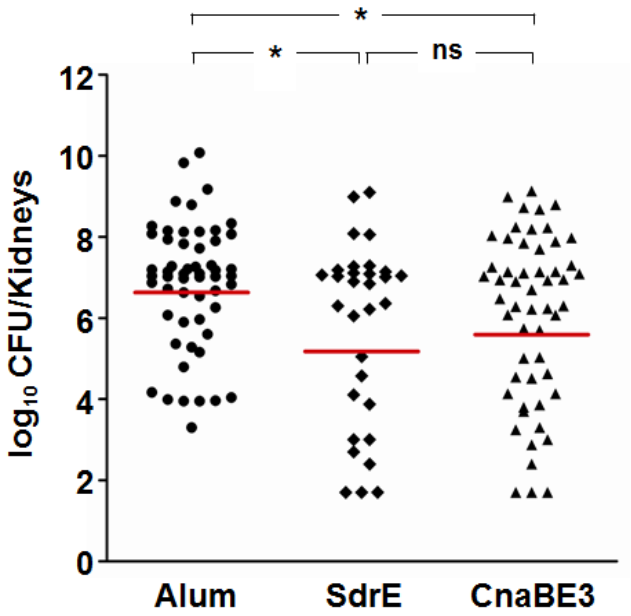

Figure 1.17: Protective efficacy of SdrE and CnaBE3 vaccination in murine abscess model. CD1 mice (N = 16 per group, 2-3 separate experiments) were immunized with alum alone (control) or full-length SdrE, CnaBE3 and then challenged with *S. aureus* Newman strain. The protection efficacy of full-length SdrE and CnaBE3 domain was assessed

evaluating bacteremia in the collected kidneys after an intravenous challenge. Mice were vaccinated intraperitoneally. SdrE vaccination induced 1.5 $\log_{10}$ CFU/ml reduction of bacterial burden in the kidneys of mice, whereas CnaBE3 vaccination reduced the bacterial burden in kidney by 1 $\log_{10}$ CFU/ml. Statistical analysis was performed by Mann-Whitney U-test.

4.10: CnaBE3 vaccination cross-protects against *S. aureus* NCTC8325 strain, a strain naturally devoid of *sdrE* gene

The cross-protection efficacy of CnaBE3 domain was tested by challenging immunized mice using the *S. aureus* strain NCTC8325, which is naturally devoid of *sdrE* gene. Three groups of 16 mice were immunized twice with rCnaBE3 at 14 days interval in three different experiments using the same conditions and challenged by approximately 2–6 × 10^7 CFU of *S. aurues* NCTC8325 strain for each mouse. Immunization with CnaBE3 domain caused a significant ($p < 0.01$) CFU reduction when compared with mice treated with alum alone (Figure 1.18).

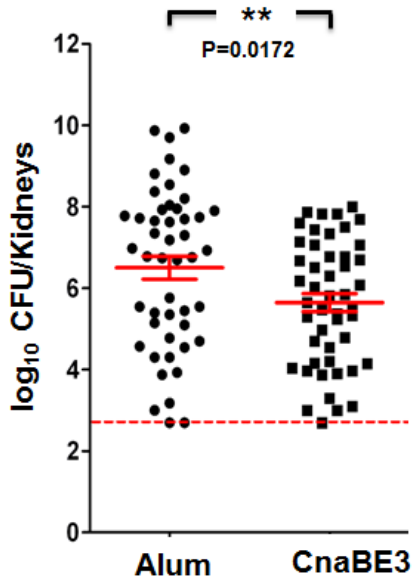

Figure 1.18: Cross-protective efficacy of CnaBE3 vaccination in murine abscess model. CD1 mice (N = 16 per group, 3 separate experiments) were immunized with alum alone (control) or CnaBE3 and then challenged with *S. aureus* NCTC8325 strain. The cross-protective efficacy of CnaBE3 was assessed evaluating bacteremia in the collected kidneys

after an intravenous challenge. Mice were vaccinated intraperitoneally. Statistical analysis was performed by Mann-Whitney U-test.

4.11: Calcium increases rSdrE stability

In 1998 Josefsson and colleagues reported that the protein segment encompassing the five B repeats (D1-D5) of SdrD protein, is subjected to large conformational changes with the addition of micro molar amounts of Ca^{+2}, suggesting the presence of high affinity Ca^{+2} binding sites. Herein we used differential scanning fluorimetry (DSF) to demonstrate the effect of calcium on the stability of rSdrE. Thermal stability of rSdrE (Tm = 45.26 ± 0.6°C) increased by 9°C in the presence of 1 mM calcium. Conversely, SdrE in presence of 1 mM EDTA (bivalent ion chelator) shows decrease in melting temperature (Tm: 39.72°C) (Figure 1.19). These data indicate that calcium binds to SdrE protein, favoring the structural stability of the protein.

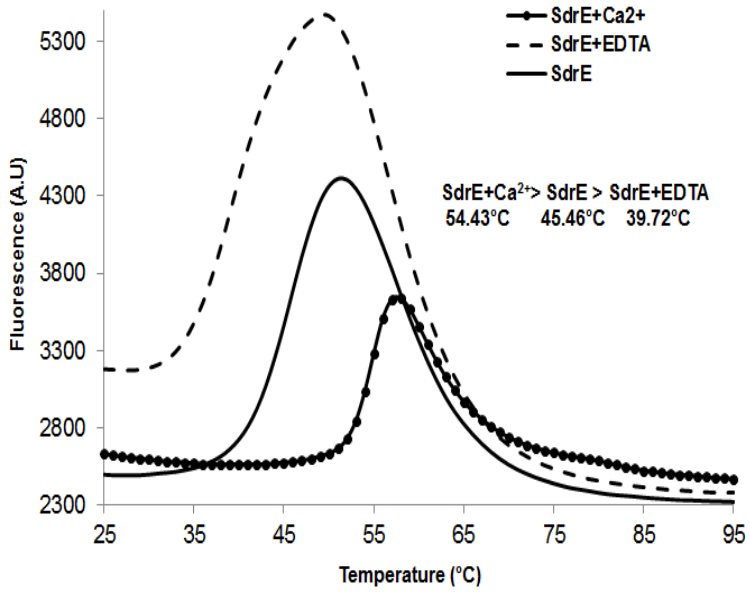

Figure 1.19: Thermal shift assay of SdrE by differential scanning fluorimetry (DSF). DSF experiments were performed on SdrE alone (line), and in the presence of 1 mM Ca^{2+} (line with closed circle) and 1 mM EDTA (dotted line). Calcium induced a large thermal stabilization (ΔTm > 9°C) of SdrE, while EDTA resulted in decrease of ΔTm < 6°C of SdrE (The smaller peak height for SdrE in presence of Ca^{2+} was due to a lower protein concentration).

5: DISCUSSION

In the present study, we tried to answer this question: Is the presence of intramolecular isopeptide bond in rSdrE makes it resistant to protease digestion, which are predicted to be present in most of the gram positive bacterial surface proteins containing CnaB domain.
If they are, could we use these signatures as predictive features to identify new vaccine candidates? There has been speculation in the past that such internal bonds could exist, but no mechanism has been put forward, and none has previously been proven to exist in *S. aureus* proteins. However, the presence of intramolecular isopeptide bond in Cna protein of *S. aureus* has been predicted (Deivanayagam, Rich et al. 2000; Kang, Coulibaly et al. 2007). According to our knowledge, this is for the first time that the loss of 17 Da is experimentally proven in one of the CnaB domain containing protein of *S. aureus*, which suggests the presence of intramolecular isopeptide bond or some other post-translational modification.
As from the literature it is known that isopeptide bond formation is a general post-translational protein modification in which an amide linkage occurs between an amino group of one amino acid and a carboxyl group of a different amino acid, with one or both of the functional groups provided by an amino acid side chain. In the examples of enzyme-catalyzed and spontaneous isopeptide bond formation described to date, the selection of the amino acids that participate in the formation of the covalent bond is typically highly specific (Osicka, Prochazkova et al. 2004; Kang, Coulibaly et al. 2007; Dierkes, Peebles et al. 2009). This is especially true in the case of the amino acid that donates the carboxyl group, where selection appears to be absolutely specific. This constraint presumably reflects the need to activate the carboxyl group as the first step in the formation of both enzyme-catalyzed (intermolecular) and spontaneous (intramolecular) isopeptide bonds (Marraffini, Dedent et al. 2006; Kang, Coulibaly et al. 2007; Dierkes, Peebles et al. 2009; Striebel, Imkamp et al. 2009).

Isopeptide bonds have until now been recognized for their importance in the intermolecular cross-linking of a variety of proteins, such as in ubiquitination (Pickart 2001), transglutamination (Greenberg, Birckbichler et al. 1991), and sortase-mediated cell-wall anchoring of surface proteins (Ton-That, Marraffini et al. 2004), as well as pilus polymerization. Isopeptide bonds make proteins resistant to proteases (Kang and Baker 2009), stabilize protein structures (Alegre-Cebollada, Badilla et al. 2010), attach proteins to cell surface (Marraffini, Dedent et al. 2006), and cross-link proteins in complex structures such as bacteriophage capsids (Wikoff, Liljas et al. 2000), bacterial pili (Kang, Coulibaly et al. 2007), and blood clots in humans (Ariens, Lai et al. 2002).

As most of the surface protein in gram-positive bacteria harboring the intramolecular isopeptide bonds are found to be resistant to trypsin digestion. In our study, interestingly we found full-length SdrE recombinant protein resistant to trypsin. Overnight trypsin incubation of SdrE resulted in the appearance of one major resistant fragment at approximately 37 kDa, and this resistant fragment remains stable even after protein incubation with trypsin for longer duration. PMF and N-terminal sequncing of the resistant fragment indicated that the resistant fragment comprise in the CnaBE3 domain of SdrE. Moreover, our study also revealed CnaBE3 domain of SdrE partially resistant to trypsin, as trypsin incubation of rCnaBE3 resulted in the appearance of one major resistant fragment at 12 kDa.

Furthermore, mass spectrometry was utilized to measure the intact mass of recombinant CnaBE3 and check for the presence of internal isopeptide bonds. Each bond should result in a 17 Da mass difference due to the loss of an NH_3 group, eliminated when the lysine ε-amino group bonds to the asparagine carboxyamide group. Indeed, we found a mass difference of 17 Da between theoretical and experimental mass of rCnaBE3, which is indicative for the formation of an intramolecular isopeptide bond due to loss of NH_3 in CnaBE3 domain of SdrE or some other post-translational modifiacation. Moreover, these data is in support of the report from Kang *et al.*, 2009 where they showed that the protein with CnaB domain is resistant to protease, and are predicted to contain intramolecular isopeptide bond. These data suggest that isopeptide bond cross-links could be important features in many surface proteins involved in adhesive functions, where stability against physical and chemical stresses is important. In addition these results, together with previous studies, will contribute to the knowledge on intra-domain bond as a stabilizing factor in CnaB domain of gram positive bacteria. Moreover, from our study, the intramolecular isopeptide bond or some other modification is expected in the CnaBE3 domain. At present we do not know whether proteins with such bonds all share a common evolutionary ancestor and are present in other Sdr proteins of *S. aureus*, or are present in a similar fashion to disulfide bonds, which is a feature of a variety of proteins in gram-negative bacteria.

For further investigation if the partial resistance and loss of 17 Da in rCnaBE3 is due to involvement of asparagine residue in CnaBE3 we designed, cloned, expressed and purified the point mutation constructs in CnaBE3, where asparagine is substituted with alanine based on literature. Asparagine is reported to be potential residues involved in the formation of intramolecular isopeptide bond (Kang, Coulibaly et al. 2007). Our study on protease resistant behavior of all the five asparagine point mutation constructs showed that none of the five asparagines are involved in the formation of intramolecular isopeptide bond, as neither of five

constructs is sensitive to protease treatment. This could be possible due to the reason that either the last asparagine to be checked, is involved or not asparagine but some other residue is involved in the formation of isopeptide bond. However, low intensity of the peptide with 17 Da loss and scarce reproducibility of the data do not allow us to definitively conclude that CnaBE3 contains an isopeptide bond. In particular, this could also be due to some other post-translational modification.

