Staphylococcus aureus transpeptidase sortase: Development of new substrates and products

THESIS SUBMITTED TO

JAWAHARLAL NEHRU UNIVERSITY

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Sharmishtha Samantaray

National Institute of Immunology

New Delhi

2008

राष्ट्रीय प्रतिरक्षाविज्ञान संस्थान
NATIONAL INSTITUTE OF IMMUNOLOGY

Certificate

This is to certify that the thesis entitled "***Staphylococcus aureus* transpeptidase sortase: Development of new substrates and products**" by **Sharmishtha Samantaray**, in partial fulfillment of the Ph.D. degree from the **Jawaharlal Nehru University**, New Delhi, embodies the work done by the candidate under my guidance at the National Institute of Immunology. The work is original and has not been submitted in part or in full for any other degree or diploma of any university.

RP Roy

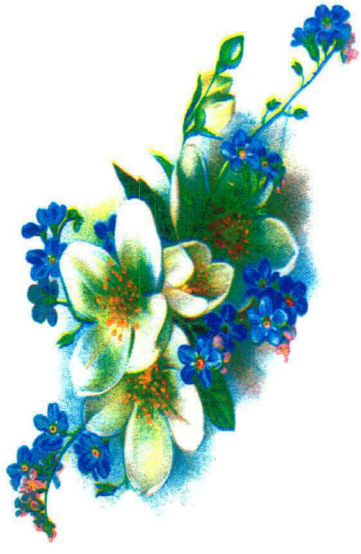
Dr. Rajendra P. Roy

(Thesis Supervisor)

July 29, 2008

अरूणा आसफ अली मार्ग, नई दिल्ली - 110067 फैक्स : 91-11-26162125, 26177626 तार : IMMUNOLOGY
Aruna Asaf Ali Marg, New Delhi - 110067 Fax : 91-11-26162125, 26177626 Grams : IMMUNOLOGY
दूरभाष (Telephones) EPABX : 26717121 - 45 & 26717010 - 19

Dedicated to
my Parents
and
my Teachers



My deepest gratitude ...

On paper 'five years' sounds like such a long time....half a decade....but today as I stand at the threshold of the beginning of the end, five years seems like something that passed by so quickly, it was over before it even began!...and now as it's time to move on and move away, I would like to express my heartfelt thanks towards all those who made this journey such an astonishing and extraordinary experience.

Above everything and everyone, I would like to thank my thesis supervisor Dr. R P Roy. It was my good fortune to have worked under his tutelage and what I owe him cannot be demonstrated through deed or words. His insight and encouragement have been the guiding light in my endeavors for these last five years. His optimism and enthusiasm has always encouraged me to try that little bit harder, go that extra mile, to strive for nothing short of excellence. I thank you Sir, for guiding me throughout, while giving me the autonomy to pursue my own ideas so I could test myself. I will consider myself fortunate if I can carry forward even a fraction of your scientific vision and acumen.

I will always be indebted to Dr. Vinay Nandicoori for taking out time from his busy schedule to instruct me. His endless optimism, infectious enthusiasm and firm belief that everything will work out when it looks like nothing will, were instrumental in the completion of this work. Thank you Sir for always making time to clear the smallest of my doubts and for listening to all my absurd ideas and opinions with a lot more patience and respect than they might have deserved.

I would like to express my deepest gratitude towards my Doctoral Committee members: Dr. Devapriya Choudhury, Dr. Sudhanshu Vrati, Dr. Devinder Sehgal and Dr. Ayub Qadri for their review, constructive criticism and excellent advice in the evolution of this thesis.

My sincere thanks to our ex-Director, Dr. Sandip Basu and our present Director, Dr. Avadesha Surolia, for ensuring that every opportunity and facility was made available for the successful completion this work.

The help rendered by the non-scientific staff is gratefully acknowledged. Mr. Akhilesh, Mr. Rana and Sanju ma'am deserve special thanks for liberating me of the nitty-gritty's of academic paperwork. I would also like to take this opportunity to express my gratitude

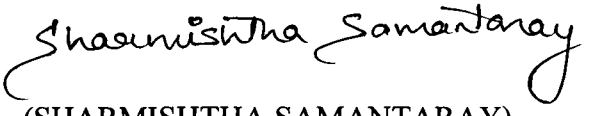
towards Mr. Dahra and Mr. Dey for the innumerable occasions where they have gone out of their way to render help. Thank you Satish ji for ensuring the thesis got bound and submitted on time.

A degree is immense hard work, but it gets simplified when your colleagues are your friends. I consider myself fortunate in having had some of the best people around as lab mates. It was a pleasure to work alongside seniors like Lavanya and Srijita; their friendship made the lab atmosphere so much fun that work ceased to feel like work at all! I will always be indebted to them (and also to Pramod) for guiding my baby steps during the initial few days in the lab. Uttara, I treasure memories of our legendary 'HPLC wars' as much as I value our newfound friendship. I will never forget how your crazy sense of humor had everyone rolling in the aisles.....and will always associate home-delivery from 'Rainbows' with you!! Sayani, words are inadequate for expressing how much you mean to me. Thanks for the innumerable occasions where you have covered up for me, made my excuses, picked up the slack....your friendship is something I will treasure forever. Kanchan, Tora and Charu, thank you for always being ready with help and sound advice. Special thanks to Amit, Divya, Shazia, Kalpana, Rajneesh, Devanand and Kasturi for welcoming me so graciously into their lab.

Rohit, Priti, Soma, Jayita, Beena, Divya....I will always remember the times we have spent together...fun-filled moments with you have been the perfect antidote to the PhD stress and tension. All the trips, treats, the all-night-chats will be part of my most fond memories for a long time to come. Harshita, Arush, Tarun and Smriti, words are not enough to express what you guys mean to me....not a day goes by when I don't thank my stars for blessing me with such wonderful friends....just the knowledge of my friends beside me has made every disappointment bearable and magnified every joy ten-fold . This camaraderie will be one of the most valuable things I will take with me from NII.

Lastly, and most importantly, I would like to thank my family for their unconditional love and support. My father, for always leading by example; his code of behavior, values and principles have been a constant source of inspiration and courage for me. I cannot thank my mother enough for her warm and caring nature, and for putting up with my erratic schedule and uncertain temperament with a loving smile; and my little sister, for being ever-willing in lending a helping hand and a ready shoulder to lean on.

This thesis would have been unimaginable had it not been for the unequivocal support of all these people and many others whom I might not have been able to name here, but feel deeply indebted towards.



(SHARMISHTHA SAMANTARAY)

Contents

Abbreviations

µl/ml	micro/milli liter
γ-^{32}PATP	P^{32} labeled Adenosine Triphosphate
β-ME	β-Mercaptoethanol
2-DOS	2-Deoxystreptamine
4A2HBA	4-amino-2-hydroxy-butyric acid
4A3HBA	4-amino-3-hydroxy-butyric acid
AAEK	Aryl(β-amino)ethyl Ketone
Abz	Aminobenzoic Acid
ACN	Acetonitirile
AIDS	Acquired Immunodeficiency Syndrome
APS	Ammonium Persulfate
ATCC	American Type Culture Collection
BPB	Bromophenol Blue
Calc	Theoretically Calculated Mass
CB	Cross Bridges
Clf	Clumping Factor
CM	Cell Membrane
CPP	Cationic Polypeptide
CW	Cell Wall
CWS	Cell Wall Sorting
CY	Cytoplasm
DCM	Dichloromethane
DEPC	Diethylenepyrocarbonate
D-iGlu	D-isoGlutamic Acid
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
Dnp	Dinitrophenyl
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
Efe	Enterococcus faecalis
EMSA	Electrophoretic Mobility Gel Shift Assay
Env	Envelope protein
ESI-TOF	Electrospray Ionization-Time of Flight

ESMS	Electrospray Mass Spectrometry
EtBr	Ethidium Bromide
Exp	Experimentally Determined Mass
Fmoc	9-Fluoromethoxycarbonyl
Fnbp	Fibronectin Binding Protein
FRET	Florescence Resonance Energy Transfer
Fur	Ferric Uptake Regulator
GFP	Green Florescent Protein
GlcNAc	N-Acetylglucosamine
GW	Glycine-Tryptophan
HBI	Heart Brain Infusion
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic Acid
HIV-1	Human Immunodeficiency Virus-1
hr/hrs	Hour (s)
IC	Inhibitory Concentration
Ig	Immunoglobulin
IMAC	Immobilized Metal Ion Affinity Chromatography
IPTG	Isopropyl-β-D-Galactopyranoside
isd	Iron-Responsive Surface Determinant
Kb/bp	Kilo Base/Base Pair
LB	Luria Bertani
LD	Lethal Dose
LTA	Lipotechoic Acid
LTR	Long Terminal Repeat
MALDI-TOF	Matrix Assisted Laser Desorption/Ionization-Time of Flight
m-DPM	*meso*-Diaminopimelic Acid
min	Minute
MN	N-Acetylmuramic Acid
Mrp	Multidrug Resistance Protein (residues 1-202)
Mrp_{1-170}	SrtA cleaved Mrp (residues 1-170)

$Mrp_{171-202}$	SrtA cleaved Mrp (residues 171-202)
MRSA	Methicillin Resistant *S. aureus*
MSCRAMM	Microbial Surface Components Recognizing Adhesive Matrix Molecules
MurNAc	N-Acetylmuramic acid
MWCO	Molecular Weight Cut Off
NCBI	National Center for Biotechnology Information
NEAT	Near Transporter
ng/μg/mg	Nano/Micro/Milli Gram
Ni-NTA	Nickel-Nitrilotriacetic Acid
NMR	Nuclear Magnetic Resonance
nt	Nucleotide
OM	Outer Membrane
PAGE	Polyacrylamide Gel Electrophoresis
PBP	Pennicillin Binding Protein
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PEI	Polyethylene Imine
PEP	Phosphoenol Pyruvate
PM	Plasma Membrane
PNA	Peptide Nucleic Acid
PNK	Polynucleotide Kinase
PTD	Protein Transduction Domain
RBD	RNA Binding Domain
Rev	Regulator of Virion Expression
RNA	Ribonucleic Acid
RP-HPLC	Reverse Phase High Performance Liquid Chromatography
rpm	Revolutions Per Minute
RRE IIB	47nt High Affinity Rev Binding Site of RRE
RRE	Rev Response Element
s	Second
Sas	*S. aureus* Surface Proteins

SC	Spore Coat
Sdr	Serine-Aspartate Repeat Containing Protein
SDS	Sodium Dodecyl Sulfate
Spa	*S. aureus* Protein A
SpaA	Sortase Mediated Pilin Assembly Protein A
Spn	*Streptococcus pneumoniae*
SPPS	Solid Phase Peptide Synthesis
Spy	*Streptococcus pyogenes*
SrtA	*Staphylococcus aureus* Sortase A
$SrtA_{\Delta 59}$	Amino terminal 59 residues deleted version of SrtA
$SrtA_{\Delta N}$	Amino terminal signal peptide/transmembrane sequence deleted version of SrtA
SrtB	*Staphylococcus aureus* Sortase B
$SrtB_{\Delta 30}$	Amino terminal 30 residues deleted version of SrtB
SrtC	Sortase C
SrtD	Sortase D
TA	Techoic Acid
TAE	Tris-Acetate-EDTA
TAR	Transactivation Response
Tat	Transactivator of Transcription
TBE	Tris-Borate-EDTA Buffer
TEMED	Tetramethylethylenediamine
TFA	Trifluroacetic Acid
TNFR	Tumor Necrosis Factor Receptor
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
UMP/UDP/UTP	Uridine Mono/Di/Tri Phosphate
UV	Ultraviolet
V	Volt
v/v	Volume by Volume
w/v	Weight by Volume
WP	Wall Peptide

X-gal	5-bromo-4-chloro-3-indoyl β-D-galactopyranoside

Introduction

The cell wall of gram positive bacteria is host to a wide variety of molecules and serves a multitude of functions most of which are critical to the viability of the cell. The primary purpose of the cell wall is to provide a rigid exoskeleton for protection against both mechanical and osmotic lysis. It also serves as an attachment site for proteins that interact with the bacterial environment, particularly with the host, where the proteins aid the initiation and propagation of infection via adherence to host tissue and immune system evasion[1,2,3].

The structure of gram-positive bacterial cell wall

The cell wall of gram-positive bacteria is a peptidoglycan macromolecule with attached accessory molecules such as teichoic acid, teichuronic acid and muramic acid[4]. The glycan strands of the cell wall consist of repeating units of the disaccharide N-acetylmuramic acid-(β1-4)-N-acetylglucosamine (MurNAc-GlcNAc) (Figure 1)[5,6]. In most cases, the D-lactyl moiety of each MurNAc is linked to a short peptide component through an amide bond[7-9]. Such wall peptides (WP) are in turn cross-linked with other peptides called cross bridges (CB) that are attached to neighboring glycan strands. This generates a three dimensional molecular network that surrounds the cell and provides the desired exoskeletal function[10-13].

While the repeating disaccharide MurNAc-GlcNAc unit is found in all bacterial peptidoglycans, WP and CB composition varies between species. In *Staphylococcus aureus*, the WP attached to MurNAc is composed of the amino acids L-Ala, D-iGlu, L-Lys, and D-Ala, and a pentaglycine moiety functions as the CB. On the other hand, in *Listeria monocytogens*, the L-Lys is replaced by *meso*-diaminopimelic acid (*m*-DPM) and *m*-DPM itself acts as the CB[14].

The synthesis of gram-positive bacterial cell wall

Cell wall synthesis in gram-positive bacteria can be divided into three separate stages that occur in distinct subcellular compartments, the cytoplasm, the membrane, and, finally, the cell wall itself (Figure 2)[15-17]. In *Staphylococcus aureus*, peptidoglycan synthesis begins by generating UDP-MurNAc from UDP-GlcNAc and phosphoenolpyruvate (PEP)[18]. Five amino acids, L-Ala, D-isoGlu, L-Lys, and the D-Ala-D-Ala dipeptide, are linked to UDP-MurNAc in four consecutive steps, to generate UDP-MurNAc-pentapeptide or Park's nucleotide[19-23]. Park's nucleotide is

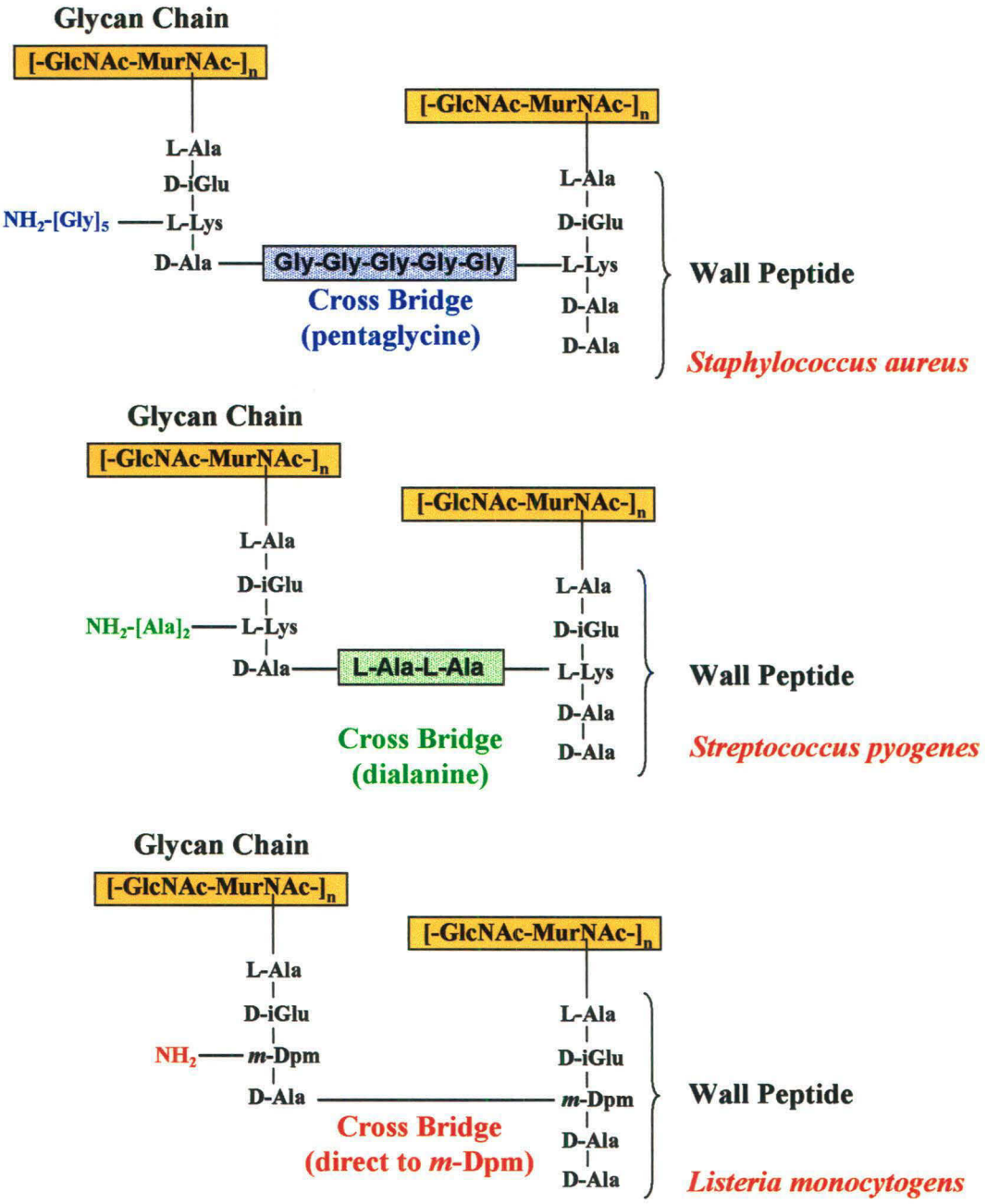


Figure 1. Peptidoglycan structures from *S. aureus*, *S. pyogenes*, and *L. monocytogenes*[15]. Glycan chains are linked to short wall peptides through the lactyl moiety of MurNAc. Adjacent wall peptides can be linked through crossbridge peptides (pentaglycine in *S. aureus* or dialanine in *S. pyogenes*). In some cases, such as *L. monocytogenes*, adjacent wall peptides are linked via an amide bond between the ε-amino group of a *m*-Dpm residue at position 3 of one wall peptide and the carboxy terminus of the D-Ala residue at position 4 of the adjacent wall peptide.

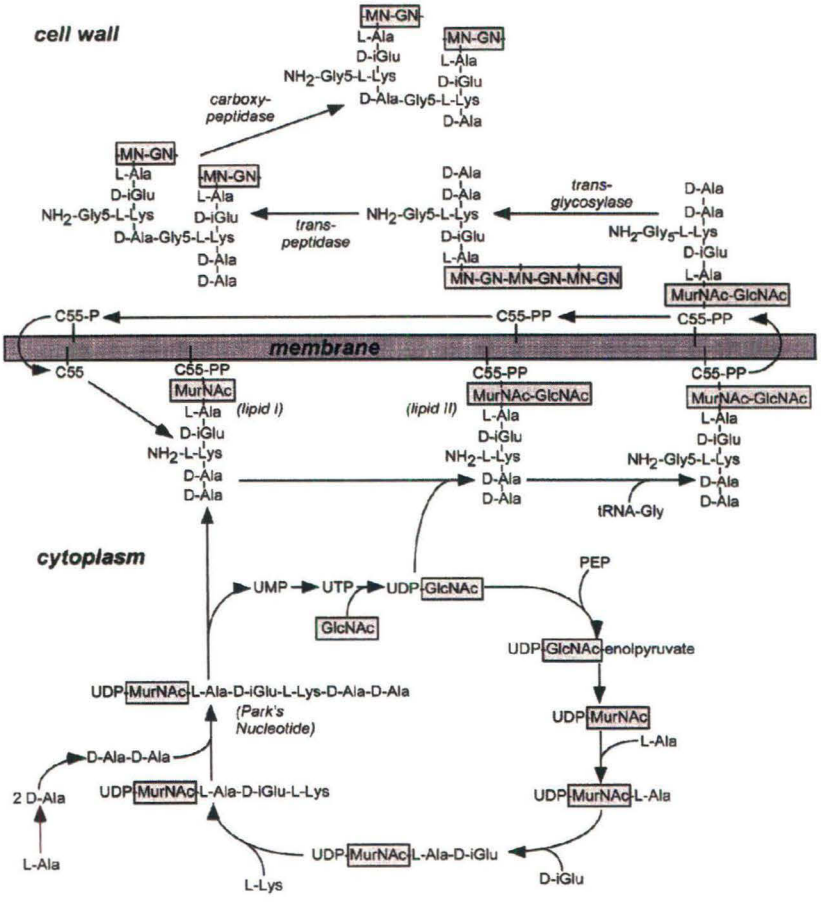


Figure 2. Biosynthetic pathway of cell wall assembly[15]. The generation of cell wall precursors begins in the cytoplasm, resulting in the synthesis of Park's nucleotide which is then transferred to a lipid carrier in the membrane to generate lipid I. After further modification, the lipid-anchored peptidoglycan precursor is translocated to the extracellular face of the cytoplasmic membrane. The peptidoglycan precursor is subsequently incorporated into the cell wall by transglycosylation and transpeptidation reactions with the concomitant displacement of the lipid carrier. The terminal D-Ala of the wall pentapeptide is subject to substitution by cross-linking to other wall subunits or can be removed by the action of a D-alanyl-D-alanine carboxypeptidase.
MN: N-acetylmuramic acid; GN: N-acetylglucosamine; D-iGlu: D-isoglutamic acid; UMP/UDP/UTP: uridine mono/di/tri phosphate; PEP: phosphoenol pyruvate; Gly_5: pentaglycine.

then linked to an undecaprenyl-pyrophosphate carrier molecule by phosphodiester bonds to generate C_{55}-PP-MurNAc-L-Ala-D-isoGlu-L-Lys-D-Ala-D-Ala, or lipid I. UDP-GlcNAc is linked to the muramoyl moiety of lipid I to generate the disaccharide lipid II precursor [C_{55}-PP-MurNAc(-L-Ala-D-isoGlu-L-Lys(Gly_5)-D-Ala-D-Ala)-β1-4-GlcNAc][24-26]. Lipid II is further modified by the addition of a pentaglycine moiety to the ε-amino of lysine[27, 28]. Finally, the modified lipid II-pentaglycine precursor is translocated across the cytoplasmic membrane to serve as the substrate for peptidoglycan assembly.

Cell wall assembly of peptidoglycan is catalyzed by penicillin binding proteins (PBPs)[29]. PBPs promote the polymerization of glycan from its disaccharide precursor, i.e., the successive addition of MurNAc(-L-Ala-D-isoGlu-L-Lys-D-Ala-D-Ala)-GlcNAc to C_{55}-PP-MurNAc(-L-Ala-D-isoGlu-L-Lys-D-Ala-D-Ala)-GlcNAc, and the transpeptidation (cross-linking) of wall peptides[30]. The latter reaction results in the proteolytic removal of the D-Ala at the C-terminal end of the pentapeptide and the formation of a new amide bond between the amino group of the crossbridge and the carbonyl group of D-Ala at position 4[13].

Importance of bacterial cell-surface proteins

As mentioned earlier, the cell wall also serves as a site of attachment of a large number of proteins. As these proteins form the primary interface between the host and the pathogen, most of them play an indispensable part in the establishment and propagation of infectious disease. Such proteins enact a wide variety of roles; they function as (i) **transporters** that allow selective passage of molecules and exclude harmful substances[31, 32]; (ii) **adhesins** that attach or adhere to specific surfaces or tissues[1, 33, 34]; (iii) **enzymes** that mediate specific reactions on the cell surface important for the survival of the organism[35]; (iv) **protective structures** to prevent phagocytic engulfment or killing[36, 37]; (v) **antigenic disguises** to bypass activation of host immune defenses[38]; (vi) **endotoxins** that cause an inflammatory response in the host; and (vii) **receptors** that can respond to temperature, osmolarity, salinity, light, oxygen, nutrients, etc., resulting in a molecular signalling cascade in the cell[39, 40].

Mechanism of attachment of bacterial cell-surface proteins

To date, five major mechanisms for cell-surface display of proteins have been elucidated (Figure 3). (i) **Choline binding proteins**, like LytA of *Streptococcus pneumoniae*, bind to choline-substituted teichoic acid (TA) or lipotechoic acid (LTA) via recognition motif present as repeat sequences (20 residues repeat) at the C-terminus[41, 42]; (ii) **Membrane anchored proteins**, like ActA of *L. monocytogenes*, are anchored to the cell membrane by an anchor sequence composed of positively charged and hydrophobic residues[43]; (iii) **lipoproteins**, like BlaZ of *Staphylococcus aureus*, have an N-terminal lipoylated cysteine that is responsible for tethering the protein to the cell-membrane[44, 45]; (iv) **LTA binding proteins**, like InlB of *Listeria*, are bound to membrane LTA through 80-amino acid tandem repeats that start with the dipeptide Gly-Trp (GW repeats)[46, 47]; (v) **LPXTG proteins**, exemplified by Protein A of *S. aureus*, are covalently linked to the cell wall peptidoglycan by the action of a transpeptidase called Sortase A. Sortase anchored proteins are the only surface proteins known to be covalently linked to the cell wall[48].

Sortase catalyzed attachment of proteins to bacterial cell wall

A large number of surface proteins are anchored to the cell wall by the action of the sortases, a recently discovered family of membrane-associated cysteine transpeptidases[49]. Sortase enzymes are characterized by an N-terminal secretory signal and a conserved TLXTC sequence motif towards the C-terminus. This motif contains the catalytic cysteine which is indispensable for transpeptidation activity[50-52].

Surface proteins anchored by this mechanism have special recognition sequence motifs at their C-terminus which are recognized, cleaved and thereafter attached to the cell wall by sortases. For example, *Staphylococcus aureus* Sortase A (SrtA), which is the prototypical sortase, recognizes a C-terminal LPXTG motif. It cleaves the motif at the Thr-Gly bond and attaches the –COOH group of the Thr to the terminal amine of the pentaglycine CB of staphylococcal peptidoglycan[53].

Sortase-catalyzed cell surface attachment of proteins is a universal phenomenon as sortase homologues have been identified in all gram-positive bacteria examined till date including sporulating gram-positives like Clostridium and

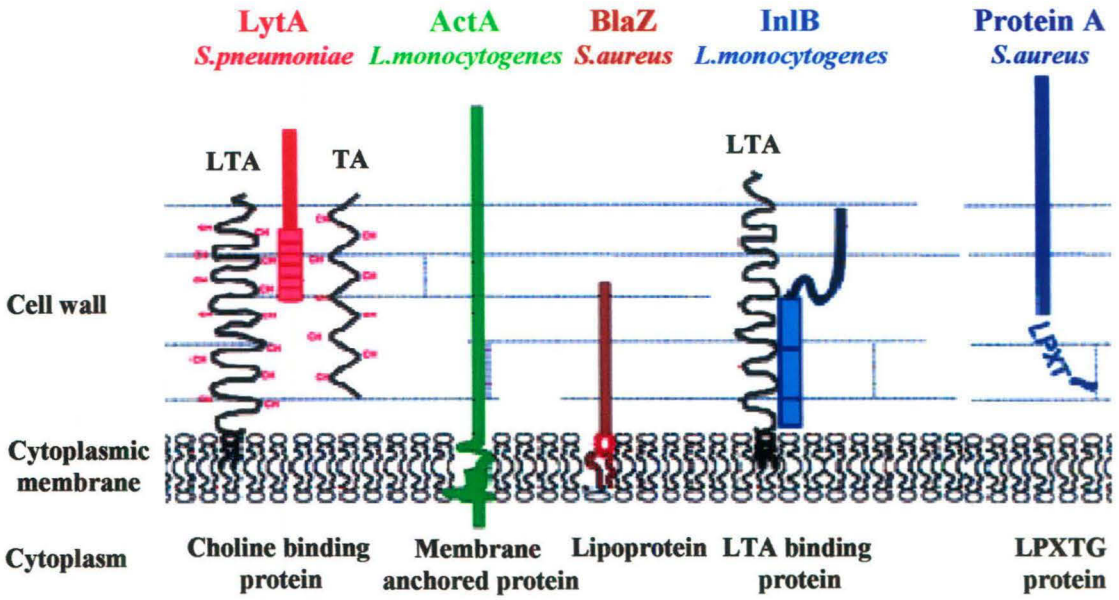


Figure 3. Major types of surface proteins in gram-positive bacteria[48].

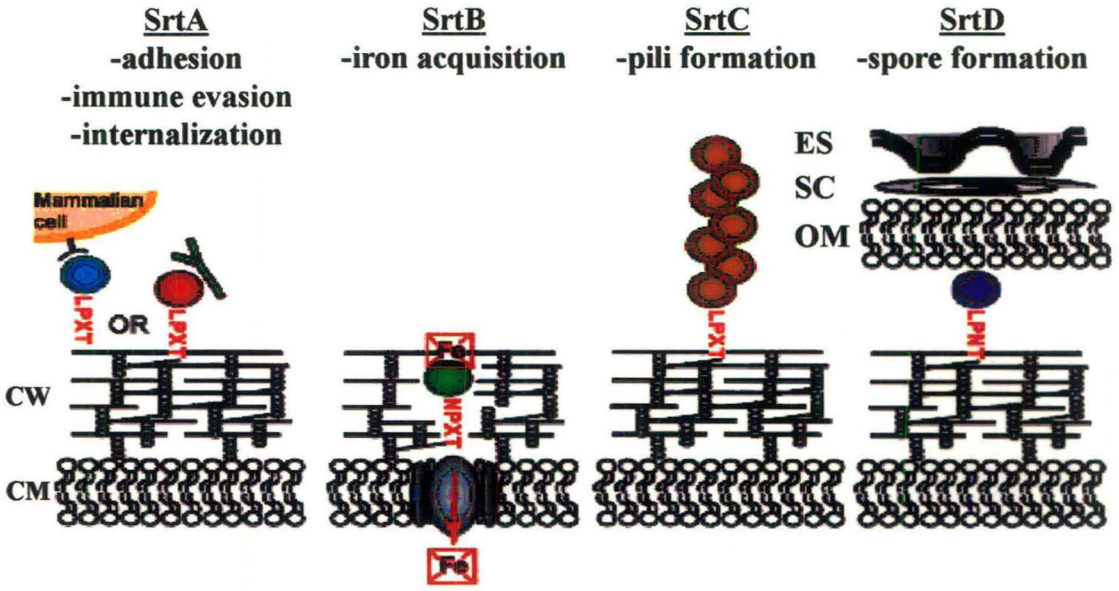


Figure 4. Diversity of sortase biology[80]. CM, cell membrane; CW, cell wall; OM, outer membrane; SC, spore coat; ES, exosporium.

Bacillus. Most bacterial genomes encode more than one sortase like protein with some, like *Bacillus cereus* and *Streptomyces coelicolor*, encoding as many as seven[54-56]. Sortases, however, seem to have been lost entirely from the two groups of gram-positives that have evolved furthest away from the archetypal gram-positive cell wall - the mycobacteria, with its highly complex cell wall and the mycoplasmas, which lack a cell wall. Curiously, sortase homologues have also been found in a few gram-negative organisms like *Shewanella putrefasciens*; their significance, however, remains to be elucidated[55].

Types of Sortases

The large number of newly identified sortase analogues necessitated the development of a classification system. Till date, two attempts at cataloguing sortases have been reported. Dramsi and coworkers[57] analyzed complete genome sequences from 22 gram-positive bacterial species and found 61 sortase homologues which they partitioned into four classes called sortase A, B, C and D. This partitioning was based on differences in their primary structure as well as the motifs that they recognized in their target proteins (Figure 4).

In another report, Comfort and coworkers[56] analyzed 72 bacterial genomes and found a total of 176 sortase homologues with 42 genomes encoding at least two enzymes. They classified the sortase homologues into 5 classes: sortase A, sortase B, subfamily C, D and E. Their basis of classification was also differences in primary structure and recognition motifs.

The sortase A and B families are in common in the two systems; subfamily 3 corresponds to class C and subfamily 5 to class D. The additional subfamily 4 described by Comfort and coworkers clusters in the class D family of the Dramsi system.

The sortase A enzymes, including staphylococcal Sortase A (SrtA), harbor an N-terminal signal peptide cum transmembrane domain and the signature catalytic TLXTC sequence near the C-terminus. Bacteria encode only a single sortase A homolog and genes encoding sortase A are never found clustered within the same locus as their substrates. An analysis of various substrates suggests that sortase A enzymes are specific for the LPXTG motif and are responsible for anchoring a

majority of LPXTG containing surface proteins most of which play important roles in bacterial adhesion, invasion, immune evasion etc.

The sortase B enzymes, including the archetype Sortase B (SrtB) form *Staphylococcus aureus*, also harbor an N-terminal signal peptide/transmembrane domain and the C-terminal TLXTC motif. *S. aureus* SrtB is required for membrane anchoring of IsdC, an NPQTN motif containing protein that is important for iron acquisition. Other motifs recognized by SrtB homologues are NAKTN (or NPKSS) in *L. monocytogenes*, NPQTG/NSKTA in *B. halodurans*, and NPKTG/NSKTA in *B. anthracis* and *B. cereus*[58]. Genes encoding members of this class and its putative targets are often part of the same genomic locus. Analysis of substrates reveals that sortase B proteins appear to be involved in iron metabolism as they contain the NEAT (near transporter) domain implicated in iron transport[59].

The sortase C is the largest class of sortases and is characterized by the presence of a C-terminal hydrophobic membrane that is hypothesized to serve as the membrane anchor (Type I anchored membrane proteins) unlike the rest of the sortase family that contain N-terminal membrane anchoring sequences and are type II proteins. Class C sortases also recognize and anchor LPXTG containing proteins and several genomes have been shown to contain both sortase A and sortase C homologues. Studies on *Streptococcus pyogenes* indicate that despite possessing similar sorting signals the two sortase classes function in a non-redundant fashion in the cell and differentially sort proteins bearing similar motifs[60]. As class C sortases are always found clustered with their potential substrates, coordinated expression may be one way by which this substrate selectivity is achieved. It has been shown that assembly of pili on the surface of *Corynebacterium diptheriae* requires class C sortases[61]. Homologues of the gene encoding the major pilin subunit SpaA and accompanying sortase machinery required for covalent tethering of pilin subunits also seem to exist in *S. pneumoniae*, *E. faecalis* and *S. agalactiae*[62, 63]. In *Actinomycetes*, class C sortases have been found to play a role in the assembly of fimbriae, composed of LPXTG subunit proteins, at the bacterial surface[64, 65]. They have also been implicated in the formation of aerial hyphae in *S. coelicolor*[66].

Sortase D class is composed of homologues that recognize a non-standard sorting signal, LAXTG. Similar to sortase A, sortase D genes are never positioned adjacent to their predicted substrates. The function of LAXTG containing proteins

remains to be elucidated although they have been implicated in bacterial sporulation and the developmental programs of bacilli or streptomyces[66, 67].

Most gram-positive bacteria encode sortases from two or more classes and studies indicate that there is no substrate overlap and that sortases of different classes function non-redundantly in sorting proteins to the cell surface. Class A sortases appear to be responsible for anchoring a majority of surface proteins in gram-positive bacteria and hence, are often referred to as 'housekeeping' sortases. Sortases belonging to other classes are predicted to play a more specialized role, anchoring on average far fewer proteins that frequently contain unusual sequence motifs in their sorting signals.

Staphylococcus aureus Sortase A (SrtA)

Domain structure of SrtA and its substrates

Staphylococcus aureus SrtA, the archetypical sortase, is a 206 amino acid protein with an N-terminal signal peptide *cum* membrane anchor sequence and a C-terminal TLXTC motif (Figure 5). It is a typical Type II membrane anchored protein, i.e., it is anchored to the cell membrane via an N-terminal hydrophobic sequence such that its C-terminus, which contains the active site, faces the extracellular milieu[53]. Deletion of the first 59 amino acids ($SrtA_{\Delta 59}$) provides for a soluble and catalytically competent enzyme that is easily expressed and purified by routine chromatographic techniques[52].

Surface proteins that are attached to the cell wall by SrtA have an N-terminal secretory signal that is presumed to promote translocation across the bacterial cell membrane (Figure 6). They also have a characteristic C-terminal cell wall sorting (CWS) signal which comprises of a conserved LPXTG motif followed by a stretch of about 20 hydrophobic residues and a short positively charged tail[68]. The CWS signal is hypothesized to temporarily retain the translocated protein within the plasma membrane until the substrate is recognized and cleaved by SrtA[69]. For example, *S. aureus* Protein A, which is the model SrtA substrate, contains an LPETG motif followed by a 25 residue long hydrophobic domain and a 5 residue long positively charged tail[70, 71] (Figure 6).

Figure 5. Diagrammatic representation of SrtA. The N-terminal signal peptide *cum* transmembrane sequence and the C-terminal catalytic motif, TLXTC, with the active site cysteine is shown.

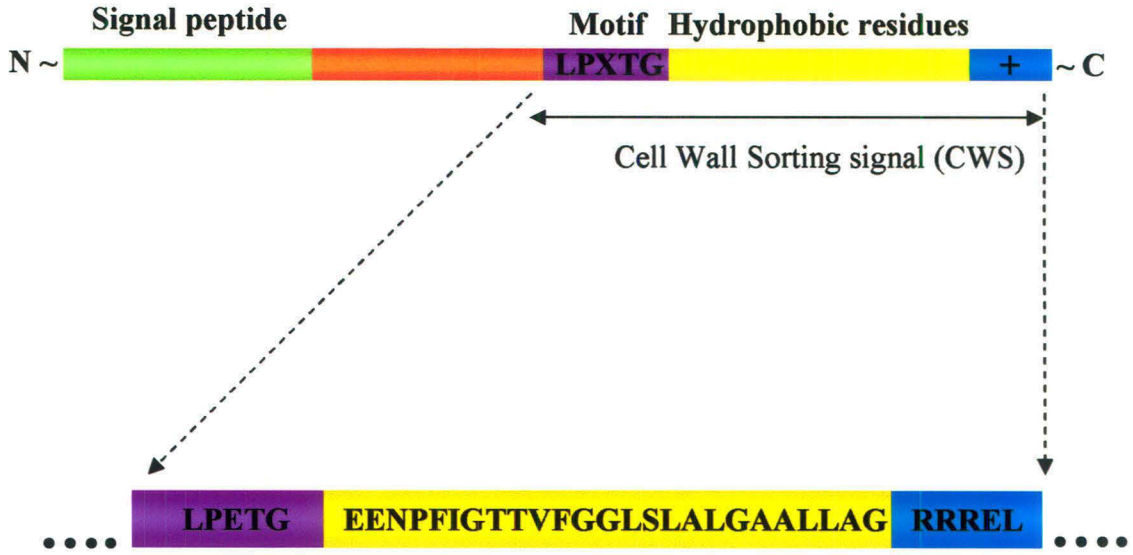


Figure 6. Diagrammatic representation of a typical SrtA anchored protein. The N-terminal signal peptide and the C-terminal CWS signal with the conserved LPXTG motif are depicted. The lower panel shows the amino acid residues in the CWS signal of staphylococcal Protein A.

Substrate specificity of SrtA

Kruger *et al* [72] analyzed the substrate specificity of SrtA using an LPXTG derived peptide library and found that the kinetically preferred residues in positions 1, 2, 4 and 5 of the LPXTG motif were consensus residues Leu, Pro, Thr and Gly respectively. In position 3, where high residue variability has been reported, all amino acid substitutions were tolerated although a 10-fold variation was observed between the best substrate (LPMTG) and the worst substrate (LPTTG). The prototypical substrate, LPETG, which mimics the sortase recognition motif of protein A, was found to be a moderate substrate for the enzyme. SrtA was found to exhibit a marked preference for Met, Tyr, Leu, Gln and Asn at this position. β-branched amino acids such as Ile, Val and Thr were poorly tolerated and neutral amino acids like Gln and Asn were preferred over their charged counterparts like Glu and Asp. In position 1, Leu could be substituted with Met with modest efficiency although other hydrophobes were not tolerated. Similarly Pro in position 2 could be substituted with an Ala residue. The greatest variance was observed at position 4 where in addition to Thr, SrtA processed substrates containing Ala, Ser, Val and Leu. No variants of the Gly residue at position 5 were tolerated. Hence, SrtA exhibits high selectivity for the Leu, Pro, Thr and Gly residues found in the LPXTG consensus motif although it is capable of accepting point alterations, *albeit* at significantly reduced efficiency. Residues following the LPXTG motif also seem to have some impact on the activity of SrtA. The results from a peptide substrate library based on an expanded motif (LPXTGX^1X^2), suggested that Gly residue at X^1 was preferred over other amino acids[73] . However, the enzyme did not seem to have any preference for amino acids at position X^2.