Furthermore, we confirm the expression of SdrE on the cell-wall of *S. aureus* Newman strain which was evidenced studying growth in TSB media. Moreover under the conditions tested, no HMW multimeric organization of SdrE on the staphylococcal cell-wall was identified indicating the protein is not forming intermolecular isopeptide bond.

Several types of insertion and deletion have been found in *sdrC* and *sdrD* genes of many bovine, ovine and human isolates due to horizontal gene transfer (Xue, Lu et al. 2011). This suggests that the *sdr* genes are variable. However, neither specific deletions of the region containing C2 or D5 domains nor any mutations in *sdrE* gene have been described so far, accounting for a strong specific stability of the region of Sdr proteins containing these CnaB domains. Moreover, the conservation analysis of Sdr proteins and their CnaB domains in our study showed that CnaBE3, CnaBD5 and CnaBC2 domains were highly conserved throughout the selected *S. aureus* strains, even compensating the absence of *sdrE* negative NCTC8325 and *sdrD* negative MRSA252 strains of *S. aureus.*

In our study by western blot analysis of the epidemiologically relevant *S. aureus* strains analyzed for homology and conservation of the Sdr proteins and last CnaB domain sequences revealed that anti-CnaBE3 antibodies were capable to detect the three Sdr proteins in the cell-wall fraction of each strain, even in the strain lacking *sdrE* and *sdrD* gene. This data shows cross-reactive behavior of CnaBE3 antibodies with the full-length Sdr proteins and therefore this study could be used to design a vaccine combination using this small highly conserved domain with more coverage.

Furthermore, biochemical characterization of full-length SdrE by differential scanning calorimetry revealed that the calcium plays an important role in maintaining the structural stability of the recombinant SdrE protein which is supported by the fact for the presence of EF-hand loop in the CnaB domain, known to have high calcium affinity and role in protein stability.

Apart from biochemical characterization of the protein we also looked for the vaccine potential of SdrE and highly conserved CnaBE3 domain. For many years, surface proteins of gram positive bacterial pathogens have been tested as antigens to generate immune responses

for the protection of human against infections. Surface protein SdrE vaccination was shown to generate protection against *S. aureus* infection in murine abscess model. These data indicate SdrE as a promising vaccine candidate to be tested in humans as reported by Stranger and Jones, 2006. Moreover, it has been previously shown by Stranger and Jones, 2006 that SdrE in combinations with other antigen can confer greater protection than single antigens. However, in our study we interestingly found that conserved CnaBE3 domain vaccination confers significant protection to *S. aureus* infection. Our data suggested a possible role of CnaBE3 domain as vaccine candidate supported by the high homology and conservation of the CnaBE3 domain, immunogenicity and stability.

As binding or functional domains which interact with host ligands are generally considered as promising vaccine candidates, our study shows that the domain acting as a spacer could also serve as a vaccine candidate. More interestingly, we found that CnaBE3 protects against a strain lacking the *sdrE* gene. And therefore it is worth noting that this relatively short domain could be easily fused to other protein antigens in a vaccine combo, thus enhancing the rate of coverage and protection of the vaccine. The results of these studies are expected to shed light on functional mechanisms of Sdr-like proteins and potentially allow the design of novel vaccine candidates against *S. aureus*.

In addition, from literature we know that the crystal structure of the B region of the *S. aureus* collagen binding (Cna) protein presents an inverse IgG like fold (Deivanayagam, Rich et al. 2000). Interestingly, the B motifs of the Sdr proteins display some homology to the B repeats of the collagen binding protein Cna of *S. aureus.* So far a number of speculations have been arisen about the potential function of the B repeats of the Sdr proteins. Perhaps the number of the repeats is linked to the level of binding capacity or maybe it is responsible for displaying the A region away from the cell surface (Deivanayagam, Rich et al. 2000). Because not much is known about the structure of full-length SdrE, it would be necessary to obtain and study the crystal of the full-length and CnaB region of SdrE protein, to entirely understand the function of this structure. Detailed structural studies on full-length SdrE in collaboration with Northwestern University, Chicago, USA are in progress to solve some of the numerous questions remaining.

CHAPTER 2

Characterization of the *Staphylococcus aureus* serine protease EpiP

2.1: INTRODUCTION

2.1.1: Overview of *S. aureus* proteases

Proteases are essential for the expression of all proteins, controlling protein composition, size, shape, turnover and ultimate destruction. Their actions are often exquisitely selective, each protease being responsible for splitting a very specific sequence of amino acids under a given set of environmental conditions. There are over 500 human proteases and proteases also occur in other vertebrates, as well as in plants, insects, marine organisms and in all microbes. Genes coding for proteases account for about 2% of the human genome and 1-5% of genomes of bacteria and viruses. Proteases play a central role in conception and birth, life, ageing, and death. An over- or under- abundance of a particular protease or altered levels of natural inhibitors/activators of proteases can lead to pathological conditions. Proteases are also essential in viruses, bacteria and parasites since they participate in their replication, spread and maintenance of infectious diseases within the human or animal hosts. Due to all these pivotal biological roles, proteases represent important potential targets for medical intervention.

The proteinases have so far been categorized into six catalytical classes according to the chemical group within the enzyme that participates in the cleavage of the substrate peptide bond. A peptide bond can be broken directly by amide hydrolysis in which a water molecule is added, as is the case with glutamic-, aspartic-, and metallo-proteases. In the case of serine-, cysteine-, and threonine-proteinases the polarized amino acid is made nucleophilic, attacks the peptide carbonyl group in the substrate and forms an intermediate before the final hydrolysis. The proteases are classified in to three different subfamilies: a) Cysteine proteases; b) Serine proteases and; c) Metalloproteases as described below:

a) Cysteine proteases

Papain originating from papaya is the best studied of all cysteine proteases. It has been used for degrading meat fibers for thousands of years. Papain is also capable to cleave the Fc (crystallisable) portion from the Fab (antigen-binding) portion of immunoglobulins. Several mammalian intracellular proteinases like cathepsins B, L, K, and S belong to this family, as do caspases and calpains. Some parasitic proteases and bacterial extracellular proteinases like the staphylococcal enzymes staphopain A (ScpA) and staphopain B (SspB) are also cysteine proteases.

b) Serine proteases

This class comprises two distinct families. The chymotrypsin family includes the digestive pancreatic enzymes such as chymotrypsin, trypsin and elastase and also other mammalian enzymes like plasmin, furin, and cathepsin G. Subtilisin is a serine protease in prokaryotes. Subtilisin is evolutionary unrelated to the chymotrypsin-families, but shares the same catalytic mechanism utilizing a catalytic triad, to create a nucleophilic serine. This is the classic example used to illustrate convergent evolution, since the same mechanism evolved twice independently during evolution. The subtilisin family includes the bacterial enzymes such as subtilisin and staphylococcal serine protease (V8, SspA).

c) Metalloproteases

The metalloproteinases exert their effect mainly extracellularly and at neutral pH, as do most of the serine proteinases. The metalloproteinases may be one of the phylogenetically older classes of proteinases and are found in bacteria, fungi as well as in all higher organisms. They differ widely in their sequence and their structure but have some highly conserved domains. The great majority of enzymes contain a zinc atom that is catalytically active. In some cases, zinc may be replaced by another metal such as cobalt or nickel without loss of the enzymatic activity. The metal ion usually serves to coordinate two to four side chains and is indispensable for the activity of the enzyme. Many of these enzymes are calcium dependent in that calcium is required for maintaining the molecule's conformation. Both the staphylococcal aureolysin (Aur) and the large group of mammalian matrix metalloproteinases (MMPs) are referred to this class of proteases.

2.1.2: Staphylococcal extracellular proteases

Until today a handful of extracellular proteases from *S. aureus* have been described. The genes of extracellular proteases are organized on the bacterial chromosome into four distinct operons: staphylococcal serine protease (ssp) operon, serine protease like proteins (spl) operon, staphylococcal cysteine protease (scp) operon, and aureolysin (aur) operon. Thus the V8 protease (SspA) and staphopain B (SspB) both are named by their encoding ssp operon although only SspA is a serine protease whilst SspB is a cysteine protease (Shaw, Golonka et al. 2004).

Serine protease V8 (SspA)

V8 protease was first purified from culture filtrate of *S. aureus*, strain V8 by Drapeau et al in 1972. Their studies revealed that V8 protease exclusively cleaves peptide bonds on the

carboxyl-terminal side of either aspartic acid or glutamic acid (Drapeau, Boily et al. 1972). Since then V8 protease has been widely used for determining the primary structures of proteins. The V8 protease does not have a high degree of sequence identity with other serine proteases and does not contain any disulphide bridges, but its tertiary structure was found to have significant structure homology with, for example the staphylococci epidermolytic toxins A and B (Prasad, Leduc et al. 2004).

Metalloprotease aureolysin (Aur)

The primary and tertiary structures of aureolysin have been determined revealing a polypeptide chain of 301 amino acids which is folded into a β-pleated N-terminal domain and an α-helical C-terminal domain, a typical fold for the thermolysin family of metalloproteinases [Banbula, A., 1998]. The diverse family of bacterial metalloproteinases encompasses several enzymes, which are acknowledged virulence factors, including *Pseudomonas aeruginosa* elastase, *Legionella pneumophila* and *Listeria monocytogenes* metalloproteinases, *Vibrio cholerae* hemagglutinin protease, *Staphylococcus epidermidis* elastase, and the lambda toxin of *Clostridium perfringens*. In contrast to these proteinases, however, little is known about the exact role of aureolysin in the pathogenicity of *S. aureus*. In vitro, aureolysin has been shown to cleave the plasma proteinase inhibitors, α1-antichymotrypsin and α1-proteinase inhibitor. Aureolysin cleaves the oxidized form of α1-proteinase inhibitor faster than it cleaves the native inhibitor, suggesting that bacteria which secrete these metalloproteinases may specifically take advantage of the host defense oxidative mechanism to accelerate elimination of α1- proteinase inhibitor and consequently increase tissue degradation by neutrophil elastase. The structure of the aureolysin gene is strongly conserved among *S. aureus* strains. This argues in favor of the likelihood that the enzyme may have an important housekeeping function (Sabat, Kosowska et al. 2000).

Cysteine proteases Staphopain B (SspB) and Staphopain A (ScpA)

The cysteine proteases exhibit, with their papain like features, potent elastinolytic activity which has been considered to be of importance for bacterial spread within tissues and also for the forming of abscesses (Potempa, Dubin et al. 1988). Kinin-releasing cysteine proteinases have been reported from various microbes. From *S. aureus* a submicromolar concentration of ScpA generated large amounts of kinin from human plasma within 5 minutes from exposure and in a guinea pig experimental model, ScpA in concert with SspB induced vascular leakage and lowering of blood pressure. Thus staphopain mediated vascular leakage was proposed to

be involved in the oedema formation at the infected sites as well as in the induction of septic shock (Imamura, Tanase et al. 2005).

2.1.3: Regulation of *S. aureus* extracellular proteases

The transcription of most of the extracellular protease genes occurs maximally at postexponential-phase, being positively regulated by the accessory gene regulator (agr) and negatively regulated by the staphylococcal accessory regulator (sarA). The enzymes are excreted as zymogens. Using mutations in each protease gene, the proteolytic cascade of activation has been analyzed.