$SrtA_{\Delta 59}$ structure and catalysis

NMR and X-ray crystallography of *S. aureus* $SrtA_{\Delta 59}$ revealed that the enzyme adopts a unique eight-stranded β-barrel with an α-helix and two 3-turn helices connecting the β strands (Figure 7)[52, 74, 75]. A pocket created by β4, β7 and β8 forms the binding site for the peptide substrate and delineates the active site of the enzyme. The catalytic cysteine, Cys^{184}, is located in the β7 strand; two residues, His^{120} and Arg^{197}, are positioned in close proximity to the sulfhydryl group of Cys^{184}

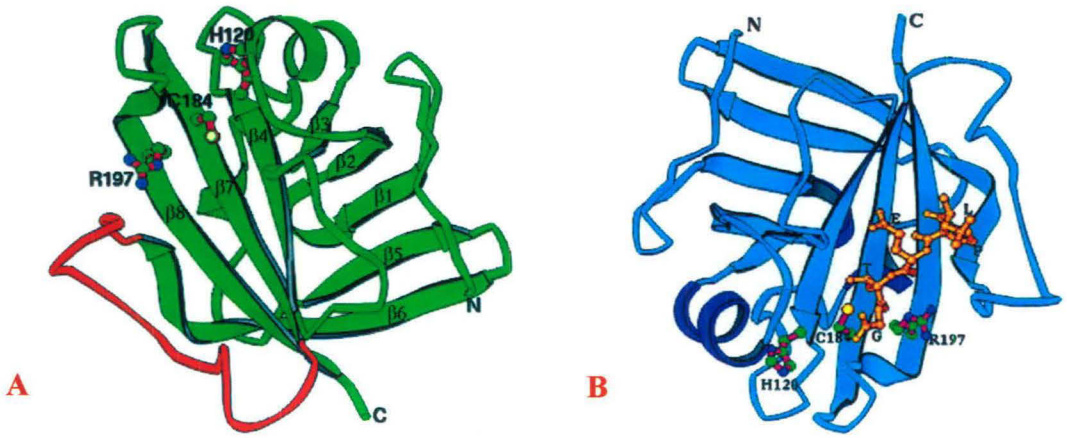


Figure 7. Crystal structure of $SrtA_{\Delta 59}$ and $SrtA_{\Delta 59Cys184Ala}$ – peptide complex[75] **(A)** The core of $SrtA_{\Delta 59}$ is an 8-strand β barrel. $\beta 4$, $\beta 7$ and $\beta 8$ form a concave structure containing Cys^{184}. **(B)** LPETG bound in the $SrtA_{\Delta 59Cys184Ala}$ active site. $SrtA_{\Delta 59Cys184Ala}$ is an active site mutant ($Cys^{184} \rightarrow Ala$) whose crystals are isomorphous with those of $SrtA_{\Delta 59}$.

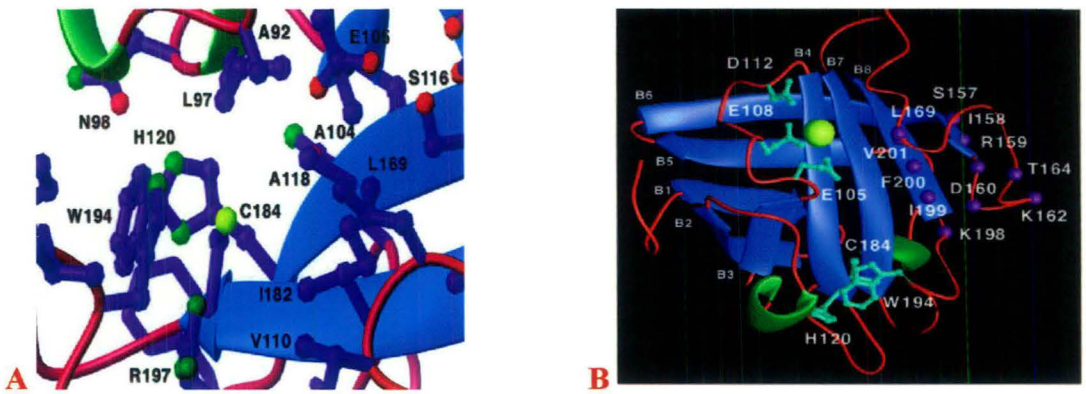


Figure 8. NMR structure showing the active site of $SrtA_{\Delta 59}$ and the Ca^{2+} binding site[52]. **(A)** an enlarged view of the active site showing the amino acid residues in close vicinity to the active site Cys^{184}. **(B)** Ribbon drawing of $SrtA_{\Delta 59}$ showing the putative Ca^{2+} binding site.

and are considered crucial for catalysis though their mechanistic contributions are still under investigation. Using a cyanoalkene analog of LPETG and NMR analysis, three residues located adjacent to the active site (Thr^{180}, Ile^{182} and Ala^{118}) were shown to undergo rearrangement upon substrate binding (Figure 8A). In agreement with the hypothesis that these three residues are crucial for recognition and stabilization of bound substrate, amino acid substitutions of these residues have profound effects on SrtA activity[76].

Calcium ions are known to stimulate $SrtA_{\Delta 59}$ activity. A conserved calcium binding site is formed by the side chains of Glu^{105}, Glu^{108} and Glu^{112} within the $\beta 3$-$\beta 4$ loop and possibly Glu^{171} from the $\beta 6$-$\beta 7$ loop (Figure 8B). The binding of Ca^{2+} ions is important for structural rearrangements of a disordered loop covering the active site that enables substrate binding [52, 75, 77]. It has been recently reported that sortase can form dimers *in-vitro*; its significance vis-à-vis catalysis has not been elucidated yet[78, 79].

Sortase catalyzed surface protein anchoring

Surface proteins that are attached to the cell wall by sortase are synthesized in the bacterial cytoplasm as precursors harboring an N-terminal signal peptide and a C-terminal sorting signal (P1 precursor; Figure 9). After initiation into the Sec secretion pathway, signal peptidase cleaves the N-terminal signal peptide to generate the P2 precursor. The hydrophobic domain of the sorting signal is thought to temporarily retain the polypeptide in the membrane, thereby allowing the LPXTG motif to be recognized by SrtA. The sulfhydryl group of Cys^{184} attacks and cleaves the scissile Thr-Gly bond and forms a thioester bond with the carboxyl group of Thr. The thioester linked acyl-enzyme intermediate, is subsequently resolved by the nucleophilic attack of the amino group of the pentaglycine cross-bridge of the peptidoglycan precursor lipid II forming the P3 intermediate. The P3 intermediates are subsequently incorporated into the envelope via the transglycosylation and transpeptidation reactions of cell wall biosynthesis, ultimately generating mature product (M), i.e., surface proteins that are covalently linked at the C-terminal end to the cell wall envelope[53, 80].

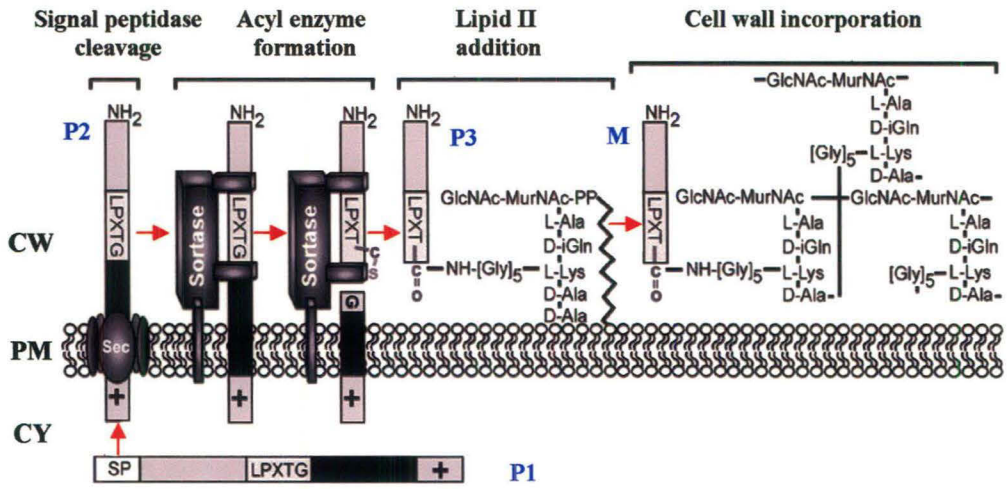


Figure 9. SrtA mediated anchoring of surface proteins[80]. The black bar and + indicate the hydrophobic domain and the positively charged tail of the CWS signal respectively. CY: cytoplasm; PM: plasma membrane; CW: cell wall.

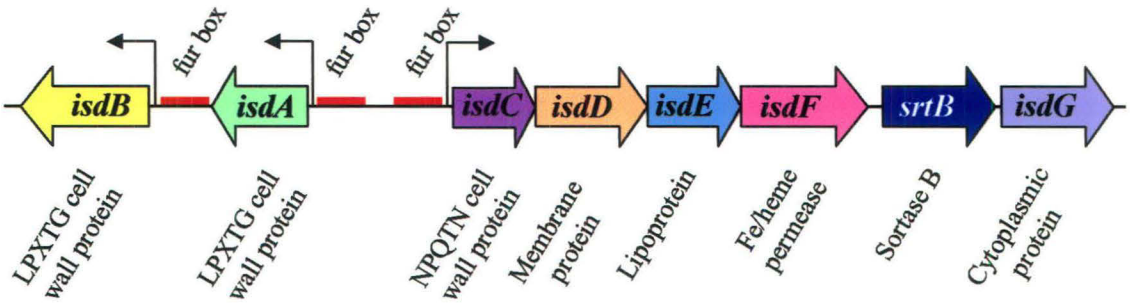


Figure 10. Genomic organization of the *S. aureus isd* locus. SrtB is located in an operon with a putative Fe transporter (isdE, isdF) and a surface protein with a C-terminal NPQTN sorting signal (isdC). Adjacent to this operon are two LPXTG containing proteins that are SrtA substrates. All transcriptional units contain a Fur box upstream of their putative promoters.

Staphylococcus aureus Sortase B (SrtB)

Domain structure and substrate specificity

The staphylococcal genome encodes for another sortase called Sortase B (SrtB). SrtB is a 246 residue protein with a putative N-terminal signal peptide/transmembrane sequence (the first 30 residues at the N-terminus) and a TLXTC motif at the C-terminus. Deletion of the transmembrane segment yields a soluble enzyme that can be expressed and purified by routine procedures. Unlike SrtA, which is transcribed as a single gene, *srtB* is part of an operon called the *isd* (*i*ron-responsive *s*urface *d*eterminant) locus (Figure 10). This locus is composed of eight proteins (IsdA-G, SrtB) which together make up an iron scavenging and transport system[81]. Specifically, the locus encodes cell-wall heme binding proteins (IsdA, IsdB, IsdC), a membrane based heme transport system (IsdD, IsdE, IsdF), a heme degrading monooxygenase (IsdG), and SrtB[82, 83]. The transcription of the locus is regulated by the ferric uptake regulator protein (Fur) which binds to a DNA sequence called the Fur box and represses transcription during iron-replete conditions[84, 85].

IsdC appears to be the only surface protein substrate in *S. aureus* that is anchored by SrtB, which cleaves the substrate at the Thr-Asn bond of the NPQTN motif and immobilizes it at the cell wall cross bridges[86]. Although the SrtB substrate specificity, mechanism of action and anchor structure of SrtB protein substrates is not very well studied, it is thought to function in a manner similar to SrtA, i.e., by attacking the Thr-Asn peptide bond of the NPQTN sorting motif, forming a thioester bond with the carboxyl group of Thr and thereafter transferring it to the pentaglycine crossbridges of lipid II.

Structure of $SrtB_{\Delta 30}$

The crystal structure of $SrtB_{\Delta 30}$ (amino terminal 30 residues deleted version of SrtB) has been elucidated (Figure 11). $SrtB_{\Delta 30}$ is composed of an eight-stranded β-barrel core containing mainly antiparalell β strands (numbered β1-β8), five α-helices (numbered H1-H5), several loops and a hairpin turn[87].The topology of the β-barrel is identical to that observed in $SrtA_{\Delta 59}$. The $SrtB_{\Delta 30}$ crystal structure differs

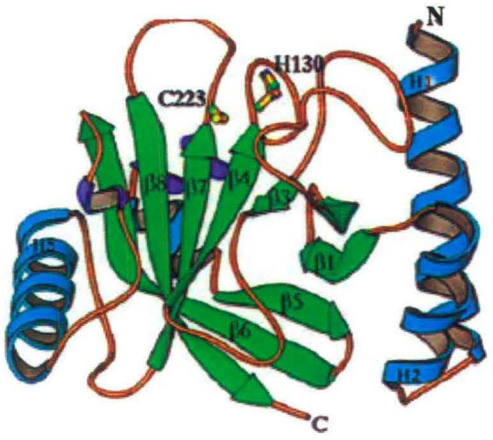


Figure 11. Structure of $SrtB_{\Delta 30}$[87]. Individual β strands and α helices are labeled. The catalytic Cys^{223} is shown as a ball-and-stick model

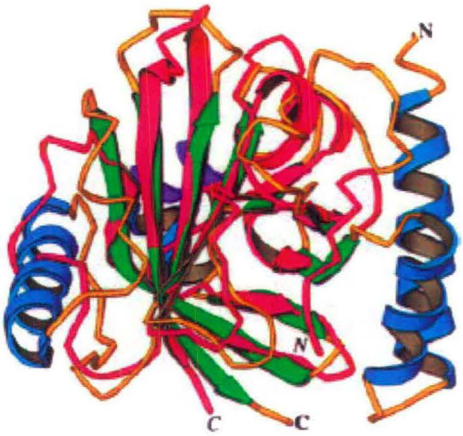


Figure 12. Superimposition of $SrtB_{\Delta 30}$ and $SrtA_{\Delta 59}$ NMR model[87]. The NMR $SrtA_{\Delta 59}$ backbone ribbons are shown in red, and the $SrtB_{\Delta 30}$ is in other colors. The N-terminal two-helix bundle and helices H4 and H5 observed in the $SrtB_{\Delta 30}$ crystal structure are not observed in the $SrtA_{\Delta 59}$ NMR structure.

from the $SrtA_{\Delta 59}$ crystal structure by the insertion of a four-turn α-helix and a one-turn 3_{10} helix between the $\beta 6$ and $\beta 7$ strands, and a two-turn 3_{10} helix and one-turn α helix between the $\beta 4$ and $\beta 5$ strands. Also, the N-terminal of $SrtB_{\Delta 30}$ forms two well-defined α helices not found in $SrtA_{\Delta 59}$ (Figure 12). The catalytically important Cys^{223} is located at the tip of the $\beta 7$ strand[88].

Sortases, surface proteins and microbial pathogenesis

Staphylococcus aureus is a human and animal pathogen that has the ability to penetrate into the deep layers of host tissue causing suppurative lesions in virtually all organ systems[89]. Staphylococci lack pili or fimbrial structures and instead rely on surface protein mediated adhesion of host cells or invasion of tissues to evade the immune system and establish disease[2]. Furthermore, *S. aureus* utilizes surface proteins to sequester iron from the host tissue during infection[86]. A majority of surface proteins involved in these aspects of staphylococcal disease are sortase substrates, i.e., they are covalently linked to the staphylococcal cell wall by the action of either SrtA or SrtB.

Significance of Sortase-anchored surface proteins in *S. aureus*

Using cell sorting signals as queries in bioinformatics searches, 18-21 genes encoding putative sortase anchored surface proteins have been identified in *S. aureus* (Table 1). The number of genes varied with the strain under investigation[90].

Microbial surface components recognizing adhesive matrix molecules (MSCRAMMs) are bacterial elements crucial for host tissue adhesion and invasion[2]. Fibronectin binding proteins, FnbpA and FnbpB, are MSCRAMMs responsible for binding to fibronectin. As fibronectin is found in the extracellular matrices of most tissues as well as in soluble form in body fluids, staphylococci can adhere to virtually all tissues as also to serum coated foreign objects like catheters or artificial heart valves[1, 91].

The collagen binding protein, Cna, which binds to collagen, is important in staphylococcal strains that cause connective tissue disorders or osteomyelitis[92-94].

S. aureus strains clump in the presence of plasma. This phenomenon, which has been exploited for diagnostic purposes, is due to the interaction of two

Table 1. *Staphylococcus aureus* cell wall anchored surface proteins

Surface protein	Ligand	Motif	Sortase
Protein A (Spa)	Ig, von Willebrand Factor, TNRF	LPETG	A
Fibronectin binding protein A (FnbpA)	Fibronectin, firbrinogen, elastin	LPETG	A
Fibronectin binding protein B (FnbpB)	Fibronectin, fibrinogen, elastin	LPETG	A
Clumping factor A (ClfA)	fibrinogen	LPDTG	A
Clumping factor B (ClfB)	Fibrinogen, keratin	LPETG	A
Collagen adhesion (Cna)	Collagen	LPKTG	A
SdrC	Unknown	LPETG	A
SdrD	Unknown	LPETG	A
SrdE	Unknown	LPETG	A
Pls	Unknown	LPDTG	A
SasA	Unknown	LPDTG	A
SasB	Unknown	LPDTG	A
SasC (Mrp)	Unknown	LPNTG	A
SasD	Unknown	LPAAG	A
SasE/IsdA	Heme	LPKTG	A
SasF	Unknown	LPKAG	A
SasG/Aap	Unknown	LPKTG	A
SasH	Unknown	LPKTG	A
SasI/HarA/IsdH	Haptoglobin	LPKTG	A
SasJ/IsdB	Hemoglobin, heme	LPQTG	A
SasK	Unknown	LPKTG	A
IsdC	Heme	NPQTN	B

MSCRAMMs, clumping factor A and B (ClfA and ClfB) with soluble fibrinogen[95, 96].

The serine(S)-aspartate(D) repeat containing proteins (Sdr C, D and E) contain putative high affinity fibrinogen as well as calcium binding domains although no such interactions have been demonstrated yet[97, 98].

S. aureus protein A (Spa) binds to the Fc terminus of immunoglobulins (Ig) resulting in the uniform coating of staphylococci with antibodies which 'masks' the microbe from the immune system. Spa also binds to the von Willebrand factor, a serum polypeptide that promotes physiological homeostasis of blood, as also to TNFR1 (tumor necrosis factor receptor 1), a signaling molecule involved in pro-inflamatory cytokine response and innate immunity[99-101]. Sequestration of these molecules is hypothesized to promote bacterial survival in the host.

Pathogenic bacteria have an absolute requirement for iron during the infectious process. Four Isd proteins, IsdA, IsdB, IsdC and IsdH/HarA, are involved in binding heme or hemoproteins and appear to play a role in scavenging of iron during staphylococcal infection. IsdA, IsdB and IsdC are encoded by the *isd* locus whereas HarA/IsdH is encoded by a gene situated outside the locus. HarA/IsdH has been shown to bind haptoglobin/hemoglobin complexes; IsdB binds to hemoglobin[86, 102, 103]. All the four proteins, including IsdC which is the only staphylococcal substrate of SrtB, are capable of binding to heme. It has been proposed that these proteins are involved in capturing hemoproteins, liberating heme, and promoting heme transport across the bacterial cell wall envelope.

The functions of twelve *S. aureus* surface proteins (SasA, B, D, SasF-H, SasK, SdrC-E and Pls) is not yet known. Sas C was recently characterized as the multidrug resistance protein (Mrp).

Significance of Sortase anchored proteins in pathogenesis

Given the importance of cell surface proteins, loss of the sortase genes leads to defective pathogenesis of gram-positive bacteria. Effect of deletion of SrtA has been assessed and reported in various models. When injected at a sub-lethal dose into the blood stream of mice, *srtA*- mutants display a 2-log reduction in bacterial growth in multiple organ systems as well as a 1.5-log increase in the lethal dose (LD_{50}) compared to wild-type strains. The mutants were also cleared form the blood

stream faster and much more effectively than their wild type counterparts[50]. In the septic arthritis model, staphylococcal inoculations into the blood stream lead to joint infections resulting in cartilage and bone destruction. In experimental endocarditis, lesions on the heart valves serve as a site of deposition of staphylococci, causing destructive damage of endocardial tissue. In both model systems, *srtA-* mutants displayed large reductions in pathogenesis. Jonsson *et al*[104] reported that in the septic arthritis model, mice infected with the knockout mutants showed improved survival as compared to those injected with wild-type staphylococci. In the endocarditis model, Weiss and coworkers[105] reported a two-fold higher mortality with the wild type *S. aureus* Newman strain as compared to the *srtA-* isolates. When non-bacterial thrombotic endocarditis was induced by the insertion of a polyethylene catheter into the heart of experimental animals, a significant difference between the mean cardiac bacterial counts for the wild-type strain and the *srtA-* strain was observed. They proposed that the *srtA* gene deletion and the ensuing loss of the ability to anchor cell surface proteins, such as those responsible for fibrin binding, affected the ability of the *srtA-* strain to infect and colonize in the cardiac environment. Similar results were also reported by Bubeck *et al*[106] using an experimental model of *S. aureus* lung infection which monitors the development of acute pneumonia and disease-associated mortality following intranasal inoculation of mice with staphylococci. Animals infected with *srtA-* mutants showed a significant reduction in mortality as compared to their wild-type counterparts.

The importance of sortase-attached surface proteins was further reinforced by studies wherein the ability of surface proteins to elicit protective immunity was assessed. Following vaccination with purified components, protective immunity was assessed for Group B streptococcal, pneumococcal, and staphylococcal infections and it was observed that by combining sortase substrates as vaccine antigens, animals could be protected against a lethal challenge with the pathogen[107]. For example, the combination of IsdA, IsdB, SdrD, and SdrE as vaccine antigens protected mice against lethal challenge with *S. aureus* strains, including several MRSA (Methicillin resistant *Staphylococcus aureus*) strains. Gaudreau *et al* [108] reinforced the significance of these results by their studies on the utility of a multivalent polyprotein DNA vaccine for immunization against *S. aureus*. They used a series of plasmids encoding cell wall attached proteins like ClfA, FnbpA and

SrtA itself, either singly or combined as a polyprotein, to immunize mice. Upon challenge with a virulent *S. aureus* isolate, a significantly larger number of vaccinated mice survived as opposed to the control group. Vaccinated mice were also observed to show no signs of arthritis when challenged the *S. aureus* Newman strain which caused reactive arthritis in the controls.

Staphylococcal *srtB*- mutants do not show as drastic a decrease in virulence as is observed for mutants lacking *srtA*. In the renal abscess model, *srtB*- mutant staphylococci exhibited defects in persistence of infection and an overall decrease in virulence in the infectious arthritis model. This corroborates the hypothesis that as heme uptake is crucial during the infectious process, deletion of *srtB* has a more profound effect on persistence of infection rather than establishment[81, 105, 109].

Virulence attributes of sortase and surface protein extend well beyond staphylococci, as defects in the pathogenesis of sortase mutants have been reported for animal infections with many different species, including Actinomyces, Enterococci, Streptococci, bacilli, and listeria[110-118]. Collectively, these studies illustrate the many important and diverse roles that surface proteins play during the pathogenesis of infections caused by Gram-positive pathogens.

Importance of sortases in the assembly of pili

Pili or fimbriae represent some of the most important virulence determinants for bacterial infection of a mammalian host[119, 120]. These macro-molecular structures consist of repeating protein subunits structured as the pilus shaft which extends from the bacterial surface into the surrounding medium. The tip or cap of these fimbrial shafts typically display adhesive properties that promote bacterial binding to extracellular matrices or host cell receptors in the context of an infection[121-123].

Unlike gram-negative bacteria whose pili subunits are noncovalently attached to each other and whose assembly is a non enzymatic process, gram positive bacterial pili are covalently connected and their assembly requires the presence of sortases[62, 64, 124, 125]. The mechanism of assembly is best characterized in *Corynebacterium diptheriae*.

The prototype: *Corynebacterium diptheriae* pilus assembly

The genome of *C. diptheriae* contains three different pili gene clusters. Each cluster forms a distinct pilus, designated as the SpaA (*s*ortase-mediated *p*ilin *a*ssembly A), SpaD and SpaH type according to the major subunit that makes up each pilus (Figure 13)[61]. The assemblies of these pili require cognate sortases that are encoded within each pilus gene cluster[61, 126-128].

The sortases involved in *C. diptheriae* pilus assembly are named SrtA, B, C, D and E. These sortases belong to the class C category as they recognize the LPXTG motif but are found clustered in the same gene loci as their substrates. The housekeeping sortase of *C. diptheriae,* SrtF, belongs to class A[129].

The biogenesis of the SpaA pili requires four adjacent genes: *spaA*, *spaB*, *spaC* and *srtA*. SpaA is the major component of the pilus shaft, SpaB is spaced at regular intervals along the shaft and SpaC is located at the pilus tip. The assembly is catalyzed by a single sortase, SrtA, located within the *spaA* locus.

The SpaD and SpaH pilus assembly requires the action of two sortase enzymes: SrtB and SrtC in case of SpaD; and SrtD and SrtE in case of SpaH.

Studies with deletion mutants have indicated that the housekeeping sortase, SrtF, is required for efficient cell-wall attachment of all three types of polymerized pilin subunits[61, 130].

Sequence analysis of the major pilin subunits SpaA, SpaD and SpaH uncovered three key motifs involved in pilus assembly: the pilin motif, WxxxVxVYPKN with an invariant lysine, the E box with an invariant glutamate and the CWS signal. Bioinformatic analysis has revealed that these three motifs are conserved in every gram-positive major pilin protein examined till date. As the ancillary pilin subunits lack both the E box and the pillin motif, it is presumed that the major subunits provide all the necessary elements for pilus assembly. This is corroborated by experimental studies of *spaA*- deletion mutants which show that SpaA is essential for formation of the pilus structure, whereas the two ancillary proteins, SpaB and SpaC, are dispensable for pilus assembly[61, 129, 130].

The first step of pilus assembly involves the Sec-dependent secretion of the pilus components all of which contain N-terminal signal peptides and C-terminal LPXTG motif containing CWS signals (Figure 14)[129]. Each component remains anchored to the cell membrane, owing to the presence of a membrane-spanning

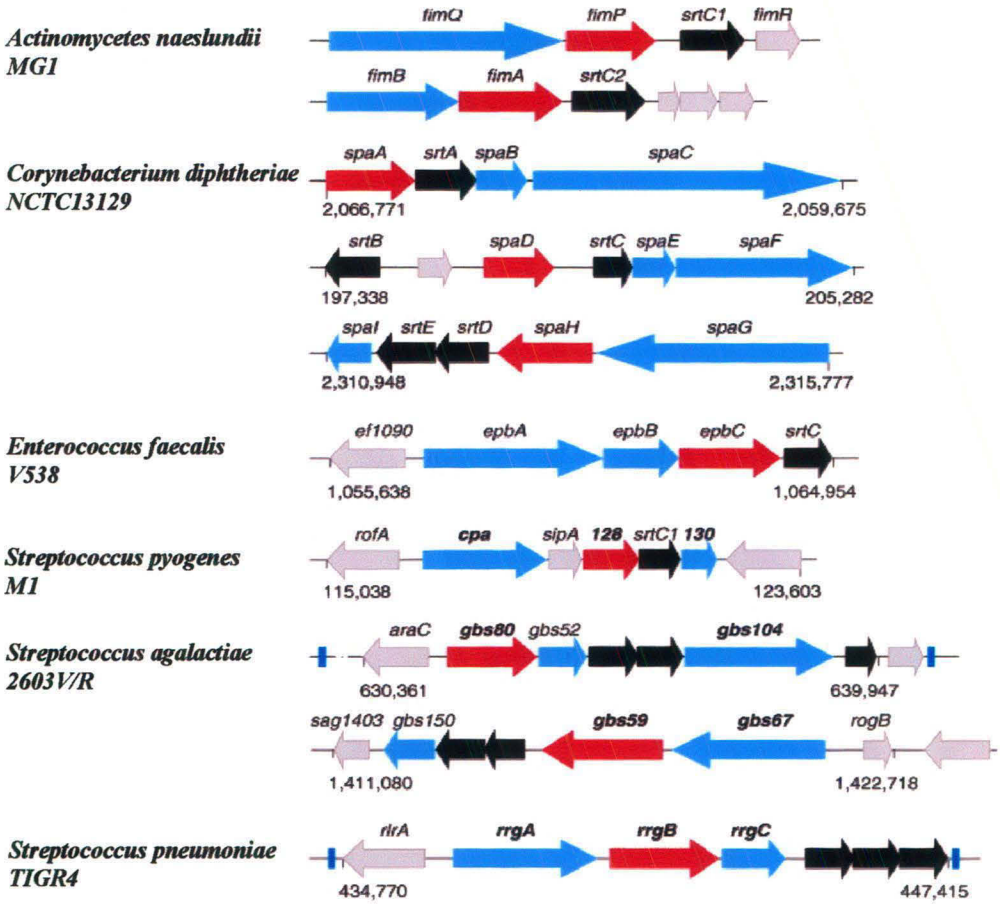


Figure 13. Pilus gene operons[129]. Graphic presentation of pilus gene clusters identified in the chromosome of *Actinomyces naeslundii* MG-1, *Corynebacterium diphtheriae* NCTC13129, *Enterococcus faecalis* V538, *Streptococcus pyogenes* M1, *Streptococcus agalactiae* 2603V/R and *Streptococcus pneumoniae* TIGR4. Each cluster contains pilus-specific sortase gene(s) (black), genes encoding a major subunit (red) and minor pilins (aqua). Some of the clusters are flanked by transposable elements (blue). Uncharacterized genes are colored in gray.

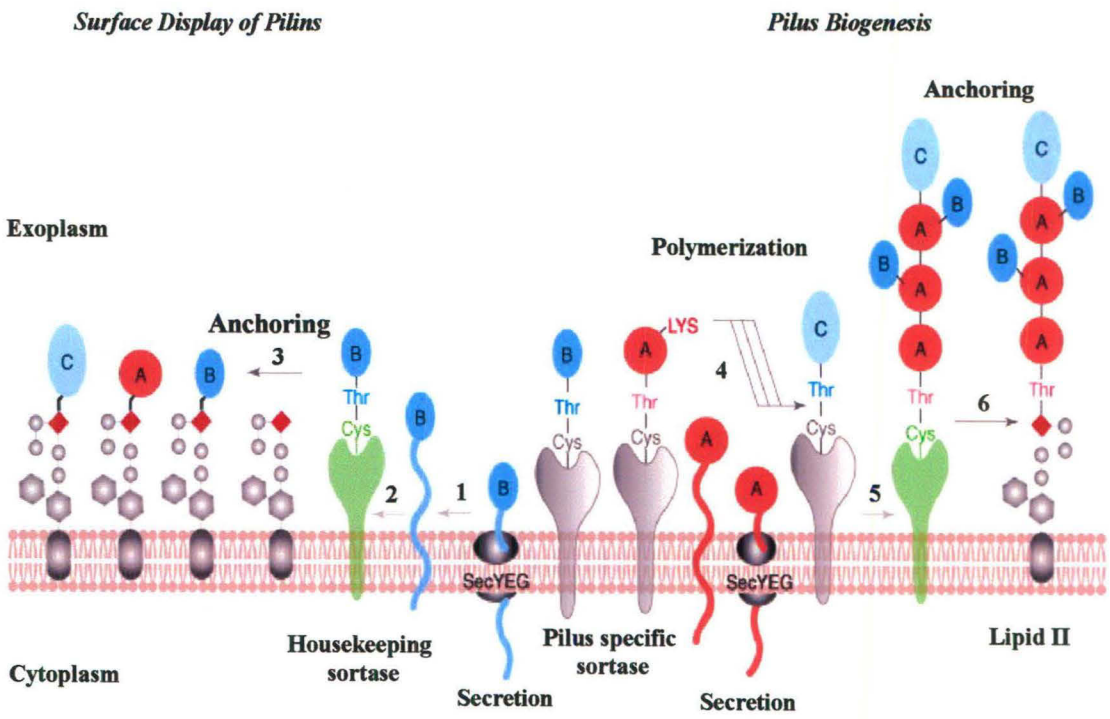


Figure 14. Model of pilus biogenesis[129]. Pilin precursors SpaA, denoted by pink circles; SpaB, denoted by dark-aqua ovals; and SpaC, denoted by light-aqua ovals. Red diamonds denote the D-diaminopimelic moiety of the cell wall pentapeptide. SecYEG stands for the three subunits of the general secretion machinery (Sec).

domain at the C-terminus. The second step is a branching point where the pilus precursor can either form an acyl-enzyme intermediate with the housekeeping SrtF or with the pilus specific sortase. In case of the former, the enzyme intermediates are transferred to lipid II thus anchoring monomeric pilin units to the cell wall (Step 3). In case of the later, pilus polymerization takes place by nucleophillic attack of the ε-amine of the lysine (K) present in the pilin motif (step 4).

Pilus polymerization is terminated when pilus polymers are transferred to lipid II in one of two possible ways. In one pathway, the housekeeping sortase having a SpaA monomer would receive the pilus polymer form the pilus-specific sortase (step 5) and transfer the polymer to lipid II (step 6). In the alternative pathway, the pilus-specific sortase would transfer the polymer directly to lipid II.

Several aspects of the pilus-assembly process remain to be elucidated. It is still not clear how the different pilus subunits are organized and what determines their order and frequency in the pilus structure. Gene-inactivation experiments in *C. diptheriae* have shown that, although the main pilus subunit can polymerize in the absence of the two ancillary proteins, the reverse is not true for the ancillary proteins[61]. Sortase specificity has been proposed to play a key role in incorporation of the ancillary proteins into the pilus structure as it has been shown that in group B Streptococcus, the sortases SAG0647 and SAG0648, encoded by pilin island 1, are each specifically required for incorporation of one of the ancillary proteins[131].

Although studies on pilus formation in other gram-positive organisms are still in their nascent stages, it is believed that sortase catalyzed pathway for pilus assembly may be common to all gram-positive bacteria. Studies of *Streptococcus agalactiae*, *S. pyogenes*, *S. pneumoniae*, *Enterococcus faecalis* and *A. naeslundii* show typical clustering of pilus genes together with multiple pilus-specific sortases (Figure 13). Similar to corynebacterial pili, the streptococcal and enterococcal pili are heterotrimeric, with the pilus shaft containing two minor pilins[62, 63, 129].

The role of pili in host cell adherence and tissue tropism has been well established from studies of Gram-negative bacteria; gram-positive pili are also hypothesized to play a major role in host–pathogen interaction and the eventual colonization of host tissues by pathogens. As pili are considered crucial for the adhesive interactions of bacteria with host cells, in biofilm formation and in modulation of host immune response leading to the establishment of commensal or

pathogenic relationships, pilus assembly provides a unique opportunity for inhibition of virulence[63, 129].

Inhibitors of sortase catalyzed transpeptidation

Since sortase catalyzed transpeptidation is crucial for so many aspects of microbial pathogenesis, it constitutes an extremely attractive target for the development of new antibacterial drugs. Another factor that increases the appeal of therapy against sortase is that as surface proteins are crucial for pathogenicity but not survival, sortase inhibitors should act as anti-infective agents and disrupt virulence without affecting microbial viability. This reduces the selection pressure towards the development of drug resistance. A number of sortase inhibitors have already been reported (Table 2).

Taking the approach that plants may harbor compounds that inhibit the activity of SrtA, Kim *et al*[132]examined the extracts of over 80 medicinal plants for inhibition of recombinant SrtA of which the extracts of *Cocculus trilobus*, *Fritillaria verticillata*, *Liriope platyphylla* and *Rhus verniciflua* exhibited the most potent anti-SrtA activity with IC_{50} values ranging form 1.6-8.5 μg/ml. The extract form *Fritillaria* was investigated further and the active principle was discovered to be a glucosylsterol, *β*-sitosterol-3-*O*-glucopyranoside[133]. The anti-sortase activity was attributed to the glucopyranoside moiety as sitosterol alone was not seen to exhibit any inhibitory activity.

Isoquinoline alkaloids (berberine chloride) from the rhizome of *Coptis chinensis* and curcuminoids (curcumin, demethoxycurcumin and bisdemethoxycurcumin) isolated form the rhizome of *Curcuma longa* exhibited strong activity against SrtA with IC_{50} values of 8.7μg/ml and 13-35μg/ml respectively. These IC_{50} values were much lower than that for pHMB (*p*-hydroxymercuribenzoic acid) which is a generic cysteine protease inhibitor known to inhibit the activity of SrtA. Curcumin treated *S. aureus* cells showed decreased ability to bind to fibronectin[135 134].

Several compound isolated form marine sponges have been demonstrated to be effective inhibitors of SrtA. Bromodeoxytopsentin, a Bis(indole)alkaloid of the Topsentin class , isolated form the marine *Spongosporites sp*, reduced the ability of *S. aureus* cells to bind to fibronectin coated plates, in a dose dependent manner. It

Table 2: Some examples of SrtA and SrtB inhibitors

Name	Source/Type	IC_{50} *in-vitro*	*In-vivo* inhibition	Mechanism
β-sitosterol-3-O-glucopyranoside	*F.verticillata*	SrtA: 18μg/ml	Inhibition of bacterial growth	N.D.
Berberine chloride	*C. chinensis*	SrtA: 8.7μg/ml	Inhibition of bacterial growth	N.D.
Isoaptamine	*Aaptos aaptos*	SrtA: 3.7μg/ml	Binding to fibronectin	N.D.
Curcumin	*C. longa*	SrtA: 13μg/ml	"	N.D.
Flavonoid phenols	*R. verniciflua*	SrtA: 37-123μg/ml SrtB: 8-38μg/ml	Clumping	N.D.
Diazo/chloromethyl ketone	Synthetic, substrate mimic	N.D.	N.D.	Most likely covalent
333-Trifluro-1-(phenylsulfonyl)-1-propene	Synthetic	SrtA: 190μM	Binding to fibronectin	Covalent modification of Cys^{184}
Phosphinic peptidomimetic	Synthetic, transition state mimic	SrtA: 10mM	N.D.	N.D.
Diacrylonitrile	Small molecule library	SrtA: 2.31μg/ml	Binding to fibronectin	Possibly noncovalent
Aryl β-amino(ethyl) ketones	Small molecule library	SrtA: 4.8μM SrtB: 14μgM SrtC: 15μM	N.D.	Covalent, mechanism based

also inhibited the activity of SrtA with an estimated IC_{50} of ~20μg/ml[136]. Isoaaptamine isolated form the tropical sponge *Aaptos aaptos* was a potent inhibitor of SrtA with an IC_{50} of ~4μg/ml. It also suppressed the ability of *S. aureus* to bind to fibronectin coated surfaces[137]. Dihydroxystyrenes isolated form the *Jaspis sp.* of sponge showed mild inhibitory activity against SrtA[138].

Three flavanoids, morin, kaempferol and galangin, along with *β*-sitosterol-3-*O*-glucopyranoside, displayed at least a 3-fold difference in inhibition between SrtA and SrtB (SrtA: 37.39, 77.94, 123.22μg/ml; SrtB: 8.54, 24.55, 38.37μg/ml respectively) suggesting that it may be possible to identify inhibitors that discriminate between specific members of the sortase family[139]. Two cyclic peptides isolated from an actinomycete obtained form an unidentified marine sponge exhibited weak inhibition against SrtB with EC_{50} values in the range of 88-126μg/ml[140].

Several investigators have taken advantage of the active site thiol features of sortase or the conserved structure of its peptide substrate for designing inhibitors. The first such report investigated the use of synthetic substrate-derived inhibitor sequences with reactive electrophiles aimed at irreversibly modifying the active site thiol of sortase[141]. Three substrate-derived, active-site directed, irreversible inhibitors of SrtA were designed based on the substrate recognition motif of staphylococcal SrtA (LPXTG). The authors replaced the scissile amide bond between the Thr and Gly with diazoketone (-$COCHN_2$) or chloromethyl ketone (-$COCH_2Cl$) grouping. Both diazomethane and chloromethane were chosen on the basis of their recognized ability to alkylate the active-site thiol groups of cysteine proteases. Both compounds displayed micromolar inhibitor constants with the chloromethyl ketone demonstrating faster kinetics of inactivation. Along these lines, vinyl sulfones, electrophilic inhibitors of cysteine proteases, were examined for inhibition of sortase. Collectively, these compounds displayed time-dependent irreversible inhibition of sortase[142]. Among the vinyl sulfones, 3,3,3-trifluoro-1-(phenylsulfonyl)-1-propene, was most effective with an IC_{50} of 190μM. In addition to substrate-derived inactivators, inhibitors aimed at imitating the transition state of the sorting reaction were evaluated for activity. A nonhydrolyzable phosphinic peptidomimetic (NH_2-YALPE-Alaψ[PO_2H-CH_2]Gly-EE-NH_2) showed competitive inhibition with an IC_{50} of 10mM[143]. Analogues of threonine (T*), where the

carboxyl group of Thr was replaced with a mercaptomethyl (CH_2SH) unit, were used to make tetrapeptide substrate mimics for SrtA (LPAT*) and SrtB (NPQT*)[144]. Both these peptides were observed to modify their respective enzymes by reacting to the active site Cys residue.

High throughput screening has also been used to search for novel inhibitors of sortase activity. In-silico virtual screening of SrtA against commercial compound libraries led to the identification of 108 potential inhibitors[145]. Eight of the 108 compounds exhibited *in-vitro* inhibition of catalytic activity of SrtA with IC_{50} values ranging from 75-400μM. The most active compound identified from the virtual screening was a compound designated Compound 1 which had an IC_{50} value of 75μM. Similarly, a screen of 1000 diverse compounds for inhibition of SrtA yielded a diacrylonitrile with an IC_{50} of 231μM[146]. In another study, screening of 135,625 small molecules for inhibition yielded aryl(β-amino)ethyl ketones (AAEK) that inhibit sortase enzymes from staphylococci and bacilli[147]. Aryl(β-amino)ethyl ketones are mechanism based inhibitors; sortases specifically activate this class of molecules via β-elimination, generating a reactive olefin intermediate that covalently modifies the cysteine thiol.

Sortase as a synthetic tool in chemistry and biology

SrtA mediated transpeptidation involving LPXTG motif containing surface proteins and pentaglycine CB of peptidoglycan can be easily reproduced *in-vitro* using LPXTG motif containing short peptides and mono-, di- or triglycine as the nucleophile. Lack of requirement for high-energy intermediates or complex cofactors further simplifies the *in-vitro* reaction.