2.1.4: Novel serine protease EpiP of *S. aureus*

RiPPs (ribosomally-synthesized and post-translationally-modified peptides) are a class of natural products that exist in all forms of life (Arnison, Bibb et al. 2013). Many gram-positive bacteria produce RiPPs that have antimicrobial activity and are called lantibiotics. These lantibiotics are effective antibacterial agents against other gram-positive bacteria (Schnell, Entian et al. 1988; Willey and van der Donk 2007) and have been investigated as possible alternatives for the treatment of bacterial infections (Cotter, Hill et al. 2005; Smith and Hillman 2008). A cascade of proteins is required to both post-translationally modify the lantibiotic into its mature form and to protect the producing organism from the effects of the lantibiotic via immunity proteins (Schnell, Engelke et al. 1992). Generally, the genes coding for these proteins are found in clusters on either plasmids or chromosomes (Augustin, Rosenstein et al. 1992; Kuipers, Rollema et al. 1993), and have been identified in several bacteria such as *Lactococcus, Bacillus, Staphylococcus, Streptococcus*, and *Enterococcus*. Recently, *in silico* screenings have uncovered 49 unidentified clusters from bacteria not known to produce lantibiotics (Marsh, O'Sullivan et al. 2010), which indicates that these clusters are more common than previously thought. Not all genes in the clusters are conserved, nor are they arranged in the same order amongst strains and species (Siezen, Kuipers et al. 1996). Moreover, not all bacteria that have these clusters can produce active lantibiotic, and if a lantibiotic is produced it does not always have antibacterial activity (Smith and Hillman 2008; Bierbaum and Sahl 2009).

Most lantibiotic gene clusters have a gene that codes for a lantibiotic leader peptide protease (Siezen, Kuipers et al. 1996). EpiP is a subtilisin-like extracellular epidermin leader peptidase and is required for proteolytic processing of the mature lantibiotic epidermin in *Staphylococcus epidermidis* (Schnell, Engelke et al. 1992; Geissler, Gotz et al. 1996);

however, its function in *S. aureus* remains unknown since there has been debate regarding if *S. aureus* produces epidermin (Otto and Gotz 2001; Daly, Upton et al. 2010; Joo, Cheung et al. 2011) or just retains the lantibiotic immunity genes for self-protection against other lantibiotics or for increased virulence. It has been shown that bacteria that have these lantibiotic gene clusters, even if they do not produce the lantibiotic, have increased virulence and resistance. Many lantibiotic peptidases are cytoplasmic, but some like NisP, CylP and EpiP reside extracellularly (Siezen, Kuipers et al. 1996). The regulation of expression of exoproteins, surface proteins and virulence factors for *S. epidermidis* and *S. aureus* is controlled by the accessory gene regulator (*agr*) quorum sensing system (Kies, Vuong et al. 2003). When *agr* was deleted in *S. epidermidis* a decrease in mature epidermin production was seen, which was due to the decreased ability of EpiP to process the pro-peptide of epidermin rather than a decrease in the transcription of genes in the lantibiotic cluster (Kies, Vuong et al. 2003). Therefore, *agr* quorum sensing does not interfere with the transcription of epidermin biosynthetic genes, but controls the extracellular processing of the N-terminal leader peptide by the EpiP protease.

Herein we characterized a *S. aureus* protein annotated as epidermin leader peptide processing serine protease (EpiP). The *epiP* gene contains a peptidase_S8 domain present in subtilisin-like serine proteases, and which is present in protective antigens of several other species. In addition, homologous proteins expressed by other bacterial species, have been shown to play important roles during pathogenesis. In particular, the *Streptococcus pyogenes* homologue, SpyCEP (also named ScpC), inactivates IL-8 by catalyzing its C-terminal cleavage (Edwards, Taylor et al. 2005). As a consequence, SpyCEP impairs the recruitment of neutrophils at the site of infection and subsequent bacterial clearance (Hidalgo-Grass, Dan-Goor et al. 2004; Zinkernagel, Timmer et al. 2008). Neutrophils are probably one of the most important elements of the host immune system against *S. aureus* infections. Indeed, defects in the number or function of neutrophils can result in an increased susceptibility to *S. aureus* infections in humans (Andrews and Sullivan 2003). However, despite the emerging relevance of proteolytic bacterial antigens, there are currently no reports describing characterization of the *S. aureus* EpiP protein. Herein, we present a study aimed at revealing the first fundamental biochemical and functional properties of this novel peptidase_S8 domain containing protein.

2.2: AIM OF THE STUDY

- Characterization of auto-cleavage mechanism of serine protease EpiP
- Identify the optimal growth conditions for *epiP* gene expression in *S. aureus* by qRT-PCR
- Investigate expression and processing of the EpiP protein in *S. aureus* by western blot
- Investigate the vaccine potential of *S. aureus* extracellular protease EpiP

2.3: MATERIALS AND METHODS

2.3.1: Antigen identification by *in silico* analysis

In silico antigen identification was performed analyzing the *S. aureus* NCTC8325 genome by several bioinformatic algorithms as described in material and method of Chapter 1.

2.3.2: Cloning, expression and purification of wild type EpiP and EpiP-S393A

epiP was amplified by PCR from the *S. aureus* NCTC8325 strain and was cloned in without its putative leader sequence (aa 1–27) as an N-terminal 6X-histidine-tag (His-tag) construct. The His-tagged construct was cloned using oligonucleotides EpiP-F/EpiP-R (Table 2.1). The PCR product was subcloned into the pET-15b+ vector using the Polymerase Incomplete Primer Extension (PIPE) technique (Klock and Lesley 2009). To determine if the EpiP protein acts as a serine protease, we mutated the hypothesized catalytic serine 393 residue to alanine, generating rEpiP-S393A (Figure 2.1). The Stratagene QuikChange™ site-directed mutagenesis kit was used to construct the S393A mutant according to the procedure outlined in the manufacture's technical manual. The conditions for the PCR reaction were initial denaturation 95°C 5 min, followed by 18 cycles of denaturation 95°C 1 min, annealing 55°C 1 min, extension 68°C 7 min, the final extension was 68°C for 10 min and then cooled to 4°C. A silent mutation for L394 (TTA to CTG) was added to improve the AT-rich region of this gene for PCR.

The ampicillin resistant rEpiP and mutant His-tagged plasmids were transformed into kanamycin resistant BL21 (DE3) Magic cells, grown and expressed in terrific broth (TB) and harvested according to previously described procedures (Millard, Stols et al. 2003). Pelleted cells were resuspended in buffer, sonicated, and cleared lysates were loaded onto an IMAC Ni^{2+}-affinity sepharose column (GE Healthcare His-trap HP) as described previously (Millard, Stols et al. 2003). Fractions of protein from IMAC were pooled and diluted with 10 mM Tris-HCl pH 8.3 to reduce the salt concentration and then loaded onto a GE HiTrap SP-5 mL column for IEX chromatography. The loading buffer was composed of 10 mM Tris-HCl pH 8.3 and 5 mM beta-mercaptoethanol and the protein was eluted over 30 CVs with a linear gradient from 0-1 M NaCl in loading buffer plus 1M NaCl. rEpiP or rEpiP-S393A eluted at approximately 0.27 M NaCl and was concentrated using Amicon protein concentrators 10,000 MWCO (Millipore). Protein concentration was determined using the absorbance at 280 nm and the extinction coefficient 0.942 M^{-1} cm^{-1}. Protein purity was determined using SDS-PAGE and was purified to near homogeneity.

MKIIKRAIISLIILSLLISITMSNASASEELYYSVEYKNTATFNKLVKKKSLNVVYNIPELHVAQIKMTKMHANALANYKNDIKYINATCSTCI
SEKTIDRTSNESLFSRQWDMNKITNNGASYDDLPKHANTKIAIIDTGVMKNHDDLKNNFSTDSKNLVPLNGFRGTEPEETGDVHDVN
RKGHGTMVSGQTSANGKLIGVAPNNKFTMYRVFGSKKTELLWVSKAIVQAANDGNQVINISVGSYIILDKNDHQTFRKDEKVEYDAL
KAINYAKKKKSIVVAAAGNDGIDVNDKQKLKLQREYQGNGEVKDVPASMDNVVTVGSTDQKSNLSEFSNFGMNYTDIAAPGGSFAY
NQFGVDKWMNEGYMHKENILTTANNGRYIYQAGTSLATPKVSGALALIIDKYHLEKHPDKAIELLYQHGTSKNNKPFSRYGHGELDVY
ALNVANQKAS

Figure 2.1: Amino acid sequence of EpiP. Residue in red is the hypothesized catalytic serine 393 residue, mutated for the generation of rEpiP-S393A mutant.

Table 2.1: PCR primers designed to amplify corresponding gene

Gene	Primer	Nucleotide sequence
epiP wild-type	EpiPF	5'-CTGTACTTCCAGGGCTCAGAAGAACTATATTACAGTGTTG
	EpiPR	5'-AATTAAGTCGCGTTAACTTGCTTTTTGATTTGCTACATTTAATGCT
epiP-S393A	EpiPF	5'-GAAGATATATTTATCAAGCTGGAACTGCGCTGGCCACACCTAAAGTTTCG
	EpiPR	5'-CGAAACTTTAGGTGTGGCCAGCGCAGTTCCAGCTTGATAAATATATCTTC

2.3.3: EpiP cleavage mechanism

For the evaluation of the cleavage mechanism of EpiP (if it was intra- or inter-molecular), rEpiP-S393A was incubated with the rEpiP for 1, 4 and 24 hour in a 5:1 ratio at 37°C and the cleavage pattern was analyzed by SDS-PAGE.

2.3.4: Intact mass spectrometry of rEpiP

Intact mass measurement of rEpiP purified protein was done according to the procedure described for rCnaBE3 domain of SdrE in material and method section of Chapter 1.

2.3.5: Peptide mass finger printing of auto-cleaved rEpiP

In order to investigate the peptide coverage of the auto-cleaved rEpiP, PMF was done. SDS-PAGE analysis was performed using Nu-Page 12% BT polyacrylamide SDS-APGE according to the manufacturer's instructions. Furthermore PMF of the auto-cleaved rEpiP bands were done following the protocol used for PMF of trypsin resistant fragment of SdrE in Chapter 1.

2.3.6: EpiP expression analysis

The expression of EpiP in *S. aureus* was evaluated *in vitro* by western blot and quantitative real-time PCR (qRT-PCR). *S. aureus* strains i.e Newman, LAC (USA300) and Mu50 (USA100) were grown overnight in TSB at 250 rpm, 37^0C using aerated Erlenmeyer flasks. Overnight cultures were centrifuged at 4,000 rpm, at 4°C for 15 minute. For preparing cell-wall fractions, pellets from 5 ml overnight culture was suspended in 500 µl TSM (50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 0.5 M sucrose) buffer. 50 µl of lysostaphin (5 µg/µl) was added to samples and incubated for 1 hour or until lysis at 37°C at 400 rpm in a thermomix. After lysis, samples were centrifuged at 4000 rpm for 15 min at 4°C and supernatants containing the cell-wall fraction were used for western blot analysis using anti-EpiP mouse serum. The supernatant (extracellular) protein was prepared by concentrating 10 ml of overnight culture supernatant up to 100 X using 7,500 Da MWCO Vivapore 10/20 solvent absorption concentrator (Sigma-Aldrich). For western blot analysis, equal amounts of supernatant protein from each strains were separated on 12% BT SDS-PAGE and transferred to nitrocellulose membranes using similar protocol as mentioned in Chapter 1. The primary antibody anti-EpiP mouse serum used was with a dilution of 1/2000 and horseradish peroxidase-conjugated goat anti-mouse secondary antibody at 1/4000 dilution.

2.3.7: RNA isolation and cDNA synthesis for qRT-PCR

For *in vitro* RNA extraction, bacteria were grown in TSB till exponential phase (OD_{600} nm= 4) and stationary phase (OD_{600} nm=12, overnight). One ml culture from each growth time point was taken and immediately mixed with 2 volume of RNAprotect bacteria reagent (Qiagen) to stabilize the RNA and further incubated for 20 min at room temperature. Bacteria were harvested by centrifugation at 4,000 rpm for 5 min at 4°C. Pellet was stored at -80°C until use. The RNeasy Mini kit (Qiagen) was used for extraction of RNA according to manufacturer's instruction.