The idea that the exquisite site-specific cleavage and transpeptidation properties of SrtA could be exploited for protein engineering was introduced by Mao *et al* who used SrtA to ligate native as well as non-native peptide fragments to LPXTG containing peptides[148]. They extended the concept of peptide-peptide ligation and investigated if SrtA could be used for peptide-protein and protein-protein ligations. They reported that GFP (green florescent protein) with an LPETG tag (GFP-LPETG), when incubated with SrtA, was hydrolyzed to GFP-LPET. In the presence of an appropriate Gly nucleophile, it yielded the formation of the expected

protein-peptide conjugate with a yield of about 50% in 6 hours. Furthermore, SrtA successfully mediated the conjugation of GFP-LPETG-6His to a 29 KDa protein expressed with an N-terminal glycine. They also demonstrated the utility of SrtA in ligating non-native peptide fragments, like (D)-Tat peptide (GYGRKKRRQRRR), to the C-terminus of an LPXTG motif. Non-peptidic molecule, like folate, when derivatized with N-terminal triglycine could also yield the expected GFP-LPET-Folate conjugate. To further illustrate the functional utility of SrtA, they applied the process to generate a conjugate between a synthetically made branched PTD-5 (Protein Transduction Domain-5) peptide and recombinant GFP-LPETG-6His, a protein-peptide conjugate that would have been difficult to make by either recombinant expression or chemical synthesis.

SrtA mediated transpeptidation has also been used to develop a new protein purification system for the generation of free recombinant protein in a single affinity -chromatographic step (Figure 15)[149]. The recombinant protein was expressed in *E. coli* as a fusion consisting of an N-terminal 6-His tag, SrtA and an LPETG linker with the protein of interest at the C-terminus. Following routine affinity chromatography, the fusion protein was made to undergo an on-column SrtA mediated cleavage in the presence of Ca^{2+} and polyglycine nucleophile. The recombinant protein of interest was released into the eluate while the 6-His tag along with the rest of the fusion protein remained bound to the column. When compared to other purification systems, this method was found to yield recombinant protein of comparable quantity and quality while eliminating cumbersome steps like enzyme mediated affinity-tag removals which increase wastage while prolonging the time required and the cost incurred in purification.

Nguyen and Schumann have used SrtA as a means to immobilize proteins on the cell surface of *B. subtilis*[150]. Microbial cell-surface display has a wide range of biotechnological and industrial applications including live vaccine development, screening displayed peptide libraries, bio-adsorbants for the removal of harmful chemicals and heavy metals and biosensor development. Although *B. subtilis* encodes for two potential sortases, their substrate specificities remain unknown. This problem was circumvented by introducing the *srtA* sequence of *Listeria monocytogenes*, whose substrate specificity is known to be LPXTG, into *B. subtilis*. By fusing the coding region of the α-amylase gene to the C-terminal region of the *S.*

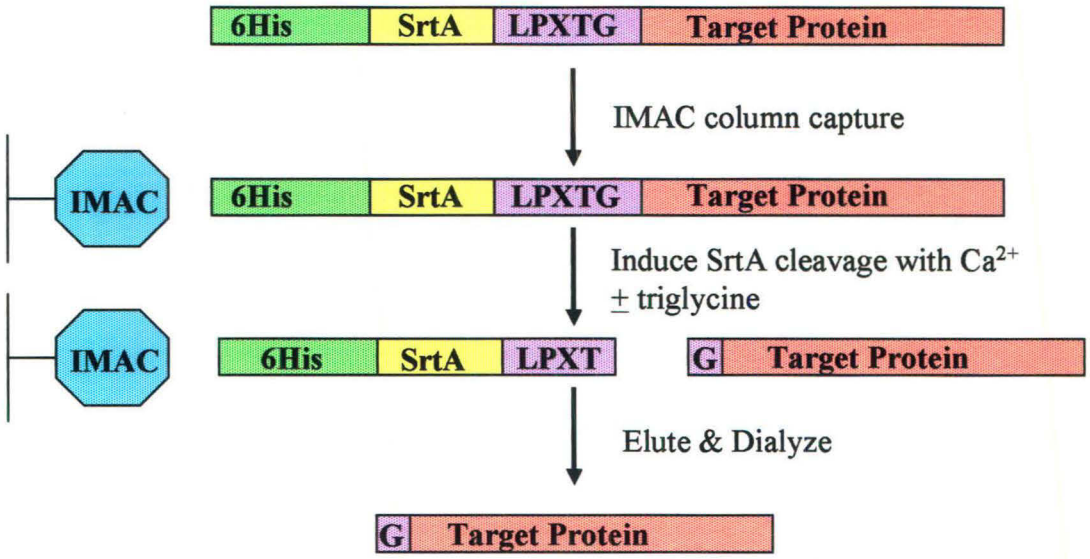


Figure 15. Schematic representation of a self-cleavable sortase fusion protein. IMAC: immobilized metal-ion affinity chromatography

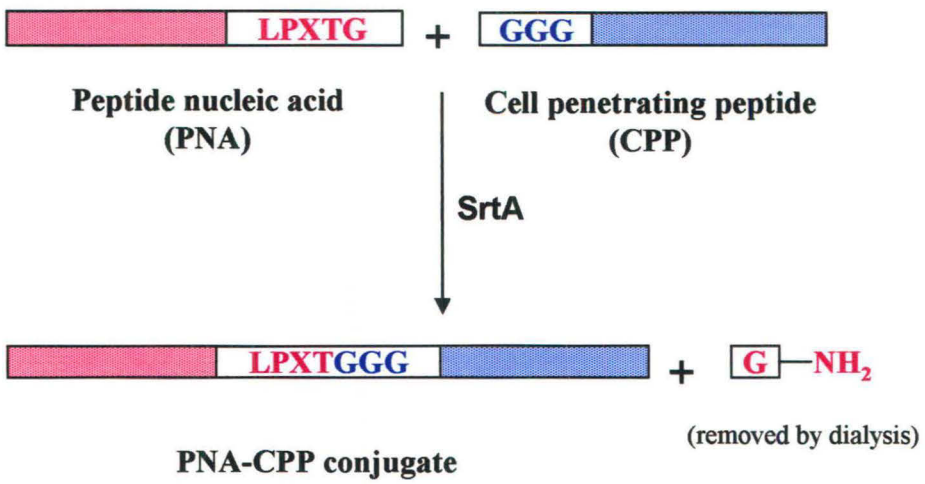


Figure 16. Synthesis of PNA-CPP conjugates. The reaction by-product, G-NH_2, was removed by dialysis to increase yield of the reaction.

aureus FnbpB, which is a natural SrtA substrate, they could achieve as many as 240,000 molecules of α-amylase immobilized per cell which is 24 times more than that achieved using other immobilization strategies.

Immobilization of proteins onto solid supports is an active area of research driven by demands for the development of supported catalysts, affinity matrices, and micro devices as well as for the development of planar and bead based protein arrays for multiplexed assays of protein concentration, interactions and activity. A critical requirement for these applications is the generation of a stable linkage between the solid support and the immobilized, yet functional, protein. Although a large number of chemical immobilization methods are available, they suffer form a lack of specificity and the use of harsh reaction conditions which often leads to impaired protein activity. Several methods that use enzymes for the formation of the linkage have been used to bypass these problems. The disadvantages of this approach are the need to fuse relatively large proteins to the protein of interest which often results in reduced recombinant expression and loss of activity. Other methods like intein-based splicing, biotinylation of target protein also suffer form similar problems.

Parthasarathy *et al* were the first to use SrtA for immobilization of proteins when they attached an LPXTG-tagged GFP to polystyrene beads that had been chemically derived with triglycine[151]. They explored the similarities between diglycine and alkylamines, and hypothesized that alkylamines could also act as nuclephiles in the transpeptidation reaction. They demonstrated this by attaching GFP-LPXTG to polystyrene beads that terminated in an amine. Considering the ready availability of amine terminated surfaces this opens up the field for SrtA catalyzed protein attachment and has the added advantage of not requiring any chemical modification or derivatization of solid supports. The authors have shown that SrtA can be used to create intra- and intermolecular attachments among proteins. It can be used to create dimers, trimers and other higher oligomers of LPXTG containing proteins or peptides and can also be used for cyclization of bifunctiontional (containing both LPXTG and nucleophile i.e., G3-GFP-LPXTG) molecules. The authors have also used SrtA to mono-PEGylate LPETG tagged protein molecules by using amine-terminated PEG (poly ethylene glycol). PEGylation is known to increase the life of proteins by preventing degradation by enzymes. It has been shown to increase the half life of proteins in blood circulation

as well as to function as homing signals on drug loaded liposomes. Over PEGylation has been known to interfere with protein function, mask important epitopes and prevent intercalation into the liposome so the ability of SrtA to mono PEGylate proteins is highly desirable.

In a similar study, Chan *et al* have also shown that SrtA provides a robust yet gentle method for attachment of LPXTG tagged proteins to a variety of solid supports like cross-linked polymer beads, beaded agarose and planar glass surfaces[152].

Clow *et al* have shown similar results using biacore sensor chips[153]. The biacore system is a real-time bio-sensor based on surface plasmon resonance allowing accurate analysis of protein binding specificity, kinetics and affinity. It requires proteins to be firmly attached to the solid support yet have access and be able to bind the ligand protein efficiently. Non specific chemical approaches, site specific attachment methods, like 6-His tags, or attachment of antibodies to Protein A suffer form problems of low specificity and random orientation of attached proteins. SrtA mediated immobilization via the C-terminus allows the protein to be oriented uniformly on the chip such that the N-terminus is exposed. The extremely high specificity of coupling allows immobilization of proteins form less than pure samples allowing for shortcuts in the purification procedure. However, this method cannot be applied for proteins whose binding sites are located within the C-terminus itself because the attachment is C-terminus mediated.

Peptide nucleic acid (PNA) is an artificially synthesized polymer similar to DNA or RNA that has found wide applications in various molecular biology techniques especially antisense therapy. Cationic polypeptides (CPPs) have been known to improve the delivery of PNAs into mammalian cells and enhance their biological activity. Conjugation of PNAs with CPPs by conventional peptide synthesis strategy is prone to several drawbacks like product aggregation and heterogeneity. SrtA mediated transpeptidation was used to ligate PNAs appended with an LPKTG motif to CPPs synthesized with N-terminal triglycine (PNA-LPKT-GGG-CPP) (Figure 16). The yield of the reaction was about 38% which increased to 61% when the reaction by-product (G-NH_2) was removed by dialysis; this pushed the reaction towards synthesis of the PNA-CPP conjugate. Synthesis of PNA-CPP conjugates using SrtA was far superior to the conventional chemical methods as the

S7
S42
616.079

TH-16309

conjugate obtained was homogenous and easily separated form reaction by-products by routine separation strategies[73].

Tagging or labeling is a very useful technique for studying protein behavior in living cells. However, not all proteins are amenable to the installation of large genetically encoded tags like GFP without compromising function or intracellular distribution. Chemical methods for the installation of labels have the advantage of ease of use but lack the precision accorded by genetic tags. Several chemoenzymatic methods and small-molecule-binding peptide sequences that allow site-specific incorporation of labels have been developed but they are hampered by the large-sized fusions required or the demands of synthetic skills involved in producing the reporter molecules that can be installed. SrtA has been used for selective labeling of proteins in cell lysates and on the surface of living cells with a diverse set of probes. Cells expressing recombinant surface proteins appended with C-terminal LPXTG motifs could be tagged, with very high specificity, by glycine derivatized probes on the surface of living cells by adding SrtA to the growth medium[154, 155].

The present thesis

Site-specific incorporation of novel functionalities like peptide analogues, non natural amino acids, fluorophores and other biochemical and biophysical probes into proteins using staphylococcal SrtA, has the potential to develop into a powerful tool for synthetic biology. The multiplicity of sortases and their distinct recognition sequences, in principle, immediately expands the possibilities of the use of sortases in various chemoenzymatic synthetic protocols. Additional expansion or diversity in synthetic utility of sortases can be achieved by innovating newer amine nucleophiles that can effectively substitute Gly_3 or Gly_5 moiety in the transpeptidation reaction. The major aim of the present thesis is to explore sortase catalyzed ligation of LPXTG peptide substrates to aminosugars with a view to develop an enzymatic approach to semisynthesis of neoglycocopeptides and proteins which are otherwise difficult to synthesize by purely chemical means. The thesis is divided into three chapters. Chapter I describes the preparation of enzymatically active sortase. The sortase catalyzed ligation of model LPXTG peptide substrates to aminosugars, aminosugar derivatives such as aminoglycoside antibiotics, and aliphatic amines is described in chapter II. Studies on sortase mediated construction of bioconjugates

comprising of aminoglycoside antibiotics and sequences derived from Rev and Tat proteins of HIV, and their respective RNA binding propensity , are discussed in chapter III.

Chapter I

Construction of Staphylococcus aureus Transpeptidase Sortase

Introduction

Staphylococcus aureus Sortase A (SrtA) is a 206 amino acid thiol transpeptidase that catalyzes the attachment of a large number of surface proteins to the cell wall peptidoglycan of gram-positive bacteria. It is a cell-membrane anchored enzyme that adopts a type II membrane topology i.e., the N-terminus is situated inside the cytoplasm while the C-terminus, which contains the catalytic site, is located across the plasma membrane[53].

The first 25 amino acids of the enzyme constitute an N-terminus membrane anchor sequence. Due to the hydrophobic nature of this segment, the full length protein is not conducive to recombinant expression and purification as it is insoluble and sediments along with the membranous fraction of the cell lysate. A mutant that lacks the anchor sequence ($SrtA_{\Delta N}$) is both soluble and amenable to purification by routine affinity chromatography techniques. Deletion of the N-terminal anchor sequence does not affect the enzymatic activity of sortase as $SrtA_{\Delta N}$ has been demonstrated to catalyze hydrolysis as well as transpeptidation of LPXTG bearing peptides[156].

A bioinformatic analysis of sortase homologues from *Staphylococcus aureus*, *Streptococcus pneumoniae* and *Enterococcus faecalis* showed that the signal peptide, the membrane anchor sequence, and a short linker domain (*S. aureus* sortase residues 26-59) display no amino acid conservation. By contrast, the core sequence (*S. aureus* sortase residues 60-204) has several highly conserved residues which suggested that this domain may comprise the catalytically active domain[52]. Indeed, recombinant sortase comprising residues 60-204, $SrtA_{\Delta 59}$ (hereby referred to as sortase) has been demonstrated to be enzymatically active indicating that removal of nonconserved residues had no effect on catalytic activity. Both sortase constructs, $SrtA_{\Delta N}$ and sortase, catalyze hydrolysis of peptide substrate in the absence of nucleophile and transpeptidation when a nucleophile like NH_2-Gly_3 is added in lieu of peptidoglycan. Further, both $SrtA_{\Delta N}$ and sortase exhibit an enhancement in enzymatic activity in the presence of calcium ions[52].

This chapter describes the construction of *Staphylococcus aureus* sortase, the synthesis of suitable peptide substrates and the establishment of an HPLC based

assay system for effective monitoring of the activity of sortase as a prelude to further investigation of enzyme activity.

Materials and Methods

Materials

Staphylococcus aureus Mu50 was procured from ATCC (ATCC No. 700699). Primers were obtained from Microsynth, Switzerland. Genomic DNA isolation kit and pGEM-T Easy vector were supplied by Promega Life Sciences, USA. All enzymes, PCR reagents, DNA and protein markers were procured from New England Biolabs, USA. Miniprep kit and gel elution kits were purchased from Qiagen, USA. HBI broth and agar and LB broth and agar were obtained from BD, USA. pET23b vector was supplied by Novagen. All Fmoc amino acids and Fmoc amino acid substituted Wang resins were purchased from Novabiochem. DCM and DMF were obtained from Merck. Abz-LPETG(Dap)Dnp was custom made by AnaSpec, USA. All antibiotics and chemicals were obtained from Sigma unless otherwise mentioned.

Methods

Preparation of Recombinant Sortase A

Cloning of sortase into pGEMT-Easy Vector

Propagation of *Staphylococcus aureus*

Staphylococcus aureus subspecies *aureus* Rosenbach (Designated Mu50, ATCC No. 700699) was obtained from ATCC in lyophilized form. It was revived, as recommended, by propagation in HBI (Heart Brain Infusion) broth, containing 4μg/ml vancomycin, at 37°C with vigorous shaking. A culture grown for 14-16 hrs was used to streak a HBI agar plate. A single *S. aureus* colony was selected for further experiments.

Preparation of Glycerol Stocks

0.5ml of a culture grown for 4-5 hrs was taken in a sterile 1.5ml microcentrifuge tube. 100µl of sterile glycerol was added to it. The microcentrifuge tube was sealed with parafilm and stored at -70°C till further use.

Isolation of Genomic DNA

Genomic DNA of *S. aureus Mu50* strain was isolated using the Wizard Genomic DNA Isolation kit by Promega, following manufacturer's protocol. Briefly, 1ml of *S. aureus* culture (grown at 37°C for about 14-16hrs) was pelleted and resuspended in 50mM EDTA. The cells were then treated with a mixture of lysozyme and lysostaphin for cell wall lysis. Following nuclear lysis and RNase treatment, the protein present in the solution was precipitated and removed by centrifugation. The supernatant, containing the genomic DNA, was transferred to a sterile microcentrifuge tube, precipitated by the addition of isopropanol and resuspended in sterile, nuclease free water. The quality of DNA was ascertained by agarose gel electrophoresis. The DNA was stored at 4°C for immediate use and at -20°C for long term usage.

Agarose Gel Electrophoresis

Agarose gels (1%) were prepared in Tris-Acetate-EDTA (TAE) buffer containing 0.5µg/ml ethidium bromide (EtBr). Samples were prepared by addition of loading dye containing bromophenol blue (BPB) as the tracker dye, and loaded into the wells of the agarose gel. The gel was run at a constant voltage of 80V for 1-3hrs and the DNA was visualized with an ultraviolet (UV) light source.

Polymerase Chain Reaction (PCR) Amplification of *S. aureus* sortase

Gene specific primers were designed and used to PCR amplify the *sortase* sequence from the *S. aureus* genomic DNA. This sequence encodes for amino acids 60-204 of staphylococcal sortase.

The sequences of primers used are as follows:

Upstream Primer (SrtA1):
5' GAT ATA **CAT ATG** CAA GCT AAA CCT CAA ATT CCG 3'

Downstream Primer (SrtA2):
5' GTG GTG **CTC GAG** TTT GAC TTC TGT AGC TAC AAA GAT 3'

These upstream and downstream primers incorporated NdeI and XhoI cleavage sites (in **bold**), respectively, at the 5' and 3' ends of the *sortase* gene for facilitating further cloning into expression vectors.

PCR Reaction Parameters:

Component	Volume (μl)	Final Concentration
10X *Taq* polymerase buffer	5	1X
Deoxynucleotide solution mix	2	400μM
Upstream primer (20μM stock)	2	0.8μM
Downstream primer (20μM stock)	2	0.8μM
Genomic DNA template	variable	5μg/ml
Taq DNA polymerase (5000units/ml)	0.5	2.5U
Nuclease free water	Bring final volume to 50μl	

PCR Cycling Parameters:

Initial Denaturation	94°C	1min 30s	
Denaturation	94°C	30s	X 30 cycles
Annealing	64°C	30s	
Extension	72°C	45s	
Final Extension	72°C	5min	
Storage	4°C	∞	

The PCR amplified product was visualized on 1% agarose gels stained with EtBr. The band corresponding to the amplicon was excised from the agarose gel

using a clean, sharp scalpel blade and the DNA was extracted using agarose gel elution kits provided by Qiagen Inc, USA. The purified PCR product was stored at -20°C till further use.

Cloning of PCR Amplified *S. aureus* sortase

The purified PCR fragment was cloned into the pGEM-T Easy vector (Promega) as per the instructions provided by the manufacturer. Ligation was performed at 4°C for 16hrs followed by transformation into XL1-Blue competent cells of *Escherichia coli*. The cells were plated onto LB agar plate supplemented with 50μg/ml ampicillin.

Preparation of Competent Cells

The following method was used to prepare competent bacterial cells. Competent cells prepared by this procedure were stored at -70°C for no longer than six months.

Requirements:

- $CaCl_2.2H_2O$ (0.1M)
- 80mM $MgCl_2$-20mM $CaCl_2$ solution, ice cold
- Ampicillin solution (100mg/ml); filtered through 0.22μ filter and stored at -20°C till use.
- Luria-Bertani (LB) broth; pH 7.2-7.4 adjusted with sterile NaOH solution; autoclaved
- LB agar plates (2% agar) contáining 50 μg/ ml ampicillin
- Sterile glycerol

A single bacterial colony of XL1-Blue strain of *E. coli* was selected from an LB agar plate that had been incubated for 14-16hrs at 37°C and used to inoculate 5ml of LB broth. This was then placed in an incubator-shaker overnight (225rpm) at 37°C to make the preinoculum. 1ml of the preinoculum was added to 100ml of LB broth (1% by volume; v/v) and incubated at 37°C for 2-3hrs, with shaking, till the A_{600nm} reached 0.5. The bacterial cells were then transferred to sterile, ice-cold 50ml

polypropylene tubes. The cultures were cooled on ice for 10-15min followed by recovery of the cells by centrifugation at 8500rpm for 8min at 4°C. The medium was decanted and the cell pellets were resuspended by vortexing in 10ml ($1/10^{th}$ of original culture volume) of ice-cold $MgCl_2$-$CaCl_2$ solution. After 10min of incubation on ice, the cells were harvested by centrifugation as done previously and resuspended in 1ml of ice-cold 0.1mM $CaCl_2$ containing 20% glycerol, for each 50ml of original culture. The cells were dispensed into aliquots of 100µl each, snap frozen in liquid nitrogen and stored at -70°C till use.