2.3.8: Quantitative real-time PCR

Quantitative real-time PCR was performed by using the RNA isolated from bacteria grown in TSB. Total RNA (1 µg) was reversed transcribed with 0.5 mM dNTP, 50 ng random hexamers and 200U of Invitrogen Superscript II Reverse Transcriptase, according to the manufacturer's recommendations. RNA was denatured and the cDNAs were purified with a QIAquick PCR purification kit. One microliter of cDNA was amplified on the Stratagene MX3000P Real-Time PCR (Sratagene, LaJolla, CA USA) with SYBR GreenER™ qPCR universal kit (Invitrogen) and 10 µM of the EpiP forward and reverse primers listed in Table 2.2. Reaction mixtures were denatured for 10 min at 94°C, followed by 40 cycles of 30 sec at 60°C, 1 min at 72°C and finished with a dissociation ramp from 55°C to 95°C. The level of expression of each gene was calculated by using the cycle threshold (Ct) of the overnight grown bacteria as the calibrator. Expression fold of genes from each experiment was then normalized with their respective 16s-rRNA expression level.

Table 2.2: PCR primers designed to amplify corresponding gene in qRT-PCR

Gene	Primer	Nucleotide sequence
epiP	EpiPF	5'- CATAAAGCGCGCTATTATTAG
	EpiPR	5'- CTTTATACACATCAAGCTCAC

2.3.9: Active immunization

Five-week-old CD1 mice were immunized intraperitoneally with a prime-booster injection of 20 µg purified recombinant EpiP adsorbed to aluminum hydroxide adjuvant (alum, 2 mg/ml) in 14' day interval. Control mice received equal amounts of PBS and alum adjuvant. Animals were bled immediately prior to the first immunization and 23 days thereafter, and sera were examined for IgG antibodies directed against purified EpiP using the Luminex technology.

2.3.10: Peritonitis model

In order to look for the protection efficacy of rEpiP and rEpiP-S393A mutant, immunized mice were challenged on day 24 by intraperitoneal injection of a lethal dose 5×10^8 CFU of *S. aureus* Newman strain. Mice were monitored daily for 7 days. Two independent experiments run under the same conditions, were performed to assess protective efficacy of rEpiP and rEpiP-S393A in the peritonitis model. Experiments were analyzed using one tailed Fisher's exact test.

2.4: RESULTS

2.4.1: Antigen selection by *in silico* prediction

The epidermin leader peptide processing serine protease (EpiP) was identified as a putative vaccine candidate by analyzing the *S. aureus* NCTC8325 genome through several bioinformatic algorithms as described in material and method of Chapter 1. *epiP* is part of an eight gene operon containing lantibiotic genes, epiABCDPFEG, which was previously identified in *Staphylococcus epidermidis* as epiABCD, epiPQ and epiFEG (ten Broeke-Smits, Pronk et al. 2010) (Figure 2.2). The *S. aureus* EpiP protein was predicted to have an extracellular localization due to the presence of a leader peptide and the lack of other known signals for membrane or cell-wall anchoring (Figure 2.3). Analysis of genomes available in public databases indicated that *epiP* is a dispensable gene and is either well conserved or absent (data not shown). By searching databases for sequence similarities and for the presence of conserved functional domains, we found that *epiP* contains a peptidase_S8 domain present in subtilisin-like serine proteases and has conserved residues (Asp140, His 187, and Ser393) corresponding to the canonical catalytic triad for serine proteases (Rawlings and Barrett 1994). This domain is present in vaccine candidates of other species including the *Streptococcus agalactiae* C5a peptidases, *Streptococcus pyogenes* SpyCEP, the *Streptococcus pneumoniae* PrtA, and the *Neisseria meningitidis* NMB1969 (AspA) (Hidalgo-Grass, Mishalian et al. 2006; Zinkernagel, Timmer et al. 2008; Zingaretti, Falugi et al. 2010). In addition, as mentioned above, proteins belonging to this family have been shown to play important roles during pathogenesis and impair the recruitment of neutrophils at the site of infection (Turner, Wooldridge et al. 2002; Hidalgo-Grass, Mishalian et al. 2006; Zinkernagel, Timmer et al. 2008; Zingaretti, Falugi et al. 2010). Therefore, on the basis of all these observations, we hypothesized that EpiP could be a protective antigen against *S. aureus* infection.

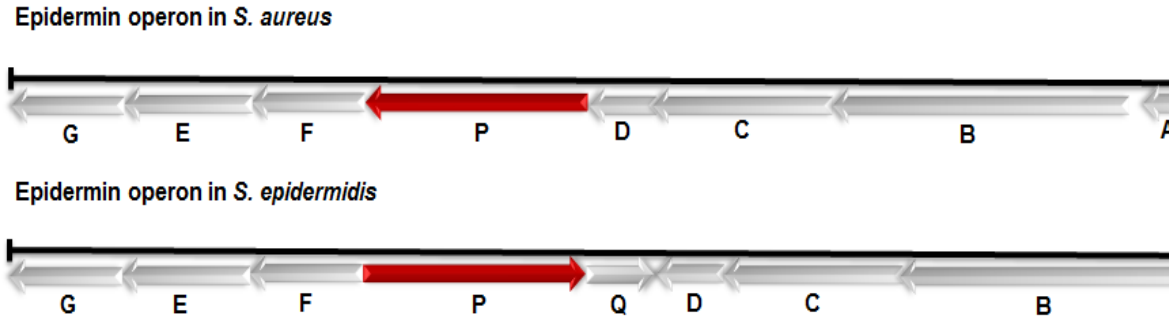

Figure 2.2: Epidermin operon in *S. aureus* and *S. epidermidis*. In *S. aureus*, all genes are located on the same strand, while in *S. epidermidis*, epiPQ are located on the other strand of epiABCD and epiFEG.

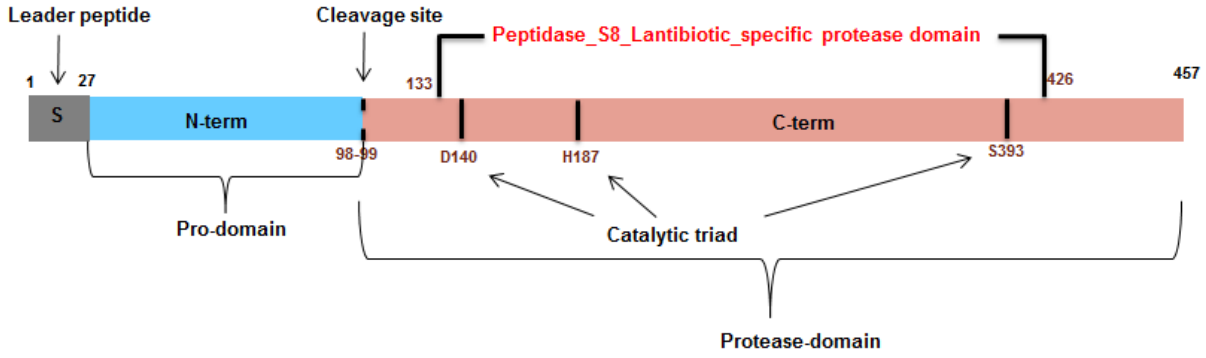

Figure 2.3: EpiP protein organization. The layout of the EpiP protein including the leader peptide, pro-domain, cleavage site, and protease-domain is shown. The catalytic triad includes residues D140, H187, and S393 and the Peptidase_S8_Lantibiotic specific protease domain includes residues 133 to 426. In gray is the leader peptide and blue the pro-domain and the cleavage site of the protein between residues 98 and 99. There is not a cell-wall anchoring domain predicted in the sequence.

2.4.2: Enzymatic activity of the recombinant EpiP

EpiP deprived of its putative leader sequence (aa 1–27) with 6X- histidine tag on N-terminus was expressed in, and purified from *E. coli*-soluble extracts. rEpiP appeared as three bands migrating with apparent molecular weights of ~ 50, 39 and 8 kDa by SDS-PAGE (Figure 2.4 A). Peptide mass fingerprinting of the three bands separated by SDS-PAGE identified peptides specific for the three fragments. More specifically, peptides covering both the N- and C-terminus of the protein were identified in the upper 50 kDa band, while coverage of the 39

kDa fragment spanned amino acids threonine-98, corresponding to threonine-88 in recombinant EpiP to lysine-446, and peptides in the 8 kDa fragment were found, mapping in the first 41 amino acid residues (Figure 2.4 B) corresponding to pro-domain of the protein.

The theoretical mass of rEpiP is 49744.27 Da (1-446 AA), but by intact mass spectrometry we observed an experimental mass of 39694.48 Da (88-446 AA; 30 ppm) (Figure 2.4 C) for the most abundant species in the sample starting with amino acid threonine at position 88. This observation fits well with the pattern obtained by SDS-PAGE in which the first peptide identified by mass spectrometry in the band migrating at 39 kDa begins with threonine-88 (Figure 2.4 B). Altogether, these data indicate that rEpiP is composed of three polypeptides resulting from the cleavage of the protein between amino acids lysine-98 and threonine-99 corresponding to residue 87 and 88 in rEpiP.