Transformation

100µl aliquots of competent XL1-Blue cells were thawed on ice and DNA (no more than 50ng in a volume of 10µl or less) was added to each microcentrifuge tube. The contents of the tubes were mixed and stored on ice for 30min. A heat shock was given by keeping the tubes at 42°C for 90s. Care was taken not to agitate the contents at this stage. The tubes were then rapidly transferred to an ice bath and the cells were allowed to cool for 1-2min. 900µl of LB broth was added to each tube (bringing the total volume to 1ml) and the cultures were incubated for 45min at 37°C, with shaking, to allow the bacteria to recover as well to express the antibiotic resistance marker encoded by the plasmid (ampicillin). The transformed cells were concentrated by centrifugation at 1000rpm for 10min and resuspended in 100µl of LB broth following which they were plated onto LB agar-ampicillin indicator plates containing 5-bromo-4-chloro-3-indolyl beta-D-galactopyranoside (X-gal) and isopropyl-β-D-galactopyranoside (IPTG).

Blue-White Screening for Recombinant Bacterial Colonies

LB Agar-ampicillin plates for blue-white screening were prepared by coating 100µl of 100mM IPTG and 40µl of 100mg/ml X-gal onto the plates. 100µl of the transformed cells were plated onto the X-gal-IPTG LB agar-ampicillin plates, air-dried and incubated in an inverted position for 14-16hrs at 37°C. Subsequently, the plates were removed from the incubator and stored for several hours at 4°C for the blue color to develop. The recombinant white colonies were then identified.
All the procedures involving IPTG and X-gal were carried out in the dark.

Plasmid Isolation from Recombinant Clones

The recombinant white colonies were inoculated into LB broth containing 50μg/ml ampicillin. After 14-16hrs incubation at 37°C, glycerol stocks of the colonies were made and stored at -70°C. The remaining cells were pelleted and plasmid DNA was isolated using the miniprep plasmid isolation kit (Qiagen). The quality of the DNA was checked on 1% agarose gels stained with EtBr and quantification was done spectrophotometrically by absorbance measurements at 260nm.

Confirmation of Recombinant Clones

The presence of the insert in the candidate recombinant colonies was confirmed by EcoRI restriction analysis, PCR amplification and DNA sequencing.

Protocol for EcoRI restriction analysis:

Constituent	Volume (μl)	Final concentration (in 20μl)
10X EcoRI Buffer	2	1X
Plasmid DNA	Variable	50
EcoRI (20,000U/ml)	0.5	10U
Nuclease Free Water	Make volume up to 20μl	

The reaction was incubated at 37°C for 3hrs. The fragmentation was analyzed on 1% agarose gels stained with EtBr. An appropriate DNA marker was used to help approximate the size of the fragments.

PCR amplification:

The presence of the desired insert in the plasmid isolated from the candidate recombinant colonies was also confirmed by PCR using the plasmid as template and the same primer combination and reaction conditions as employed for PCR of the genomic DNA. The PCR products were analyzed by 1% agarose gel electrophoresis.

Sequencing:

A single pGEMT-*sortase* clone was selected for further studies. The isolated plasmid was sequenced by Microsynth, Switzerland using the commercially available SP6 Promoter Primer.

Cloning of S. aureus sortase into pET23b Expression Vector

Preparation of the insert

The plasmid isolated from the selected clone was used to extract *sortase* insert for downstream cloning. The plasmid was fragmented with restriction enzymes NdeI and XhoI. The cleavage sites for these two enzymes had been engineered during PCR amplification. The buffer system and incubation conditions were used according to manufacturer's protocol.

Protocol for restriction endonuclease analysis:

Constituent	Volume (µl)	Final concentration (in 50µl)
10X NEB Buffer 4	5	1X
pGEMT-*sortase* Plasmid	Variable	3µg
100X BSA	0.5	1X
NdeI (20,000U/ ml)	1.5	30U
XhoI (20,000U/ ml)	1.5	30U
Nuclease Free Water	Make volume up to 50µl	

The reaction was incubated at 37°C for 14-16hrs following which a small fraction was analyzed on a 1% agarose gel stained with EtBr to evaluate the extent of fragmentation. Upon confirmation of complete fragmentation, the remaining fraction was loaded on a preparative 1% agarose gel stained with EtBr. The gel was run at 80V for 1-3hrs to separate the insert from the linear, nicked and supercoiled plasmid species. The DNA was visualized with a UV light source and the band corresponding to the insert was excised using a clean scalpel blade. The DNA was recovered from the gel slice using agarose gel extraction kits provided by Qiagen.

The prepared insert was resuspended in 30µl of sterile, nuclease free water and stored at -20°C till use.

Preparation of Vector

The pET23b vector was obtained from Novagen (0.38µg/µl). 50ng of the plasmid was transformed into XL1-Blue competent cells of *E. coli* and plated on LB agar plate supplemented with 50µg/ml ampicillin. After 14-16hrs of incubation, colonies were inoculated into LB broth containing 50µg/ml ampicillin and incubated at 37°C (225rpm) for another 14-16hrs. About 500µl of the culture was used to make glycerol stocks as previously described and stored at -70°C till use. The remaining culture was used to isolate the pET23b plasmid vector using miniprep kit from Qiagen as described previously. pET23b was linearized by fragmentation with NdeI and XhoI, the same enzymes that were used in preparing the insert, by following a similar protocol. The reaction was analyzed by agarose gel electrophoresis as described earlier. Uncut vector DNA was run in an adjacent lane to help distinguish unfragmented from linearized plasmid. The prepared vector was purified into 30µl of sterile, nuclease free water by using gel elution kits (Qiagen) and stored at -20°C till use. A small fraction was analyzed on 1% agarose gel to confirm presence and purity of the prepared vector.

Ligation of Vector and Insert

The prepared vector and insert were ligated using manufacturer's protocol. The vector and insert ratio used was 1:2.

Protocol for ligation and transformation:

Constituent	Volume (µl)	Final concentration (in 20µl)
10X Ligase Buffer	2	1X
Vector	Variable	50ng
Insert	Variable	100ng
T4 DNA Ligase (Diluted in 1X ligase buffer to 20U/µl)	1	20U
Nuclease free water	Volume to 20µl	

The ligation mixture was incubated at 4°C for 14-16hrs following which it was transformed into 100µl of XL1-Blue competent cells. The transformed cells were plated onto LB agar plate containing ampicillin to a final concentration of 50µg/ ml and kept inverted for 14-16hrs in a 37°C incubator.

Confirmation of Clones

One recombinant clone was selected for further experiments. The recombinant nature of the clone was confirmed by sequencing by Microsynth, Switzerland using the commercially available T7 promoter and terminator primers.

Heterologus Expression and Purification of Sortase

Transformation into BL21 (DE3)

50ng of the plasmid isolated from the selected recombinant XL1-Blue clone was used to transform competent cells of BL21 (DE3) expression strain of *E. coli.*

Induction of Protein Expression

A 1% inoculum of 14-16hrs culture of the recombinant BL21 (DE3) colony was used to inoculate a large scale culture. The culture was grown at 37°C till the A_{600nm} reached 0.6. Protein expression was induced by the addition of IPTG to a final concentration of 0.2mM. The culture was propagated at 30°C for a period of 3hrs following which the cells were harvested and either processed immediately or stored in -20°C for future use.

Sodium Dodecyl Sulfate-Poly Acrylamide Gel Electrophoresis (SDS-PAGE)

The protein expression profile of the induced cells was analyzed on 12% SDS-PAGE. Small-scale (3-5ml) culture pellets were resuspended in 10mM phosphate buffered saline (PBS) and lysed by sonication. Sample buffer containing BPB tracker dye was added to the sample and loaded on SDS-PAGE gels. The gels were run at 125V for 2-4hrs at room temperature. Appropriate protein markers were run in lanes adjacent to the sample to help determine the size of the recombinant protein.

Tris-Glycine-SDS Running Buffer:

25mM Tris, 200mM Glycine, 0.1% SDS; pH 8.4-8.6

5% Stacking Gel Components:

Constituent	Volume (ml)
Water	3.4
30% Acrylamide: Bis acrylamide Solution (29:1 w/w)	0.83
1M Tris-HCl, pH 6.8	0.63
10% SDS	0.05
10% Ammonium Persulfate (APS)	0.05
Tetramethylethylenediamine (TEMED)	0.005

12% Separating Gel Components:

Constituent	Volume (ml)
Water	3.3
30% Acrylamide: Bis acrylamide Solution (29:1 w/w)	4.0
1.5M Tris-HCl, pH 8.8	2.5
10% SDS	0.1
10% Ammonium Persulfate (APS)	0.1
TEMED	0.004

Staining Solution:

0.1% Coomassie Blue R in Methanol, Acetic Acid and water in the ratio 5:1:4 (v/v)

Destaining solution:

Methanol, Acetic Acid and water in the proportion 5:1:4 (v/v)

The gels were stained with coomassie blue followed by destaining for 3-4hrs to remove the non-specific staining. Both staining and destaining were carried out on a rocker at room temperature.

Preparation of Polyethylene Imine (PEI) Solution

4gm of PEI was weighed and diluted to 10ml. This solution was then transferred to a dialysis tubing (MWCO 3,500 Da) and dialyzed against 25mM Tris-HCl buffer (pH 8.0) for at least 24hrs. The PEI solution was then recovered from the tubing and measured. Percentage concentration (weight by volume; w/v) was calculated and the solution was stored at 4°C.

Purification of Recombinant Sortase

For the purpose of protein purification, the cell pellet of large scale culture was resuspended in Buffer A [10mM Tris-HCl, pH 7.5, 50mM NaCl and 1mM β-mercaptoethanol (β-ME)] and lysed by sonication. The cell debris was separated by centrifuging the cell lysate at 10,000rpm for 30min at 4°C. Nucleic acids were precipitated by addition of 0.1% PEI to the supernatant. Following centrifugation, the supernatant was applied to a 5ml Ni-NTA column pre-equilibrated with Buffer A. The column was extensively washed with Buffer B (10mM Tris-HCl, pH 7.5, 500mM NaCl, 30mM imidazole and 1mM β-ME). The protein was eluted with Buffer C (10mM Tris-HCl, pH 7.5, 50mM NaCl, 250mM imidazole and 1mM β-BE).

Fractions containing the protein were pooled and freed of imidazole on a PD-10 desalting column (Amersham) using 50mM Tris-HCl buffer (pH 7.5) containing 150mM NaCl. The protein was further concentrated using protein concentrators (Amicon, MWCO 10KDa), aliquoted and stored at -70°C till use.

Electro Spray Mass Spectrometry (ESMS) and Edman Degradation Analysis of Sortase

The identity of the recombinant sortase was established by ESMS as well as by Edman degradation. A small fraction of the protein was desalted on a C-8 (250 × 4.6 mm) reverse-phase (RP) column using HPLC, lyophilized, resuspended in 50% acetonitrile (ACN) with 0.1% formic acid and analyzed by ESMS in the positive ion mode. The lyophilized protein was also subjected to 25 cycles of N-terminal sequencing using an automated protein sequencer.

Sortase A Catalyzed Transpeptidation Reactions

Synthesis of Peptide Substrates

The peptide substrates were synthesized by standard solid phase peptide synthesis (SPPS) protocols using 9-Fluromethoxycarbonyl (Fmoc) chemistry on a semi-automated peptide synthesizer (Model 90, Advanced Chemtech). Wang resin pre-loaded with the desired amino acid was used as the starting material. The coupling and deprotection was monitored at every step by the Kaiser test for free amines. Before each coupling step and on completion of synthesis, N-terminal Fmoc group was removed using 20% piperidine in DMF (v/v, Dimethyformamide). The peptides were cleaved from the resin and the side chains deprotected with a mixture containing TFA (trifluroacetic acid):ethanediol:phenol:thioanisol:water (80:5:5:5:5, v/v). The resin was removed by filtration and the crude peptides were precipitated using cold diethyl ether. The peptides were purified to ≥98% by RP-HPLC (Reverse phase-high performance liquid chromatography; column C8, 10 mm X 250 mm, gradient: 4-72% B in 130 minutes, A; 0.1% TFA, B; acetonitrile containing 0.1% TFA, flow rate: 1 ml/ min), lyophilized and stored at -70°C. The chemical identity of the peptides was confirmed by mass spectrometry.

Sortase Catalyzed Hydrolysis of LPXTG Motif Peptides

Sortase catalyzed cleavage of LPXTG motif containing peptides was carried out in 300mM Tris-HCl, pH 7.5, containing 150mM NaCl, 5mM $CaCl_2$ and 2mM β-ME. Each assay was set up in a 100µl volume that contained 0.5mM of LPXTG motif peptide and 50µM sortase. The reaction was allowed to proceed at 37°C for 6hrs, quenched by addition of 20-fold excess of chilled 0.1% TFA and analyzed by RP-HPLC. Progress of the reaction was monitored at 210nm on an analytical C-18 column using an acetonitrile-water-TFA based solvent system (gradient: 4-72% B in 130 minutes, A; 0.1% TFA, B; acetonitrile containing 0.1% TFA, flow rate: 1 ml/min). The reaction products were characterized by ESMS.

Sortase Catalyzed Ligation of Peptides

The ligation of LPXTG containing peptides to GGGKY was carried out in 300mM Tris-HCl, pH 7.5, containing 150mM NaCl, 5mM $CaCl_2$ and 2mM β-ME.

Each assay was set up in a 100µl volume that contained 0.5mM of each peptide and 50µM sortase. The reaction was allowed to proceed at 37°C for 6hrs, quenched by addition of 20-fold excess of chilled 0.1% TFA and analyzed by RP-HPLC as described before. The reaction products were characterized by ESMS.

Kinetic Analysis of Sortase Catalyzed Ligation of Abz-LPETG(Dap)Dnp to Peptides

The kinetic assays were performed using Abz-LPETG(Dap)Dnp. The *o*-aminobenzoic (Abz) acid and dinitrophenyl (Dnp) are a florescent donor-quencher team frequently employed as peptide substrates for proteolytic enzymes. The peptide was dissolved in DMSO (dimethylsufoxide) and stored at 4°C. Quantification was done by calculating the absorbance at 365nm (Dnp, ε_{365}nm=17300$M^{-1}cm^{-1}$)[157].

Kinetic assays were performed using desired concentration of substrates in 50µl reaction mixtures containing 300mM Tris-HCl, pH 7.5, 150mM NaCl, 5mM $CaCl_2$, 2mM β-ME, 20% DMSO. Reaction was initiated by addition of 10µM sortase and incubated at 37°C for 30min; 30µl aliquots of the reaction were removed and quenched by the addition of 20-times excess 0.1% TFA. Progress of the reaction was monitored by RP-HPLC at 365nm and the peptides were separated using a linear gradient of 4-72% acetonitirile in 0.1% TFA over 30 minutes. Dnp containing peaks were detected by absorbance at 365nm and the product yield was calculated by integrating the area under the HPLC trace. Kinetic parameters were computed using Graphpad Prism software.

Results

Cloning of Sortase

The nucleotide sequence for *S. aureus Mu50* sortase (NC_002758.2) obtained from the National Center for Biotechnology Information (NCBI) GenBank database (Figure 1A) was used to design gene specific primers for PCR amplification of *sortase*. The primers were engineered with sites for restriction endonuclease NdeI, in case of the upstream primer, and XhoI, in case of the downstream primer.

Genomic DNA isolated from *Staphylococcus aureus* Mu50 was used as template for the PCR amplification of *sortase*. The DNA sequence corresponded to amino acids 60-204 of sortase (Figure 1B, C). The resulting amplicon of 453 base pairs was cloned into the pGEMT-Easy vector (Promega) following manufacturer's protocol. The presence of the insert was confirmed by EcoRI restriction analysis, PCR amplification as well as by DNA sequencing using the commercially available SP6 promoter primer in an automated sequencer (Figure 2, 3).

Expression and Purification of Sortase

The *sortase* insert was sub-cloned into the NdeI/XhoI site of the pET23b expression vector which was then transformed into *E. coli* strain XL1-Blue. The recombinant clones were sequenced using the T7 promoter and terminator primers in an automated sequencer.

The pET23b: *sortase* plasmid was transformed into *E. coli* strain BL21 (DE3) and cells were grown in the presence of IPTG to induce the expression of the 6-His tagged sortase. Sortase expression and purification was monitored by SDS-PAGE analysis (Figure 4). The results presented in Figure 4 (lane2) demonstrate the induced expression of the desired protein in the cell lysate. A certain amount of 'leaky' or uninduced protein expression was also observed (Figure 4, lane1). Cells were lysed by sonication and the recombinant protein present in the supernatant was isolared from the cell lysate using Ni-NTA agarose resin (Figure 4, lane 3). The protein purified by Ni-NTA affinity chromatography migrated as a single band on SDS-PAGE albeit showed a slightly higher molecular weight. RP-HPLC analysis yielded a single peak of the recombinant protein (Figure 5).The ESMS analysis of

A. Complete nucleotide sequence of *S. aureus* sortase (NC_002758.2)

```
ttatttgact tctgtagcta caaagatttt acgtttttcc caaacgcctg tcttttcatt
gtaatcatca caagtaatta atgttaattg tttatcttta ccttttttgtt catctagaac
ttctacatct gttggcttaa catctcttat acttgtcatt ttatacttac gtgtttcatt
accaacttta aagtacacca tactaccttt tttggctgct ttaagatttg taaattgata
gttcggacgg tcaatgaaag tgtgtcctgc aattgaaata ttttgatcat ctagtgattc
attttcttct gcaaagctta cacctctatt taattgttca ggtgttgctg gtcctggata
tactggttct ttaatatcag catctggaat ttcaatatag cctgccactt ttgatttatc
tttcggaatt tgaggtttag cttgctgctt attgtcttta ctcgcctgtt cttttacatt
tttatcatat tgttcaatct tttcatcttt atctttatcg tgaagataat tatcgatatg
tggtttagca aacaaatatg ctgccactag gataagtact acaccagcga ttgtcattaa
tcgatttgtc catttttca t
```

B. Complete protein sequence of *S. aureus* sortase (NP_373052)

mkkwtnrlmtiagvvlilvaaylfakphidnylhdkdkdekieqydknvkeqaskdnkqqakpqipkdkskvagyie
ipdadikepvypgpatpeqlnrgvsfaeenesldqnisiaghtfidrpnyqftnlkaakkgsmvyfkvgnetrkykmts
irdvkptdvevldeqkgkdkqltlitcddynektgvwekrkifvatevk

C. Genomic DNA isolated from *S. aureus* and PCR amplification of *sortase*

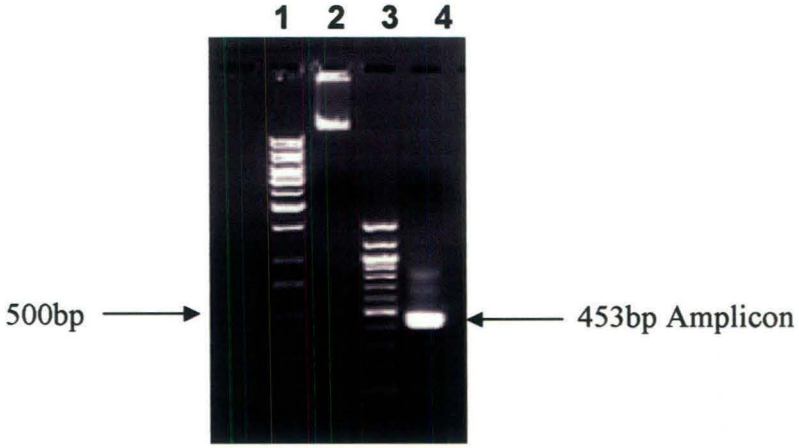


Figure 1. Isolation of genomic DNA from *S. aureus Mu50* and PCR amplification of *sortase* **(A)** The complete nucleotide sequence of *S. aureus Mu50* sortase; **(B)** The complete protein sequence of *S. aureus Mu50* sortase, the 60th and 204th amino acids are highlighted in red; **(C)** Genomic DNA isolated from *S. aureus* and the 453bp PCR amplified fragment corresponding to amino acid 60-204 of sortase. Lane1, 4: 1 kilo base (kb) and 100 base pair (bp) DNA ladder; lane 2: genomic DNA; lane 3: PCR amplified fragment

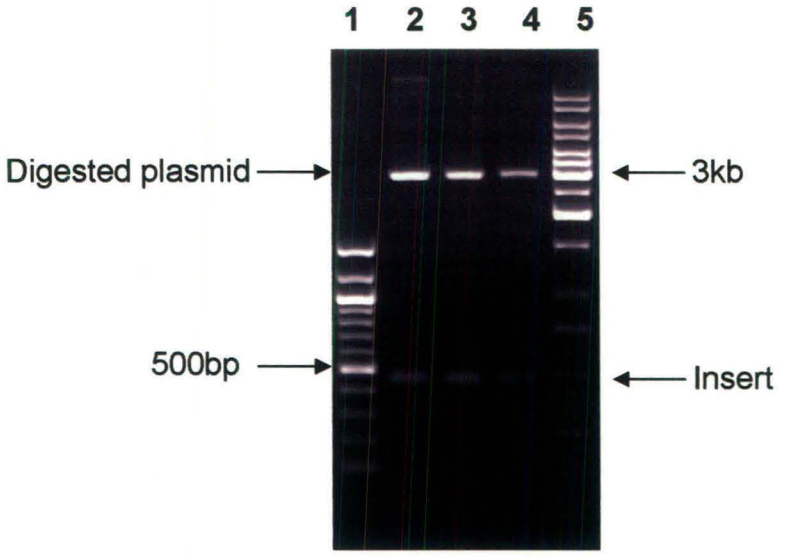


Figure 2. Restriction digest confirmation of pGEMT:*sortase* clones. Plasmids isolated from candidate recombinant clones were digested with EcoRI to check for the presence of insert. Lane 1,5:100bp DNA ladder, 1kb ladder; lane 2,3,4: EcoRI digest of 3 candidate clones The insert can be seen slightly below the DNA ladder band corresponding to 500bp (highlighted by arrow). The digested plasmid can be seen running alongside the 3kb band.

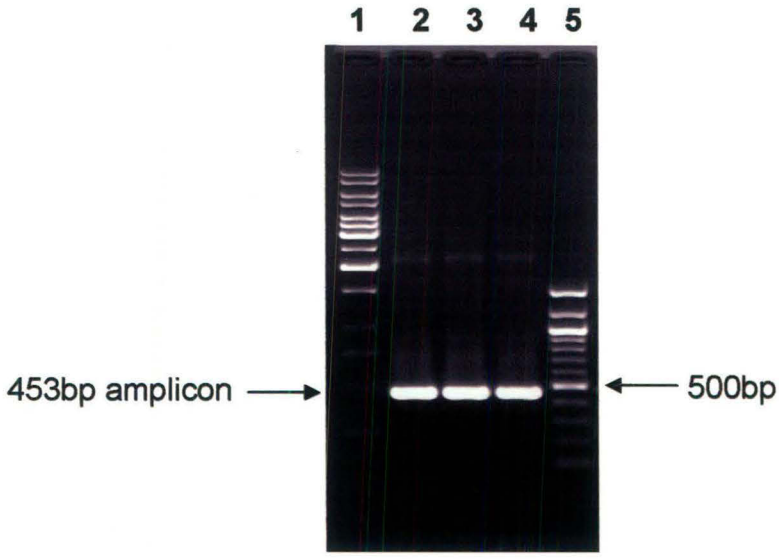


Figure 3. PCR confirmation of pGEMT:*sortase* clones. Plasmids isolated from the candidate recombinant clones were used for PCR amplification using the same primer combination and PCR parameters as was used for the original PCR. Amplicon of desired size confirms presence of insert. Lane 1,5: 1kb DNA ladder, 100bp ladder; lane 2,3,4: PCR amplification of 3 candidate clones. The amplicon can be seen running just below the 500bp marker (highlighted by arrow).

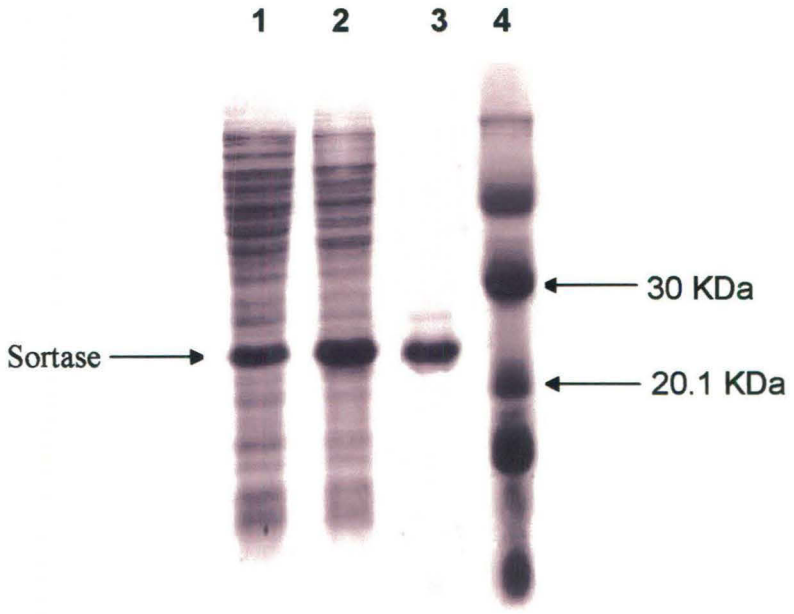


Figure 4. Expression and purification of recombinant sortase. Protein samples were subjected to 12% SDS-PAGE followed by Coomassie Brilliant Blue staining. Lane 1: uninduced cell extract; lane 2: IPTG induced cell extract; lane 3: purified sortase; lane 4: molecular weight marker. Note that the recombinant protein exhibits slightly higher molecular mass than expected.

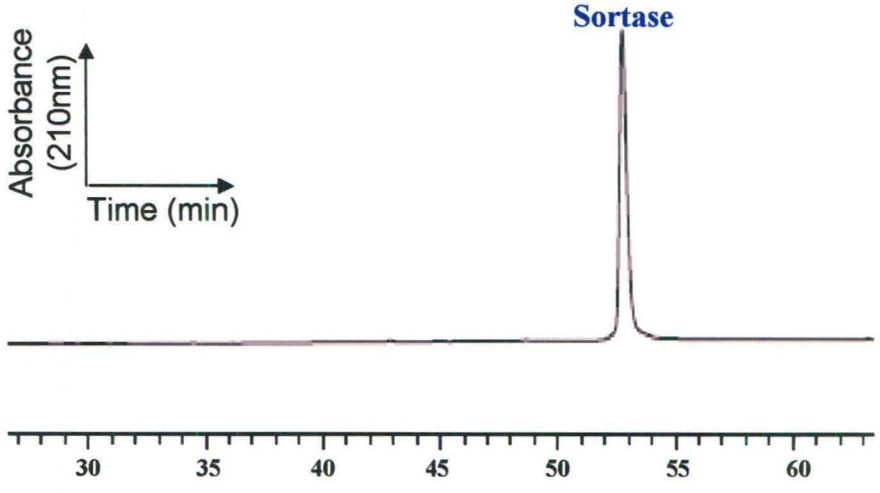


Figure 5. RP-HPLC profile of recombinant sortase.

the recombinant sortase showed a mass of 17864.78 Da which was in agreement with the theoretically calculated value of 17865 Da (Figure 6). 25 cycles of Edman degradation also gave the expected sequence results (Figure 7). Taken together the results establish the integrity of the primary sequence of the recombinant protein.

Synthesis and purification of peptide substrates

Sortase substrates YALPMTGK, YALPETGK and triglycine (GGGKY) were synthesized by SPSS as described in Methods. Following cleavage from the resin, the peptides were purified by RP-HPLC to ≥ 98% homogeneity (Figure 8). The identity of the peptides was verified by MALDI-TOF analysis. The experimentally determined mass (exp) was in agreement with the theoretically calculated values (calc) for the synthesized peptides (YALPMTGK-exp: 880.07 Da, calc: 879.53 Da; YALPETGK-exp: 877.55 Da, calc: 877.99 Da; GGGKY-exp: 480.26 Da, calc: 480.23 Da).

Sortase catalyzed proteolysis of LPXTG substrates

The proteolysis or the ability of sortase to cleave the Thr-Gly peptide bond in a peptide containing the LPXTG motif was tested. RP-HPLC analysis of a mixture of the LPXTG motif peptide (YALPMTGK/YALPETGK) and sortase incubated for 6hrs revealed the presence of a new peak (Figure 9). By ESMS analysis, the mass of the product peaks were determined to be 694.38 Da and 693.04 Da which corresponded to the mass of YALPMT and YALPET respectively (YALPMT-exp: 694.38 Da, calc: 694.34 Da; YALPET-exp: 693.04 Da, calc: 692.34 Da) indicating cleavage at the Thr-Gly peptide bond. The yield of the hydrolyzed product was about 8-10%.

Sortase-catalyzed transpeptidation

The ability of sortase to ligate LPXTG motif containing peptides to a triglycine nucleophile (GGGKY) was investigated. Analyses of the reaction products by RP-HPLC revealed the formation of LPXT-GGGKY as expected (Figure 10). ESMS analysis confirmed the results as the calculated masses were in accord with the experimentally determined values. (YALPMTGGGKY- exp: 1157.28 Da, calc: 1156.56 Da; YALPETGGGKY-exp: 1155.00 Da, calc: 1154.56

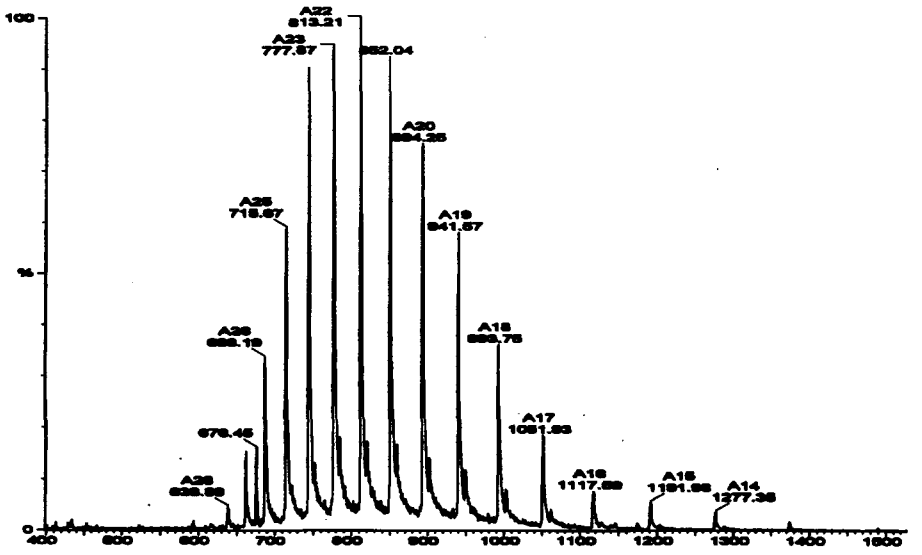


Figure 6. ESMS analysis of recombinant sortase. The experimentally determined mass of 17864.88 Da is in accordance with the calculated value of 17865 Da.

AAcid #	AAcid ID	R.Time (min)	C.Time (min)	Pmol (raw)	Pmol (-back)	Pmol (+Lag)	AAcid ID
1	M	13.08	13.07	2094.03	2080.51	2200.63	Met
2	Q	5.69	5.67	1046.56	1038.12	1038.12	Gln
3	A	8.60	8.23	2289.92	2278.33	2391.23	Ala
4	K	17.01	16.92	1988.23	1843.38	2020.39	Lys
5	P	12.58	12.55	1436.85	1401.05	1639.85	Pro
6	Q	5.69	5.67	1456.74	1429.67	1582.95	Gln
7	I	16.69	16.68	1533.27	1479.20	1637.79	Ile
8	P	12.58	12.55	924.27	920.39	1017.47	Pro
9	K	16.93	16.92	1235.57	1013.42	1122.08	Lys
10	D	4.23	4.24	1066.01	946.73	1048.24	Asp
11	K	16.95	16.92	1435.61	1182.55	1298.00	Lys
12	S	5.40	5.37	627.33	592.11	655.60	Ser
13	K	16.94	16.92	1522.59	1238.61	1358.18	Lys
14	V	13.39	13.39	908.40	724.00	801.71	Val
15	A	8.24	8.23	838.86	745.90	825.88	Ala
16	G	6.15	6.13	668.99	574.07	636.50	Gly
17	Y	10.17	10.16	727.53	631.32	699.01	Tyr
18	I	16.70	16.68	765.23	619.10	685.48	Ile
19	E	6.43	6.39	575.90	390.55	432.39	Glu
20	I	16.69	16.68	1075.62	912.75	1003.60	Ile
21	P	12.58	12.55	540.72	328.47	363.34	Pro
22	D	4.26	4.24	788.50	522.30	578.25	Asp
23	A	8.24	8.23	632.56	485.35	485.35	Ala
24	D	4.26	4.24	1049.58	758.90	778.47	Asp
25	I	16.72	16.68	620.77	416.07	416.07	Ile

Figure 7. N-terminal sequencing of recombinant sortase. N-terminal protein sequencing was done to confirm the identity of the recombinant sortase. 25 cycles of sequencing were performed and the sequence was identical to the sequence registered with the NCBI database.

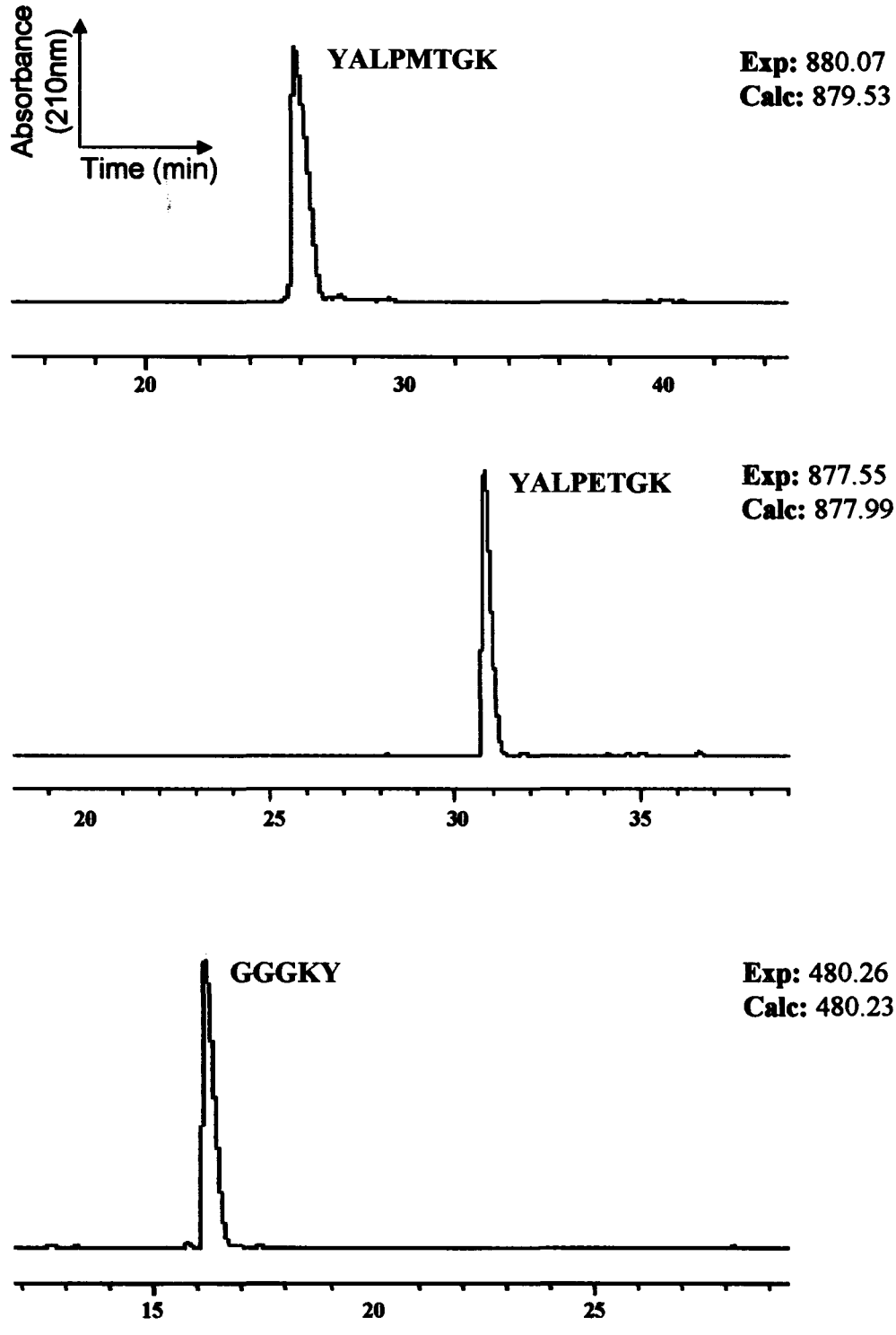


Figure 8. RP-HPLC traces of purified YALPMTGK, YALPETGK and GGGKY. The crude peptides were cleaved from the resin and purified to ≥ 98% purity by RP-HPLC.

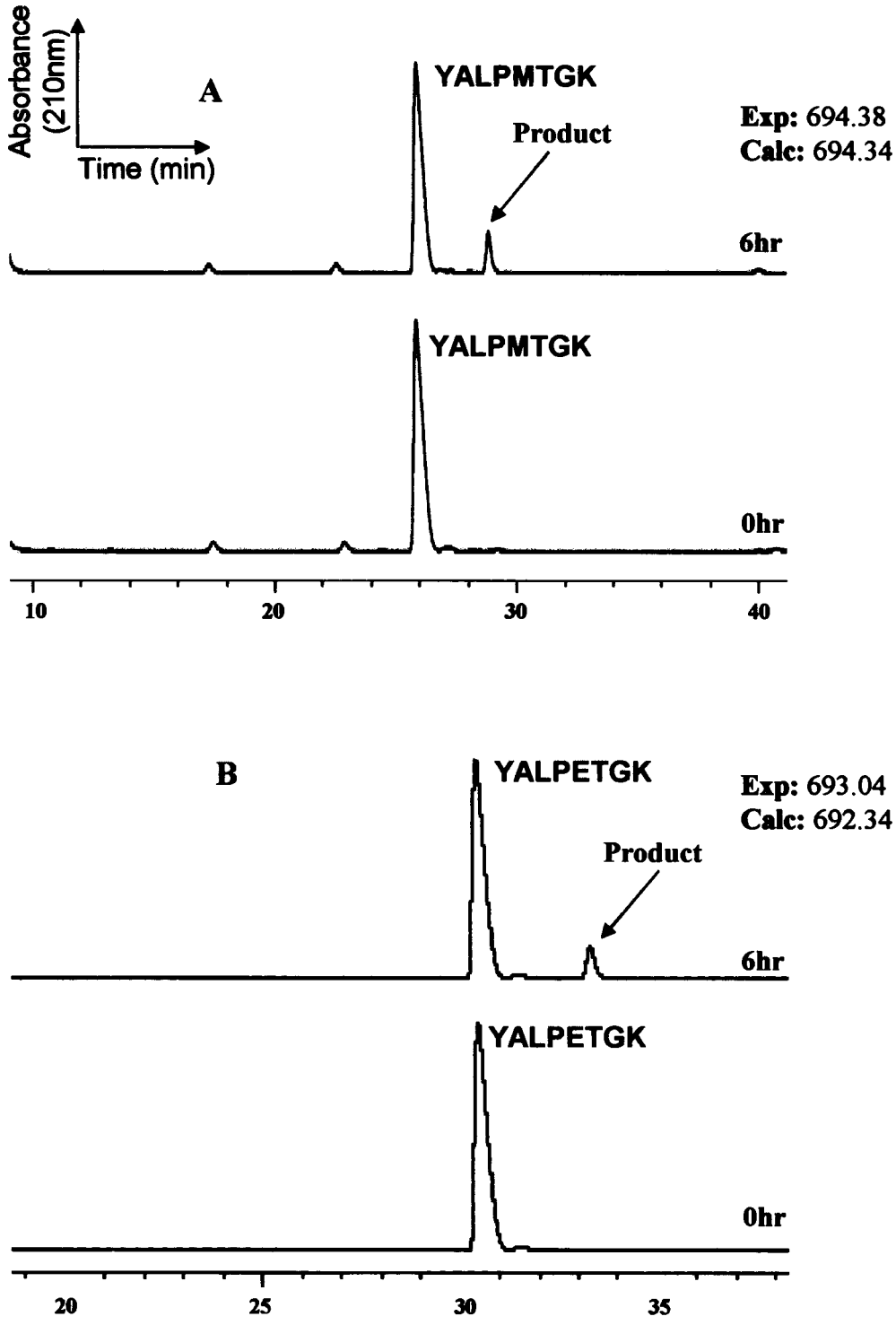


Figure 9. RP-HPLC profile of sortase catalyzed proteolysis of LPXTG peptides. Panel **A** shows the 0hr and 6hr RP-HPLC profile for YALPMTGK; Panel **B** represents the proteolysis of YALPETGK. New peaks corresponding to the cleaved product can be observed in the respective 6hr profiles (marked by arrow).