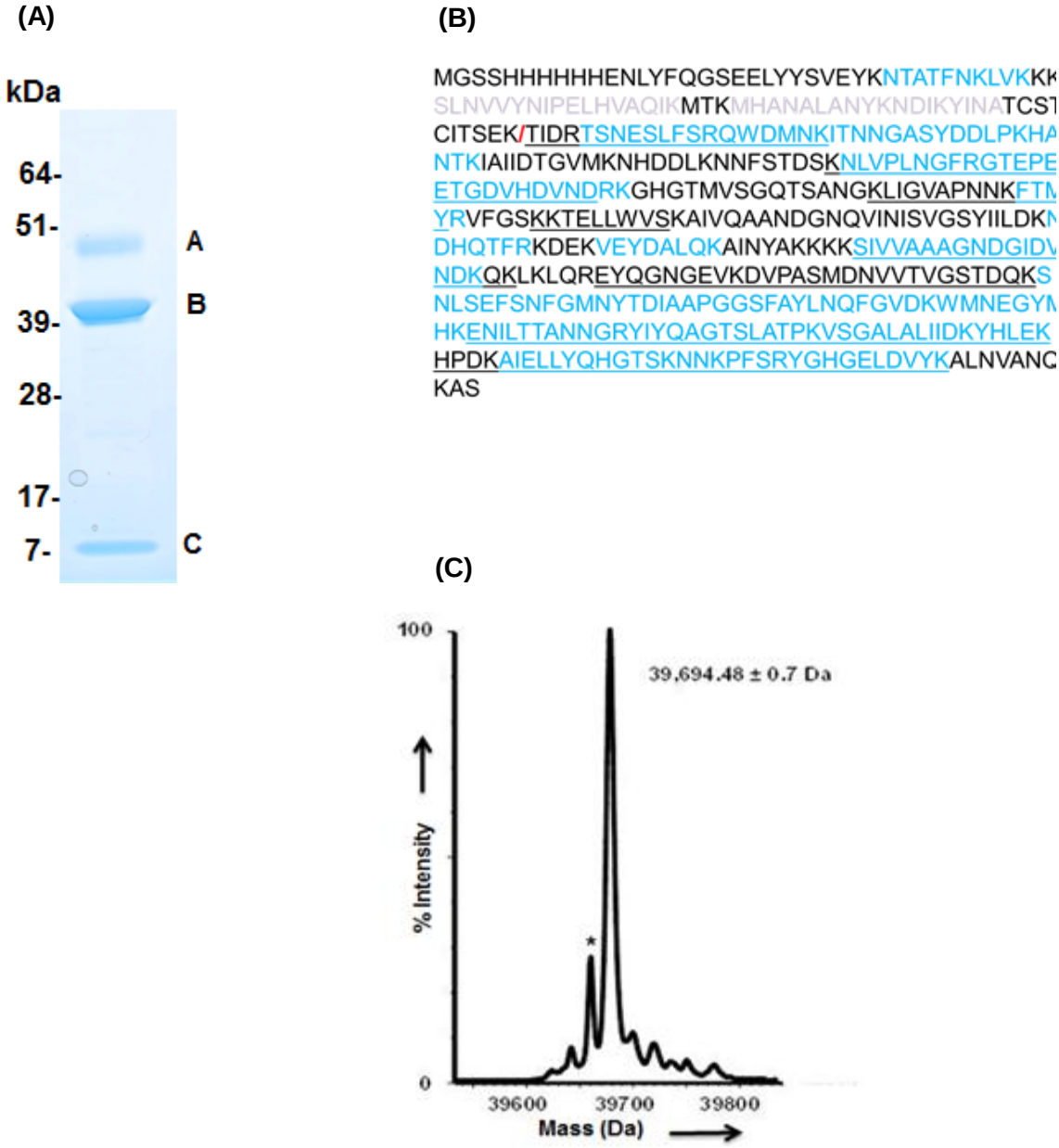

Figure 2.4: Analysis of recombinant EpiP purified from *E. coli* extracts through SDS-PAGE and mass spectrometry. A) In SDS-PAGE stained with coomassie the proteins migrated with three species (peak A, B and C). The MW of the bands from top to bottom is comparable with the size of the mature protein (≈ 50 kDa), with the protease domain (≈ 39 kDa) and the pro-domain (≈8 kDa). B) Protein sequence of rEpiP with peptides identified in bands A, B and C shown in Fig. 2.4 A by peptide mass fingerprinting are indicated on the sequence of rEpiP: blue color corresponds to the peptides identified in band A, underlining indicates peptides from band B and gray corresponds to peptides from band C. C)

Experimental mass observed by intact mass measurement of the rEpiP fits with band B present Fig. 2.4 A. The sequence starts with threonine-88 indicating that the cleavage site of the protein occurred between lysine-87 and threonine- 88 (as indicated by a red diagonal in Figure 2.4 B).

2.4.3: rEpiP processing is due to autoproteolysis

We next investigated whether cleavage of EpiP could be due to autoproteolysis, as reported for other extracellular bacterial proteases (Ikemura, Takagi et al. 1987; Anderson, Wetherell et al. 2002). To address this point, we generated rEpiP-S393A, in which the serine 393 residue of the predicted catalytic triad was substituted with an alanine. The purified mutant protein was then subjected to SDS-PAGE analysis to investigate whether the mutation impaired the cleavage pattern observed with rEpiP. As shown in Figure 2.5, the mutant protein appeared as a single band with an electrophoretic mobility comparable to that of the non-cleaved protein and comparable to the one of the upper band of rEpiP, indicating the absence of cleavage.

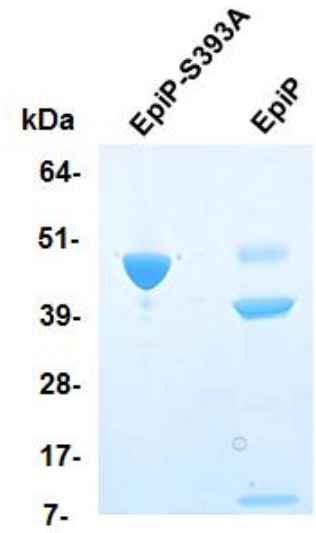

Figure 2.5: **rEpiP processing is due to autoproteolysis**. In SDS-PAGE stained with coomassie the rEpiP-S393A protein migrated with one species with the size of the mature protein (≈ 50 kDa). This pattern is in clear contrast with the three protein species present in the rEpiP.

2.4.4: EpiP cleavage occurs through an autocatalytic intra-molecular mechanism

In order to understand if the cleavage of EpiP occurs through an inter- or an intra-molecular mechanism, rEpiP-S393A and rEpiP were mixed in a 5:1 ratio and incubated for up to 24 hours. During this experiment, we observed no increase in intensity of the protease-domain of the protein by SDS-PAGE (Figure 2.6). Given that rEpiP did not cleave the mutant, these data suggest that the cleavage event of EpiP occurs through an intra-molecular autocatalytic mechanism.

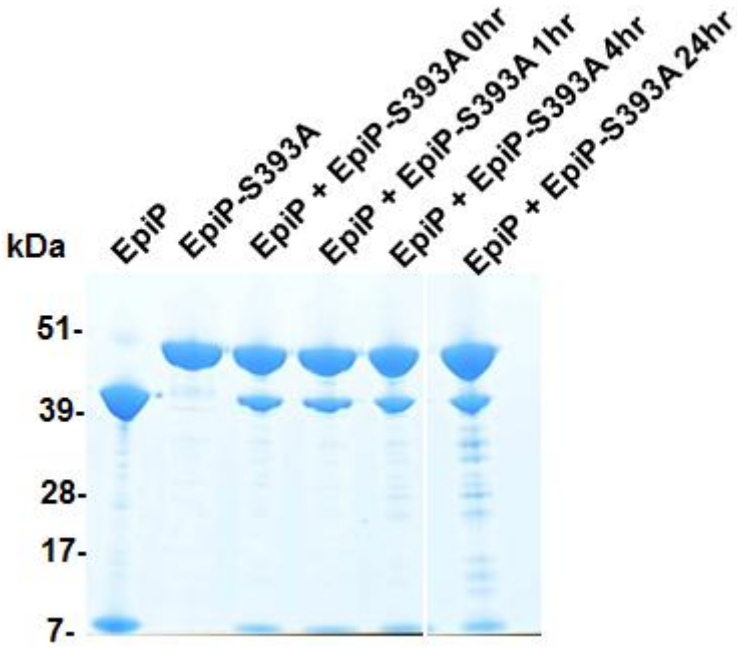

Figure 2.6: EpiP cleavage occurs through an autocatalytic intra-molecular mechanism. Analysis of the auto-proteolytic activity of rEpiP-S393A and rEpiP derivative by a time-course co-incubation of the two proteins in a 5:1 ratio; the two recombinant proteins (rEpiP and rEpiP-S393A) were used as molecular mass controls (lanes 1 and 2). No decrease of the non-cleaved rEpiP-S393A was observed through the 24 hr incubation time indicating that EpiP uses an intra-molecular autocleavage mechanism to remove the pro-domain.

2.4.5: EpiP is expressed and processed in *S. aureus* cells

In order to assess the expression of the *epiP* gene, qRT-PCR was performed on RNA isolated from exponential and stationary phase bacteria, grown in TSB, which is the same medium was used to prepare challenge inocula for infection experiments in the animal model. Under these conditions the gene was found to be expressed and increased gene expression was observed in exponential as compared to stationary growth phase (Figure 2.7).

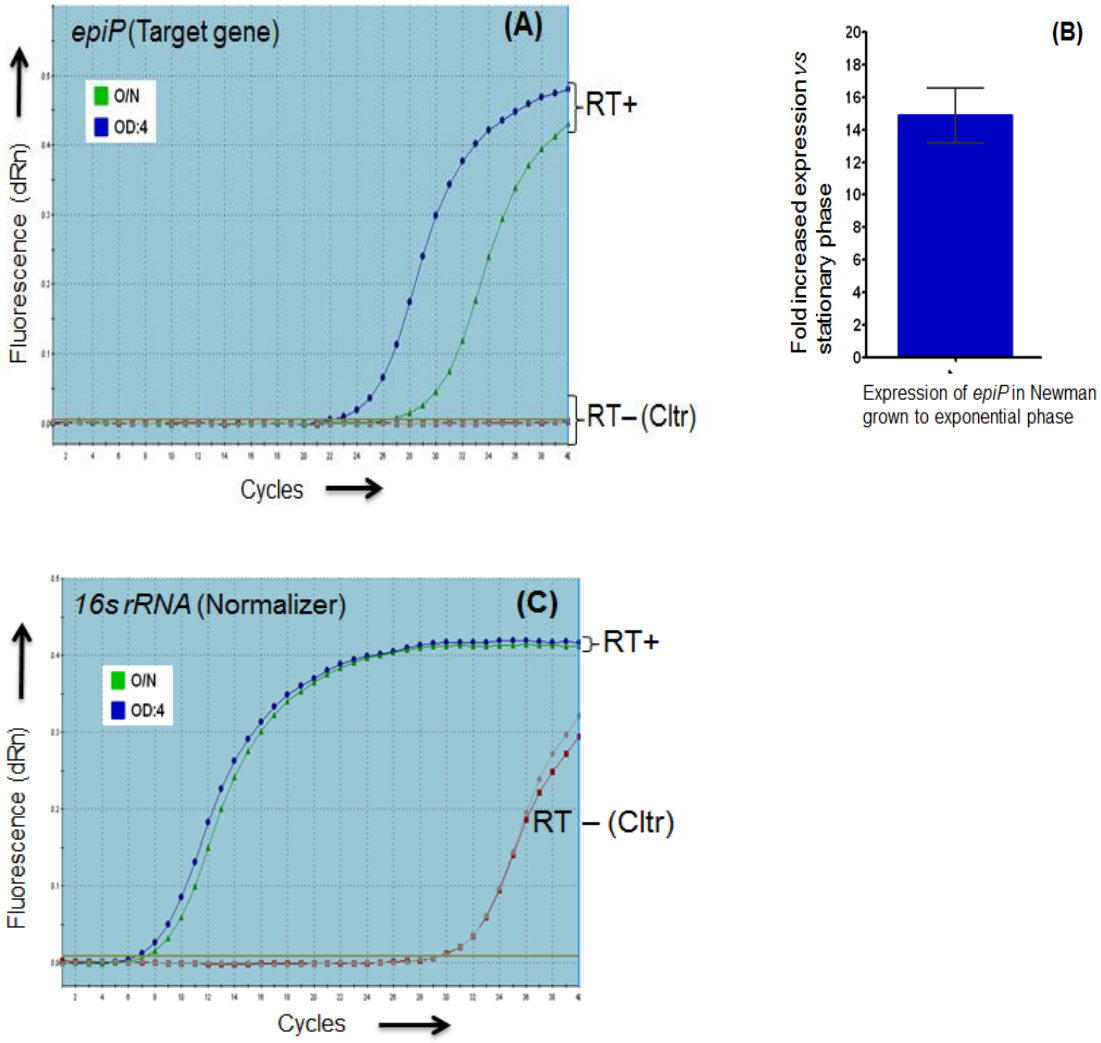

Figure 2.7: *epiP* was found to be up-regulated in exponential phase. A and B) qRT-PCR of *epiP* gene on RNA extracted from *S. aureus* Newman grown in TSB showed ~16 fold increased expression at exponential phase (OD_{600} = 4) as compared to stationary (overnight culture) phase. C) *epiP* expression was normalized against 16s-rRNA for which expression was found to be constant throughout the *S. aureus* growth phase.

Moreover, we also looked for the *epiP* gene expression in different *S. aureus* strains (i.e. Newman, LAC, and Mu50) by qRT-PCR with RNA isolated from bacteria grown to exponential phase in TSB. Under the condition tested, the *epiP* gene was expressed in Newman and LAC strains but not in Mu50. Furthermore, LAC showed ~2 fold increased gene expression as compared to Newman (Figure 2.8).

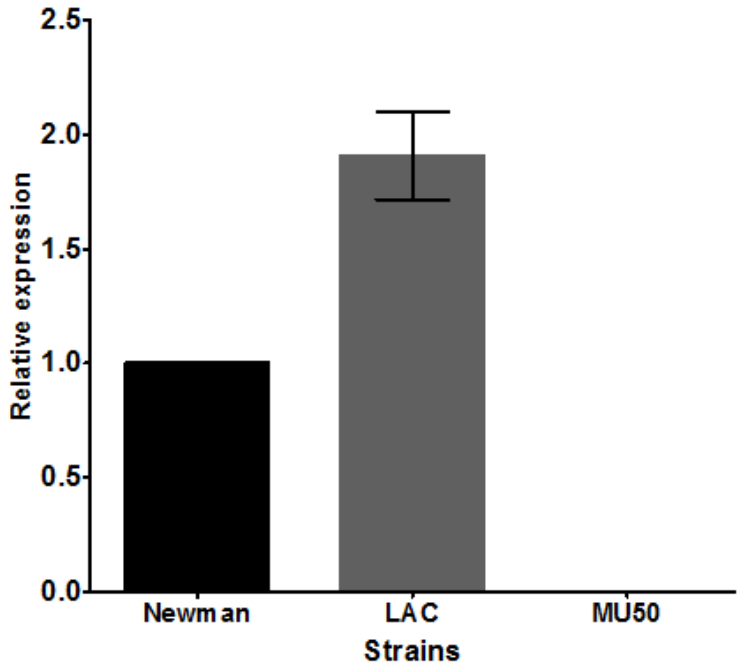

Figure 2.8: Relative quantification of *epiP* gene using qRT-PCR in three different strains of *S. aureus*. qRT-PCR of *epiP* gene on RNA extracted from *S. aureus* Newman, LAC and Mu50 strain grown in TSB. *epiP* gene was found to be expressed in Newman and LAC strains, whereas no expression was detected in Mu50 strain, as expected. Furthermore, LAC showed ~2 fold increased gene expression as compared to Newman.

The *S. aureus* EpiP lacks a known localization signal (Figure 2.3), and it may either be released in the extracellular milieu or be anchored to the cell-wall through an unknown mechanism as reported for EpiP of *S. epidermidis* (Augustin, Rosenstein et al. 1992; Schnell, Engelke et al. 1992; Engelke, Gutowski-Eckel et al. 1994). To experimentally verify its localization, we isolated the cell-wall fraction, the culture supernatant as well as protoplasts of *S. aureus* cells and we analyzed them by immunoblot using an anti-EpiP mouse serum. For this experiment we decided to use the *S. aureus* strain Newman deficient for SpA (SEJ2)

(DeDent, Bae et al. 2008; Frankel, Wojcik et al. 2010). This strain was selected to reduce as much as possible the non-specific staining in western blot analyses due to the binding of IgGs mediated by SpA (Zhang, Jacobsson et al. 1998). The anti-EpiP serum identified several immunoreactive bands, including one at approximately 50 kDa and another at 8 kDa. These bands appear to have a molecular weight comparable to the non-cleaved EpiP and the pro-domain, respectively. Furthermore, this indicates that the serum mainly detects the pro-domain of EpiP (Figure 2.9 A). At this point, the identity of the 17 kDa immunoreactive band is not clear. No immunoreactive bands were detected in the cell-wall fraction suggesting that the protein was mainly released into the extracellular milieu.

Interestingly, different patterns were observed in the supernatant and protoplast preparations. In particular, a band migrating at a molecular weight comparable to the non-cleaved form was more abundant in the cytoplasm while the pro-domain appeared more prevalent in the supernatant. This suggests that the protein was produced in its non-cleaved form and got processed during or after its release into the extracellular milieu.

We then performed an expression analysis of the culture supernatants of *S. aureus* Newman (ST-254), a standard laboratory strain, and given its epidemiological relevance, the strain LAC, which belongs to the pandemic clone USA300. In addition, we used the Mu50 strain that belongs to the hospital acquired clone USA100 and, which lacks *epiP*. A specific band, corresponding to the full- length protein was detected in *S. aureus* Newman and the LAC strain (Figure 2.9 B). However, as expected, no specific reactive bands were observed in the hospital acquired strain Mu50. Overall, these results suggest that EpiP is released into the extracellular milieu and it undergoes a processing event similar to the one observed with the recombinant protein.

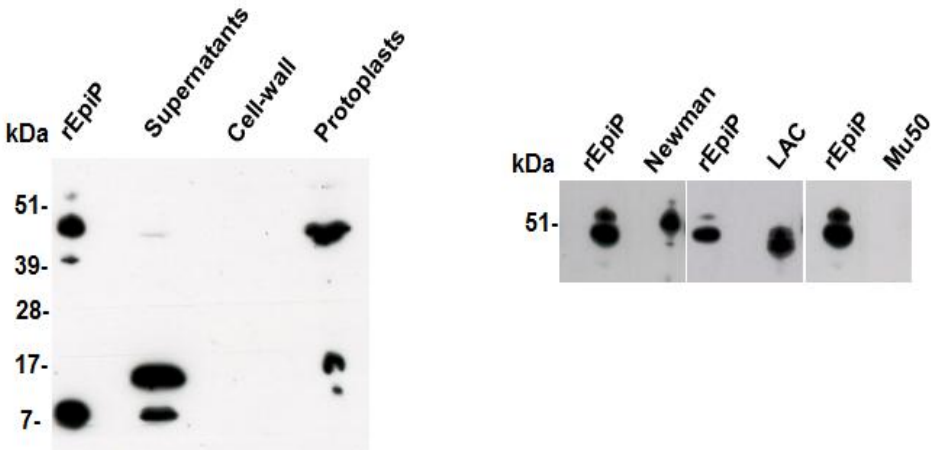

Figure 2.9: EpiP expression and processing in *S. aureus*. A) Analysis of culture supernatant, cell-wall fraction and protoplast of *S. aureus* strain Newman by western blot using anti-EpiP mouse antiserum. The three polypeptides (non-cleaved, protease domain and pro-domain) of the purified recombinant protein were recognized by the serum. In the supernatant preparation, a faint band with a MW compatible with the non-cleaved protein was visible as well an immunoreactive band with a size similar to the pro-domain. An additional band was visible which does not correspond to any of the three polypeptides of EpiP. No immunoreactivity was detected in the lane where the cell-wall preparation was loaded. A major immunoreactive band was present in the protoplast preparation comparable with the MW of the non-cleaved protein. Other minor protein species were also detected that appeared similar to the ones observed in the lower MW range of the culture supernatant preparation. B) Immunoblot analysis using anti-EpiP mouse serum of culture supernatants of *S. aureus* strains Newman, LAC showed the presence of an immunoreactive band at a MW compatible with the non-cleaved EpiP. As expected, no immunoreactive bands were detected in the culture supernatant of the *S. aureus* strain Mu50 which lacks *epiP*.

2.4.6: EpiP vaccination protects mice against the challenge with *S. aureus* clinical isolates

As predicted, mice vaccinated with EpiP were protected from staphylococcal lethal infection. Given that the protein was found to be released extracellularly, protection might be associated to EpiP antibodies blocking the function of the protein. Moreover, protective efficacy of rEpiP and rEpiP-S393A was found comparable (Figure 2.10).

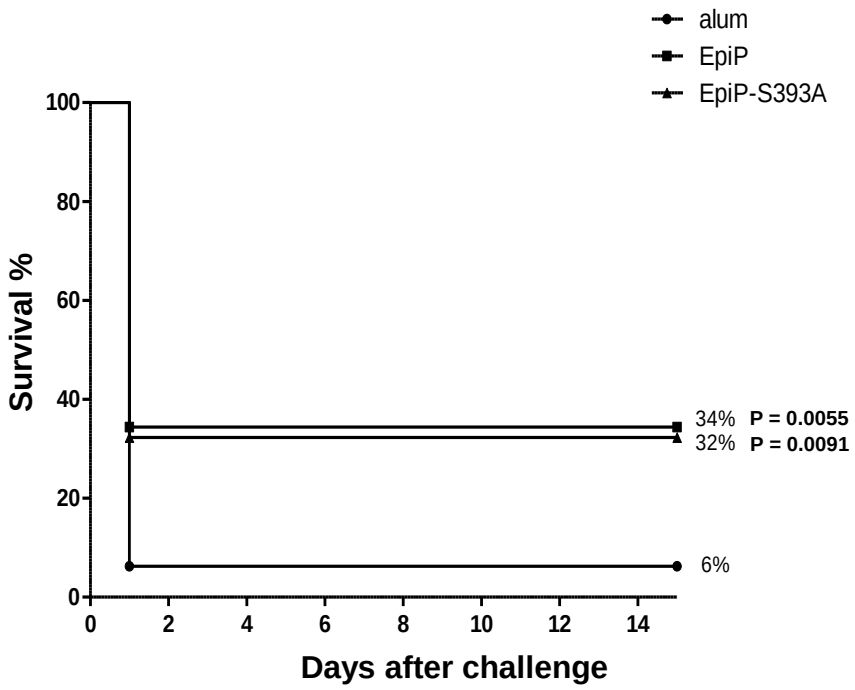

Figure 2.10: Protective efficacy of EpiP vaccination in the peritonitis model. A lethal dose of *S. aureus* Newman strain was used to challenge CD1 mice (N = 32 per group, 2 separate experiments) and survival rates were followed for 1 week. Mice were vaccinated intraperitoneally and challenged with the *S. aureus* strain Newman. Statistical analysis was performed using one tailed Fisher's exact test.

2.5: DISCUSSION

Based on bioinformatics antigen prediction (reverse vaccinology) and sequence homology we identified the *S. aureus* epidermin leader peptide processing serine protease (EpiP) as a novel vaccine candidate. Like other extracellular bacterial proteases, EpiP enzymes are known to be synthesized with an N-terminal leader peptide involved in translocation across the cell membrane, followed by a pro-peptide, which assists the folding of the mature enzyme. On protease secretion, the leader peptide is removed by a signal peptidase, and the pro-region is auto-catalytically cleaved and degraded, leading to enzyme switch from the zymogen to the active form, however it does not contain a known cell-wall anchoring motif. Indeed, our data indicate that the protein is released into the extracellular milieu. Furthermore, we found that the recombinant purified protein undergoes through an autocatalytic intra-molecular cleavage which separates an N-terminal pro-peptide from the protease domain. The protein expressed by staphylococcal cells appears to be similarly processed. Most likely this event occurs after its translocation beyond the cell membrane. Indeed, while most of it is present with a molecular weight comparable with the full length protein in the protoplast preparation, which contains cytoplasmic and membrane proteins, in the extracellular fraction the most abundant reactive band appears to be the pro-peptide. The biological meaning of these observations requires further investigation.

To better characterize the protein we decided to perform a crystallographic study and solved the structure of rEpiP and rEpiP-S393A in collaboration with Northwestern University, Chicago, USA. We found rEpiP to be cleaved somewhere between residue 95 and 100. Moreover we also solved the structure of rEpiP-S393A and found no cleavage (data not shown).

Indeed, by removing the predicted catalytic serine its protease activity was impaired. rEpiP-S393A did not undergo through auto-cleavage confirmed by SDS-PAGE. This data confirms the role of serine 393 residue as catalytic residue. As predicted, mice vaccinated with EpiP were protected from staphylococcal lethal infection. Given that the protein was found to be released extracellularly, protection might be associated to EpiP antibodies blocking the function of the protein. This advocates in favor of a role of EpiP in invasive infection, as it has been shown for some of its homologs. The function of the *Streptococcus pyogenes* homolog SpyCEP in impairing the recruitment of neutrophils at the site of infection (Hidalgo-Grass, Dan-Goor et al. 2004; Zinkernagel, Timmer et al. 2008), is particularly intriguing at this regard. Given the prominent role of neutrophils against *S. aureus* infections, we are now investigating if EpiP has a similar function.

Protective efficacy of rEpiP and rEpiP-S393A was found comparable. Comparison of the crystal structure of rEpiP and rEpiP-S393A protein showed that the mutation does not alter the protein conformation providing an explanation for the protective equivalence of the two proteins. rEpiP-S393A may be better suited for vaccine development because it consists of a single polypeptide, while the rEpiP may present some issues particularly in terms of product characterization. Therefore, these observations highlight the importance of structural biology in antigen design.

CONCLUSIONS AND FUTURE PERSPECTIVES

(Chapter 1 and 2)

Mortality, morbidity, and cost from invasive *S. aureus* infections remain disturbingly high despite the introduction of several new antibiotics to treat methicillin-resistant *S. aureus* infections. Because of that designing a vaccine is a high research priority.

In our study, the reactivity of sera raised against CnaBE3 domain of SdrE with all the three Sdr proteins of different *S. aureus* strains, even in strain lacking *sdrE* showed that in strains lacking the respective gene cross-reacts with the sera and hence showed the cross-reactive potential of the highly conserved domain. Interestingly, our study also showed the cross-protective efficacy of the CnaBE3 domain of SdrE against the strain devoid of *sdrE* gene. This is encouraging as they represent conserved domain in immunogenic antigen SdrE, which may be suitable as vaccine candidate. And we know that the ideal candidate would be a conserved surface protein, stable, expressed across the different clinical strains, and evoke protective antibodies. Moreover, we also found the full-length SdrE protein to be resistant to trypsin incubation. Indeed, we found the highly conserved CnaBE3 domain partial resistant to trypsin which shows the stability characteristics of the domain. Additional testing would be required to determine if resistance is due to presence of intramolecular isopeptide bond or some other post-translational modifications. Once the structure of full-length SdrE is solved, which is in progress in collaboration with Northwestern University, Chicago, USA, an even better understanding of the stability feature of the SdrE and its spatial characteristics should emerge.

In 2010 McCarthy and Lindsay studied the relationship between surface and immune evasion gene variation and genetic backgrounds in *S. aureus*. They demonstrated that variation in genes encoding surface proteins and genes encoding secreted proteins predicted to interact with host immune responses is lineage specific. Most of the variations occurred in predicted functional domains and some surface proteins were missing or truncated in some lineages. Some domains were found to be conserved across the lineages and they concluded that successful staphylococcal vaccines should contain cocktails of antigens representing all variants. Our work led to the identification of highly conserved small domain in different clinical strains of *S. aureus*, representative members of the predominant lineages and therefore our study provides information using this conserved, stable antigen to extend the concept for vaccine development to *S. aureus*.

Furthermore, the second part of this thesis describes investigations on the novel serine protease EpiP of *S. aureus*. In this work, increased *epiP* gene expression was observed in exponential as compared to stationary growth phase. EpiP was shown to be released extracellularly and to undergo through a processing event similarly to what we observed in

vitro with recombinant EpiP. Moreover, EpiP vaccination protects mice against the challenge with *S. aureus* clinical isolates which might be associated with the fact that the protein is present extracellularly and protection is due to EpiP antibodies blocking the function of the protein. Based on protective efficacy of the wild-type EpiP and the mutant EpiP-S393A we suggest that EpiP-S393A may be better suited for vaccine development because it consists of a single polypeptide, while the wild-type may present some issues particularly in terms of product characterization. Therefore, these observations highlight the importance of the novel serine protease EpiP in *S. aureus* infection.

LIST OF PUBLICATIONS

a) **Prachi P**, Biagini M, Bagnoli B (2012) Vaccinology Is Turning into an Omics-Based Science. Drug Development Research 73: 547–558.

b) Mishra RPN, Oviedo-Orta E, **Prachi P**, Rappuoli R, Bagnoli F (2012) Vaccines and antibiotic resistance. Current Opinion in Microbiology 15(5):596–602 2.

c) **Prachi P**, Donati C, Masciopinto F, Rappuoli R & Bagnoli F (2013) Deep sequencing in pre- and clinical vaccine research. Public Health Genomics (DOI: 10.1159/000345611)

d) **Prachi P**, Kuhn M L, Biagini M, Liberatori S, Anderson W F, Bagnoli F, Grandi G (2013) Structure and protective efficacy of novel *Staphylococcus aureus* auto-cleaving protease EpiP (Manuscript under preparation)

e) Becherelli M*, **Prachi P***, Viciani E*, Biagini M, Fiaschi L, Chiarot E, Nosari S, Brettoni C, Marchi S, Biancucci M,. Fontana M.R, Grandi G, Bagnoli F, Barocchi M.A and Manetti A.G.O (2013). Protective activity of the CnaBE3 domain conserved among *Staphylococcus aureus* Sdr proteins (Manuscript under preparation)

BIBLIOGRAPHY

1. Alegre-Cebollada, J., C. L. Badilla, et al. (2010). "Isopeptide bonds block the mechanical extension of pili in pathogenic Streptococcus pyogenes." J Biol Chem **285**(15): 11235-11242.
2. Anderson, E. T., M. G. Wetherell, et al. (2002). "Processing, stability, and kinetic parameters of C5a peptidase from Streptococcus pyogenes." Eur J Biochem **269**(19): 4839-4851.
3. Andrews, T. and K. E. Sullivan (2003). "Infections in patients with inherited defects in phagocytic function." Clin Microbiol Rev **16**(4): 597-621.
4. Ariens, R. A., T. S. Lai, et al. (2002). "Role of factor XIII in fibrin clot formation and effects of genetic polymorphisms." Blood **100**(3): 743-754.
5. Arnison, P. G., M. J. Bibb, et al. (2013). "Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature." Nat Prod Rep **30**(1): 108-160.
6. Augustin, J., R. Rosenstein, et al. (1992). "Genetic analysis of epidermin biosynthetic genes and epidermin-negative mutants of Staphylococcus epidermidis." Eur J Biochem **204**(3): 1149-1154.
7. Bagnoli, F., B. Baudner, et al. (2011). "Designing the next generation of vaccines for global public health." OMICS **15**(9): 545-566.
8. Berlanda Scorza, F., F. Doro, et al. (2008). "Proteomics characterization of outer membrane vesicles from the extraintestinal pathogenic Escherichia coli DeltatolR IHE3034 mutant." Mol Cell Proteomics **7**(3): 473-485.
9. Bierbaum, G. and H. G. Sahl (2009). "Lantibiotics: mode of action, biosynthesis and bioengineering." Curr Pharm Biotechnol **10**(1): 2-18.
10. Brumfitt, W. and J. Hamilton-Miller (1989). "Methicillin-resistant Staphylococcus aureus." N Engl J Med **320**(18): 1188-1196.
11. Budzik, J. M., L. A. Marraffini, et al. (2008). "Amide bonds assemble pili on the surface of bacilli." Proc Natl Acad Sci U S A **105**(29): 10215-10220.
12. Budzik, J. M., C. B. Poor, et al. (2009). "Intramolecular amide bonds stabilize pili on the surface of bacilli." Proc Natl Acad Sci U S A **106**(47): 19992-19997.
13. Cotter, P. D., C. Hill, et al. (2005). "Bacteriocins: developing innate immunity for food." Nat Rev Microbiol **3**(10): 777-788.

14. Daly, K. M., M. Upton, et al. (2010). "Production of the Bsa lantibiotic by community-acquired Staphylococcus aureus strains." J Bacteriol **192**(4): 1131-1142.
15. DeDent, A., T. Bae, et al. (2008). "Signal peptides direct surface proteins to two distinct envelope locations of Staphylococcus aureus." EMBO J **27**(20): 2656-2668.
16. Deivanayagam, C. C., R. L. Rich, et al. (2000). "Novel fold and assembly of the repetitive B region of the Staphylococcus aureus collagen-binding surface protein." Structure **8**(1): 67-78.
17. Dierkes, L. E., C. L. Peebles, et al. (2009). "Mutational analysis of a conserved glutamic acid required for self-catalyzed cross-linking of bacteriophage HK97 capsids." J Virol **83**(5): 2088-2098.
18. Doro, F., S. Liberatori, et al. (2009). "Surfome analysis as a fast track to vaccine discovery: identification of a novel protective antigen for Group B Streptococcus hypervirulent strain COH1." Mol Cell Proteomics **8**(7): 1728-1737.
19. Drapeau, G. R., Y. Boily, et al. (1972). "Purification and properties of an extracellular protease of Staphylococcus aureus." J Biol Chem **247**(20): 6720-6726.
20. Edwards, A. M., J. R. Potts, et al. (2010). "Staphylococcus aureus host cell invasion and virulence in sepsis is facilitated by the multiple repeats within FnBPA." PLoS Pathog **6**(6): e1000964.
21. Edwards, R. J., G. W. Taylor, et al. (2005). "Specific C-terminal cleavage and inactivation of interleukin-8 by invasive disease isolates of Streptococcus pyogenes." J Infect Dis **192**(5): 783-790.
22. Engelke, G., Z. Gutowski-Eckel, et al. (1994). "Regulation of nisin biosynthesis and immunity in Lactococcus lactis 6F3." Appl Environ Microbiol **60**(3): 814-825.
23. Fleishmann, J. A., V. Mor, et al. (1995). "Longitudinal patterns of medical service use and costs among people with AIDS." Health Serv Res **30**(3): 403-424.
24. Forsgren, N., R. J. Lamont, et al. (2010). "Two intramolecular isopeptide bonds are identified in the crystal structure of the Streptococcus gordonii SspB C-terminal domain." J Mol Biol **397**(3): 740-751.
25. Francois, P., A. Scherl, et al. (2007). "Proteomic approach to investigate MRSA." Methods Mol Biol **391**: 179-199.
26. Frankel, M. B., B. M. Wojcik, et al. (2010). "ABI domain-containing proteins contribute to surface protein display and cell division in Staphylococcus aureus." Mol Microbiol **78**(1): 238-252.

27. Geissler, S., F. Gotz, et al. (1996). "Serine protease EpiP from Staphylococcus epidermidis catalyzes the processing of the epidermin precursor peptide." J Bacteriol **178**(1): 284-288.
28. Greenberg, C. S., P. J. Birckbichler, et al. (1991). "Transglutaminases: multifunctional cross-linking enzymes that stabilize tissues." FASEB J **5**(15): 3071-3077.
29. Guttilla, I. K., A. H. Gaspar, et al. (2009). "Acyl enzyme intermediates in sortase-catalyzed pilus morphogenesis in gram-positive bacteria." J Bacteriol **191**(18): 5603-5612.
30. Hagan, R. M., R. Bjornsson, et al. (2010). "NMR spectroscopic and theoretical analysis of a spontaneously formed Lys-Asp isopeptide bond." Angew Chem Int Ed Engl **49**(45): 8421-8425.
31. Hendrickx, A. P., J. M. Budzik, et al. (2011). "Architects at the bacterial surface - sortases and the assembly of pili with isopeptide bonds." Nat Rev Microbiol **9**(3): 166-176.
32. Hidalgo-Grass, C., M. Dan-Goor, et al. (2004). "Effect of a bacterial pheromone peptide on host chemokine degradation in group A streptococcal necrotising soft-tissue infections." Lancet **363**(9410): 696-703.
33. Hidalgo-Grass, C., I. Mishalian, et al. (2006). "A streptococcal protease that degrades CXC chemokines and impairs bacterial clearance from infected tissues." EMBO J **25**(19): 4628-4637.
34. Ikemura, H., H. Takagi, et al. (1987). "Requirement of pro-sequence for the production of active subtilisin E in Escherichia coli." J Biol Chem **262**(16): 7859-7864.
35. Imamura, T., S. Tanase, et al. (2005). "Induction of vascular leakage through release of bradykinin and a novel kinin by cysteine proteinases from Staphylococcus aureus." J Exp Med **201**(10): 1669-1676.
36. Izore, T., C. Contreras-Martel, et al. (2010). "Structural basis of host cell recognition by the pilus adhesin from Streptococcus pneumoniae." Structure **18**(1): 106-115.
37. Joo, H. S., G. Y. Cheung, et al. (2011). "Antimicrobial activity of community-associated methicillin-resistant Staphylococcus aureus is caused by phenol-soluble modulin derivatives." J Biol Chem **286**(11): 8933-8940.
38. Josefsson, E., K. W. McCrea, et al. (1998). "Three new members of the serine-aspartate repeat protein multigene family of Staphylococcus aureus." Microbiology **144 (Pt 12)**: 3387-3395.

39. Josefsson, E., D. O'Connell, et al. (1998). "The binding of calcium to the B-repeat segment of SdrD, a cell surface protein of Staphylococcus aureus." J Biol Chem **273**(47): 31145-31152.
40. Kang, H. J. and E. N. Baker (2009). "Intramolecular isopeptide bonds give thermodynamic and proteolytic stability to the major pilin protein of Streptococcus pyogenes." J Biol Chem **284**(31): 20729-20737.
41. Kang, H. J. and E. N. Baker (2011). "Intramolecular isopeptide bonds: protein crosslinks built for stress?" Trends Biochem Sci **36**(4): 229-237.
42. Kang, H. J., F. Coulibaly, et al. (2007). "Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure." Science **318**(5856): 1625-1628.
43. Kang, H. J., M. Middleditch, et al. (2009). "Isopeptide bonds in bacterial pili and their characterization by X-ray crystallography and mass spectrometry." Biopolymers **91**(12): 1126-1134.
44. Kies, S., C. Vuong, et al. (2003). "Control of antimicrobial peptide synthesis by the agr quorum sensing system in Staphylococcus epidermidis: activity of the lantibiotic epidermin is regulated at the level of precursor peptide processing." Peptides **24**(3): 329-338.
45. Klevens, R. M., M. A. Morrison, et al. (2007). "Invasive methicillin-resistant Staphylococcus aureus infections in the United States." JAMA **298**(15): 1763-1771.
46. Klock, H. E. and S. A. Lesley (2009). "The Polymerase Incomplete Primer Extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis." Methods Mol Biol **498**: 91-103.
47. Kuipers, O. P., H. S. Rollema, et al. (1993). "Biosynthesis and secretion of a precursor of nisin Z by Lactococcus lactis, directed by the leader peptide of the homologous lantibiotic subtilin from Bacillus subtilis." FEBS Lett **330**(1): 23-27.
48. Lieutaud, P., B. Canard, et al. (2008). "MeDor: a metaserver for predicting protein disorder." BMC Genomics **9 Suppl 2**: S25.
49. Lindenthal, C. and E. A. Elsinghorst (1999). "Identification of a glycoprotein produced by enterotoxigenic Escherichia coli." Infect Immun **67**(8): 4084-4091.
50. Lowy, F. D. (1998). "Staphylococcus aureus infections." N Engl J Med **339**(8): 520-532.
51. Lowy, F. D. (2003). "Antimicrobial resistance: the example of Staphylococcus aureus." J Clin Invest **111**(9): 1265-1273.

52. Mariotti, P., E. Malito, et al. (2013). "Structural and functional characterization of the Staphylococcus aureus virulence factor and vaccine candidate FhuD2." Biochem J **449**(3): 683-693.
53. Marraffini, L. A., A. C. Dedent, et al. (2006). "Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria." Microbiol Mol Biol Rev **70**(1): 192-221.
54. Marraffini, L. A., A. C. DeDent, et al. (2006). "Sortases and the Art of Anchoring Proteins to the Envelopes of Gram-Positive Bacteria." Microbiology and Molecular Biology Reviews **70**(1): 192-221.
55. Marsh, A. J., O. O'Sullivan, et al. (2010). "In silico analysis highlights the frequency and diversity of type 1 lantibiotic gene clusters in genome sequenced bacteria." BMC Genomics **11**: 679.
56. Maskalyk, J. (2002). "Antimicrobial resistance takes another step forward." CMAJ **167**(4): 375.
57. Miajlovic, H., A. Loughman, et al. (2007). "Both complement- and fibrinogen-dependent mechanisms contribute to platelet aggregation mediated by Staphylococcus aureus clumping factor B." Infect Immun **75**(7): 3335-3343.
58. Mickelson, M. N. (1964). "Chemically Defined Medium for Growth Streptococcus Pyogenes." J Bacteriol **88**: 158-164.
59. Millard, C. S., L. Stols, et al. (2003). "A less laborious approach to the high-throughput production of recombinant proteins in Escherichia coli using 2-liter plastic bottles." Protein Expr Purif **29**(2): 311-320.
60. Mishra, R. P., P. Mariotti, et al. (2012). "Staphylococcus aureus FhuD2 is involved in the early phase of staphylococcal dissemination and generates protective immunity in mice." J Infect Dis **206**(7): 1041-1049.
61. Miyafusa, T., J. M. Caaveiro, et al. (2012). "Crystal structure of the enzyme CapF of Staphylococcus aureus reveals a unique architecture composed of two functional domains." Biochem J **443**(3): 671-680.
62. Osicka, R., K. Prochazkova, et al. (2004). "A novel "clip-and-link" activity of repeat in toxin (RTX) proteins from gram-negative pathogens. Covalent protein cross-linking by an Asp-Lys isopeptide bond upon calcium-dependent processing at an Asp-Pro bond." J Biol Chem **279**(24): 24944-24956.
63. Otto, M. and F. Gotz (2001). "ABC transporters of staphylococci." Res Microbiol **152**(3-4): 351-356.

64. Palumbo, E., L. Fiaschi, et al. (2012). "Antigen identification starting from the genome: a "Reverse Vaccinology" approach applied to MenB." Methods Mol Biol **799**: 361-403.
65. Panula-Perala, J., J. Siurkus, et al. (2008). "Enzyme controlled glucose auto-delivery for high cell density cultivations in microplates and shake flasks." Microb Cell Fact **7**: 31.
66. Patti, J. M. (2011). "Will we ever see the approval of a Staphylococcus aureus vaccine?" Expert Rev Anti Infect Ther **9**(10): 845-846.
67. Pickart, C. M. (2001). "Mechanisms underlying ubiquitination." Annu Rev Biochem **70**: 503-533.
68. Pizza, M., V. Scarlato, et al. (2000). "Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing." Science **287**(5459): 1816-1820.
69. Potempa, J., A. Dubin, et al. (1988). "Degradation of elastin by a cysteine proteinase from Staphylococcus aureus." J Biol Chem **263**(6): 2664-2667.
70. Prasad, L., Y. Leduc, et al. (2004). "The structure of a universally employed enzyme: V8 protease from Staphylococcus aureus." Acta Crystallogr D Biol Crystallogr **60**(Pt 2): 256-259.
71. Rappuoli, R. (2000). "Reverse vaccinology." Curr Opin Microbiol **3**(5): 445-450.
72. Rappuoli, R. (2001). "Reverse vaccinology, a genome-based approach to vaccine development." Vaccine **19**(17-19): 2688-2691.
73. Rasmussen, R. V., V. G. Fowler, Jr., et al. (2011). "Future challenges and treatment of Staphylococcus aureus bacteremia with emphasis on MRSA." Future Microbiol **6**(1): 43-56.
74. Ravipaty, S. and J. P. Reilly (2010). "Comprehensive characterization of methicillin-resistant Staphylococcus aureus subsp. aureus COL secretome by two-dimensional liquid chromatography and mass spectrometry." Mol Cell Proteomics **9**(9): 1898-1919.
75. Rawlings, N. D. and A. J. Barrett (1994). "Families of serine peptidases." Methods Enzymol **244**: 19-61.
76. Roche, F. M., R. Massey, et al. (2003). "Characterization of novel LPXTG-containing proteins of Staphylococcus aureus identified from genome sequences." Microbiology **149**(Pt 3): 643-654.

77. Rodriguez-Ortega, M. J., N. Norais, et al. (2006). "Characterization and identification of vaccine candidate proteins through analysis of the group A Streptococcus surface proteome." Nat Biotechnol **24**(2): 191-197.
78. Sabat, A., K. Kosowska, et al. (2000). "Two allelic forms of the aureolysin gene (aur) within Staphylococcus aureus." Infect Immun **68**(2): 973-976.
79. Schnell, N., G. Engelke, et al. (1992). "Analysis of genes involved in the biosynthesis of lantibiotic epidermin." Eur J Biochem **204**(1): 57-68.
80. Schnell, N., K. D. Entian, et al. (1988). "Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with four sulphide-rings." Nature **333**(6170): 276-278.
81. Sharp, J. A., C. G. Echague, et al. (2012). "Staphylococcus aureus surface protein SdrE binds complement regulator factor H as an immune evasion tactic." PLoS One **7**(5): e38407.
82. Shaw, L., E. Golonka, et al. (2004). "The role and regulation of the extracellular proteases of Staphylococcus aureus." Microbiology **150**(Pt 1): 217-228.
83. Shorr, A. F. (2007). "Epidemiology and economic impact of meticillin-resistant Staphylococcus aureus: review and analysis of the literature." Pharmacoeconomics **25**(9): 751-768.
84. Siezen, R. J., O. P. Kuipers, et al. (1996). "Comparison of lantibiotic gene clusters and encoded proteins." Antonie Van Leeuwenhoek **69**(2): 171-184.
85. Sinha, B., P. Francois, et al. (2000). "Heterologously expressed Staphylococcus aureus fibronectin-binding proteins are sufficient for invasion of host cells." Infect Immun **68**(12): 6871-6878.
86. Sitkiewicz, I., I. Babiak, et al. (2011). "Characterization of transcription within sdr region of Staphylococcus aureus." Antonie Van Leeuwenhoek **99**(2): 409-416.
87. Smith, L. and J. Hillman (2008). "Therapeutic potential of type A (I) lantibiotics, a group of cationic peptide antibiotics." Curr Opin Microbiol **11**(5): 401-408.
88. Striebel, F., F. Imkamp, et al. (2009). "Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes." Nat Struct Mol Biol **16**(6): 647-651.
89. ten Broeke-Smits, N. J., T. E. Pronk, et al. (2010). "Operon structure of Staphylococcus aureus." Nucleic Acids Res **38**(10): 3263-3274.
90. Ton-That, H., L. A. Marraffini, et al. (2004). "Protein sorting to the cell wall envelope of Gram-positive bacteria." Biochim Biophys Acta **1694**(1-3): 269-278.

91. Turner, D. P., K. G. Wooldridge, et al. (2002). "Autotransported serine protease A of Neisseria meningitidis: an immunogenic, surface-exposed outer membrane, and secreted protein." Infect Immun **70**(8): 4447-4461.
92. Ward, J. J., L. J. McGuffin, et al. (2004). "The DISOPRED server for the prediction of protein disorder." Bioinformatics **20**(13): 2138-2139.
93. Wikoff, W. R., L. Liljas, et al. (2000). "Topologically linked protein rings in the bacteriophage HK97 capsid." Science **289**(5487): 2129-2133.
94. Willey, J. M. and W. A. van der Donk (2007). "Lantibiotics: peptides of diverse structure and function." Annu Rev Microbiol **61**: 477-501.
95. Xue, H., H. Lu, et al. (2011). "Sequence diversities of serine-aspartate repeat genes among Staphylococcus aureus isolates from different hosts presumably by horizontal gene transfer." PLoS One **6**(5): e20332.
96. Zhang, L., K. Jacobsson, et al. (1998). "A second IgG-binding protein in Staphylococcus aureus." Microbiology **144 (Pt 4)**: 985-991.
97. Zhou, X., K. Szeker, et al. (2013). "Recombinant purine nucleoside phosphorylases from thermophiles: preparation, properties and activity towards purine and pyrimidine nucleosides." FEBS J.
98. Zingaretti, C., F. Falugi, et al. (2010). "Streptococcus pyogenes SpyCEP: a chemokine-inactivating protease with unique structural and biochemical features." FASEB J **24**(8): 2839-2848.
99. Zinkernagel, A. S., A. M. Timmer, et al. (2008). "The IL-8 protease SpyCEP/ScpC of group A Streptococcus promotes resistance to neutrophil killing." Cell Host Microbe **4**(2): 170-178.