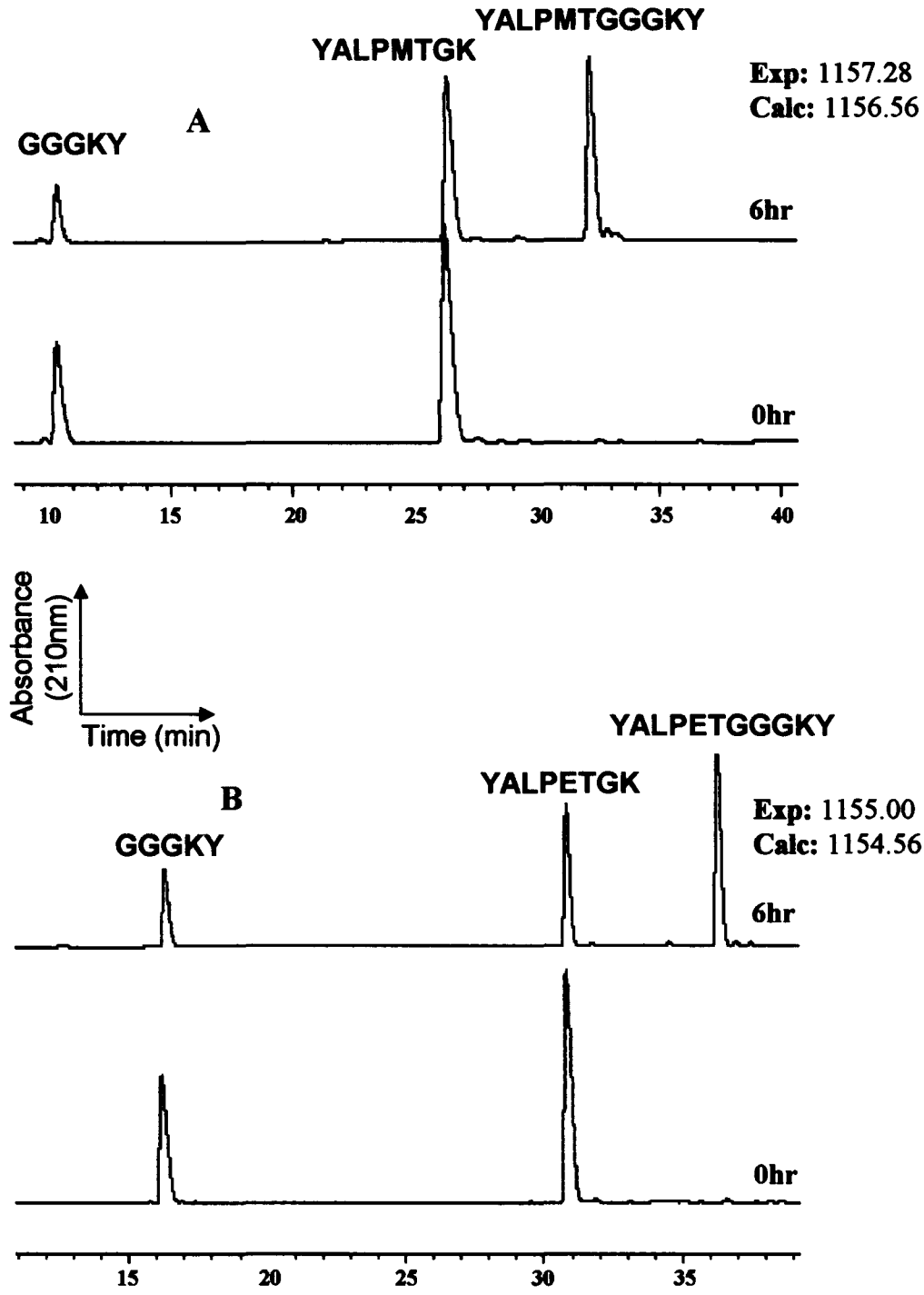


Figure 10. RP-HPLC profile of sortase catalyzed transpeptidation reactions. Panel **A** shows the 0hr and 6hr RP-HPLC profile for the reaction between YALPMTGK and GGGKY; Panel **B** represents the transpeptidation of YALPETGK and GGGKY. New peaks corresponding to the ligated products can be observed in the respective 6hr profiles.

Da). The yield of the transpeptidation product for both LPXTG motif peptides with GGGKY was about 40-50%.

Kinetic analysis of Sortase catalyzed ligation of Abz-LPETG(dap)Dnp to peptides

For kinetic analysis of the transpeptidation reaction, the Abz-LPETG(Dap)Dnp peptide was used (Figure 11, 12). The progress of the reaction was analyzed by RP-HPLC by monitoring at 365nm. For the GGGKY peptide, kinetic analysis yielded K_m and K_{cat} values of 0.099mM and 5.294×10^{-3} s^{-1} respectively (Figure 13).

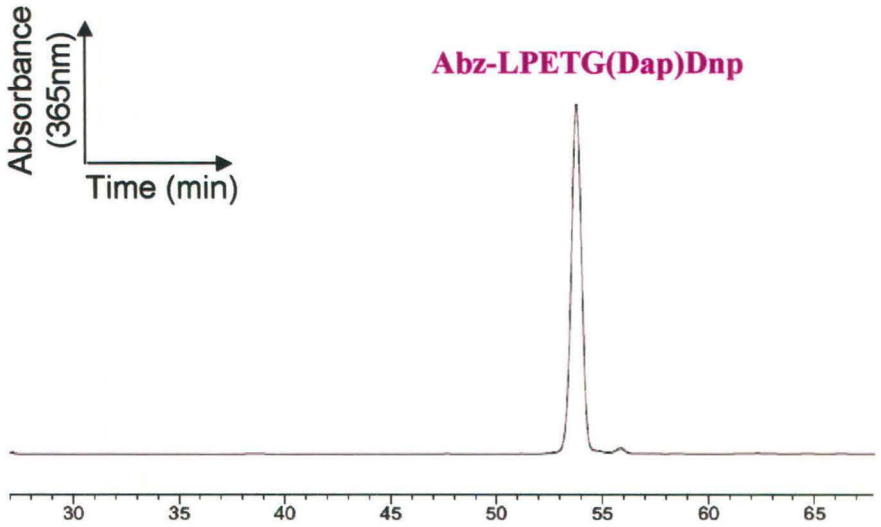
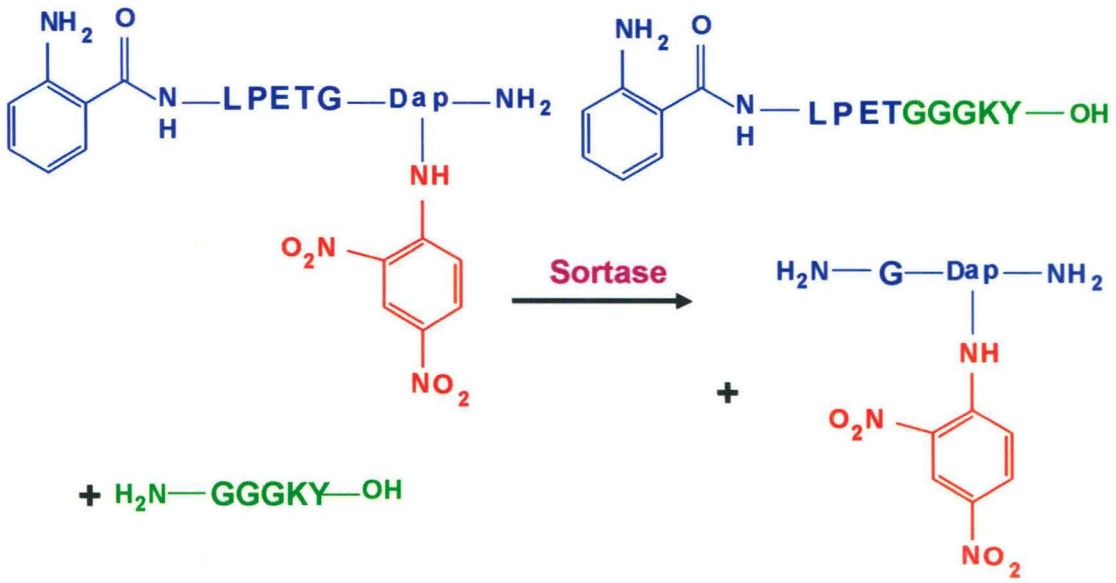


Figure 11. Elution profile of Abz-LPETG(Dap)Dnp. The RP-HPLC run was monitored at 365nm.



Abz-LPETG(Dap)Dnp + GGGKY → Abz-LPETGGGKY + G(Dap)Dnp

(Monitored at 365nm) (Monitored at 365nm)

Figure 12. Schematic representation of sortase catalyzed transpeptidation reaction of Abz-LPETG(Dap)Dnp and GGGKY

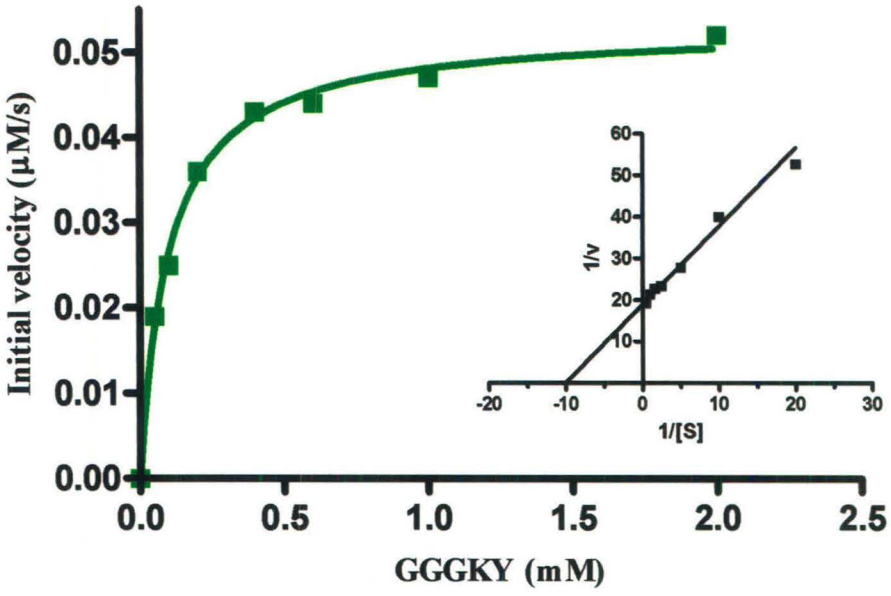


Figure 13. Determination of kinetic parameters for sortase (GGGKY). Estimates of kinetic parameters of K_m=0.099mM and K_{cat}=5.294 × 10^{-3} s-1 were obtained. Inset shows the Lineweaver-Burke plot.

Discussion

The recombinant sortase that encodes for the amino acids 60-204 which represents the catalytic core of the enzyme was cloned from *Staphylococcus aureus Mu50* and expressed as a 6-His tagged protein and purified by Ni-NTA affinity chromatography. The sequence integrity of the enzyme was established by Edman degradation and mass spectrometry.

Two peptides containing the LPXTG-motif were synthesized by standard Fmoc chemistry. The sequences of both peptides were identical except for the third position of the motif; YALPETGK had a glutamine at the third position while YALPMTGK had a methionine. In addition, a Tyr residue was placed at the N-terminus of the peptides to facilitate concentration determination. Although pentaglycine acts as the nucleophile *in-vivo*, it is known that tri-, di- or even mono-glycine can function as nucleophiles in the transpeptidation reaction[158]. Hence, to substitute for pentaglycine-lipid II, a triglycine, GGGKY, was also synthesized. To monitor the progress of the sortase catalyzed reaction, a RP-HPLC based assay was established. As the substrate peptides and products have different elution times, progress of the reaction was easily followed by analyzing aliquots at different time points.

When sortase was incubated with the LPXTG-motif peptides alone, formation of the product (i.e., YALPMT and YALPET) due to cleavage of the T-G peptide bond was observed. In the presence of GGGKY, sortase catalyzed the transpeptidation reaction. The yields for the reactions were in the range of 8-10% for proteolysis and 40-50% for transpeptidation.

To further investigate the transpeptidation reactions, kinetic analysis of the sortase catalyzed reactions was carried out. The peptide used for this purpose, Abz-LPETG(Dap)Dnp, is a fluorescently labeled peptide suitable for florescence resonance energy transfer (FRET) analysis. It was procured with the objective of establishing a florescence quenching based assay for the analysis of sortase activity. However, while the work was in progress, a report by Kruger *et al*[159] described marked fluorescence inner filter quenching in such assays. This quenching resulted in prematurely hyperbolic velocity versus substrate profiles causing underestimation

of the true kinetic parameters. Employing the same substrates, they developed an alternative methodology based on a discontinuous HPLC based system to monitor the sortase catalyzed reaction. Their analysis, using the HPLC assay, revealed K_m and K_{cat} values of 5.5mM and $0.27s^{-1}$ respectively for Abz-LPETG(Dap)Dnp. For Gly_5 substrate K_m value reported was 140μM[159] that compares well with a value of 100μM determined for Gly_3KY in the present study. Hence, the recombinant sortase enzyme prepared here is fully functional capable of catalyzing both the proteolysis and the transpeptidation.

Chapter II

Sortase Catalyzed Ligation to Aminosugars and Aliphatic amines

Introduction

Enzymes are traditionally considered to be highly specialized biocatalysts, capable of rapid conversion of a specific substrate into a specific product with high efficiency. In recent years, however, there has been a growing realization that this picture is oversimplified. A large number of enzymes are now known to serve additional functions: several enzymes are found to 'moonlight' i.e. serve structural or regulatory purposes unrelated to enzymatic activity; in addition, many enzymes are 'catalytically promiscuous' i.e. capable of processing or reacting with multiple substrates at an active site that is specialized to catalyze a primary reaction[160, 161].

Catalytic promiscuity is the ability of an enzyme to catalyze an adventitious secondary activity at the active site responsible for the primary reaction. Promiscuity arises because active sites are inherently reactive environments packed with nucleophiles, electrophiles, acids, bases, cofactors etc. Another reason for promiscuity is that most active sites are endowed with a certain amount of flexibility as a result of which several compounds, other than their natural substrate, can be accommodated.

Substrate promiscuity can be very useful in synthetic chemistry and biology because richness of substrate can lead to a range of useful products. The product diversity obtained through a chemoenzymatic route becomes extremely attractive in the synthesis of complex and high-value molecules especially where chemical routes are difficult to implement[162, 163]. Enzymes also have the added advantage of operating most effectively under mild, physiological conditions which minimizes problems associated with undesired side-reactions such as decomposition, isomerization, racemization and rearrangement, which often plague traditional chemical methodology.

In the *in-vivo* sortase catalyzed transpeptidation reaction, the terminal amine of the Gly_5 cross bridges of Lipid II acts as the nucleophile in resolving the acyl-enzyme intermediate formed between the carboxyl group of Thr and the thiol of the active-site Cys. Studies have revealed that peptidoglycan surrogates like Gly_5, Gly_4, Gly_3 and Gly_2 can efficiently stimulate sortase activity *in-vitro*. Even a single Gly molecule can function as a nucleophile *albeit* with 10-fold less efficiency[156]. This observation of seemingly relaxed substrate specificity hinted at catalytic promiscuity

and suggested that other amino compounds may also function as nucleophilic substrates. Accordingly, sortase catalyzed ligation of LPXTG substrates to aminosugars has been explored with a view to develop a chemo-enzymatic approach for neoglycopeptide synthesis. The ability of certain aliphatic amines to act as nucleophile is also investigated.

Materials and Methods

Materials

Staphylococcus aureus Mu50 was procured from ATCC (ATCC No. 700699). Primers were obtained from Microsynth, Switzerland. Genomic DNA isolation kit and pGEM-T Easy vector were supplied by Sigma, USA. All enzymes, PCR reagents, DNA and protein markers were procured from New England Biolabs, USA. Miniprep kit and gel elution kits were purchased from Qiagen, USA. HBI broth and agar and LB broth and agar were obtained from BD, USA. pET23b vector was supplied by Novagen. All Fmoc protected amino acids and amino acid substituted Wang resins were purchased from Novabiochem. DCM, DMF and acetonitrile were obtained from Merck. 6-deoxy-6-amino mannose and 6-deoxy-6-amino glucose were supplied by Glycoteam, Germany. Abz-LPETG(Dap)Dnp was obtained from AnaSpec, USA. Aminoglycoside antibiotics, glucosamine, DMSO and other chemicals were obtained from Sigma, unless otherwise mentioned, and used without further purification.

Methods

Synthesis of Peptide Substrates

All peptide substrates were synthesized and purified as described earlier.

Sortase catalyzed ligation of peptides to aminosugars, amines and aminoglycosides

The ligation of YALPXTGK (X= E or M) peptide to aminosugars, aminoglycosides and other amino compounds was carried out in 300mM Tris-HCl, pH 7.5, containing 150mM NaCl, 5mM $CaCl_2$ and 2mM β-ME. Each assay containing YALPXTGK peptide (0.5mM), amine nucleophile (2.5mM) and sortase (50μM) was set up to a total volume of 100μl volume. The reaction was allowed to proceed at 37°C for 6hrs. The reaction products were analyzed and characterized as

described before. All RP-HPLC runs were monitored at 210nm unless otherwise mentioned.

Kinetic analysis of sortase catalyzed ligation of Abz-LPETG(Dap)Dnp to 6-aminosugars and aminoglycosides

Kinetic assays were performed in a 50µl reaction volume containing 300mM Tris-HCl, pH 7.5, 150mM NaCl, 5mM $CaCl_2$, 2mM β-ME, 20% DMSO. The concentration of Abz-LPETG(Dap)Dnp was kept constant at 1mM. Reactions were initiated by addition of 10µM sortase and incubated at 37°C for 30min; 30µl aliquots of the reaction were removed and quenched by the addition of 20-times excess chilled 0.1% TFA. Progress of the reaction was analyzed on RP-HPLC by monitoring at 365nm and the peptides were separated using a linear gradient of 4-72% acetonitirile in 0.1% TFA over 30min. Dnp containing peaks were detected by absorbance at 365nm and the percentage of substrate converted to product was calculated by integrating the area under the HPLC trace. Kinetic parameters were computed using Graphpad Prism software.

Preparation of Recombinant Multidrug resistance Protein (Mrp)

The Mrp protein (NP_372281) of *Staphylococcus aureus* that carries a sortase recognition pentapeptide sequence motif (LPNTG) near the C-terminus was cloned expressed and purified similar to sortase. Primers, 5' **GCT AGC** GTG CAA AAT TAT CGA AAA GTA AGT A 3' and 5' **CTC GAG** TGA TTC TTT TTC GTT TTT AGT ACG T 3', were used for PCR amplification of the gene encoding Mrp protein. The primers were engineered with cleavage sites for the restriction endonucleases NheI and XhoI respectively (in **Bold**). The amplified PCR product was ligated into the pGEMT-Easy vector and the identity of the insert verified by EcoRI cleavage, PCR amplification and DNA sequencing. Subsequently, the insert was ligated into the NheI-XhoI site of the pET23b expression vector and transformed into *E. coli* strain BL21 (DE3) for expressing recombinant Mrp with a C-terminal hexa-histidine tag. The expression was induced by addition of IPTG to a final concentration of 0.5mM for 6hrs and the protein was purified by standard Ni-NTA affinity chromatography.

Sortase catalyzed ligation of aminoglycosides to recombinant Mrp

About 40μM Mrp was incubated with 80μM sortase at pH 7.5 in the presence of desired aminoglycosides for 6hrs at 37°C. The reaction was quenched by addition of chilled 0.1% TFA. The products were analyzed by RP-HPLC and SDS-PAGE. The identity of the reaction products was confirmed by mass spectrometry.

Results

Sortase catalyzed ligation of LPETG peptides to aminohexoses

6-aminohexoses, such as 6-deoxy-6-aminoglucose and 6-deoxy-6-aminomannose were considered as amine nucleophile in the transpeptidation reaction with the idea that $-CH_2-NH_2$ moiety present in these sugars might mimic some elements of the glycine structure (Figure 1). The potential of sortase to ligate the aforementioned aminosugars to a model YALPXTGK (X = E or M) peptide substrate was tested. HPLC assays followed by MALDI-TOF or ESMS analysis revealed the formation of respective YALPET-sugar adducts with both aminohexoses, suggesting that the above amino sugars indeed acted as nucleophiles in the transamidation reaction (Figure 2, 3, 4; Table 1).

Time course analysis of sortase catalyzed ligation of YALPETGK to 6-amino-6-deoxymannose showed that the reaction reached equilibrium in about 6-8hrs accompanied by a product yield of ~ 34 % (Figure 5, Table 1). Kinetic analysis of the sortase catalyzed transpeptidation to 6-deoxy-6-amino mannose using Abz-LPETG(Dap)Dnp as substrate was carried out. The K_m and K_{cat} for aminomannose was determined as 4.85mM and 0.0033 s^{-1} respectively (Figure 6). Thus, the K_m for the amino sugar is quite weak (about 500-fold) as compared with GGGKY substrate. The K_{cat} values are, however, comparable.

2-deoxy-2-aminoglucose (glucosamine) was also tested as a potential nucleophile to further elaborate the specificity of sortase for the aminohexose. However, incubation of YALPETGK and glucosamine in the presence of sortase yielded only YALPET without the formation of the YALPET-sugar adduct suggesting that the substrate peptide was hydrolyzed and that 2-amino group could not function as an amine donor in the transpeptidation reaction (Figure 2, 3).

Sortase catalyzed ligation of LPETG peptides to aminoglycosides

To further probe the specificity requirements as well as to see if 6-aminohexoses can serve as recognition tags for peptide-sugar ligations, the ability of sortase to ligate peptides to the aminoglycoside class of therapeutically important antibiotics was investigated. These antibiotics are built up by a variety of amino

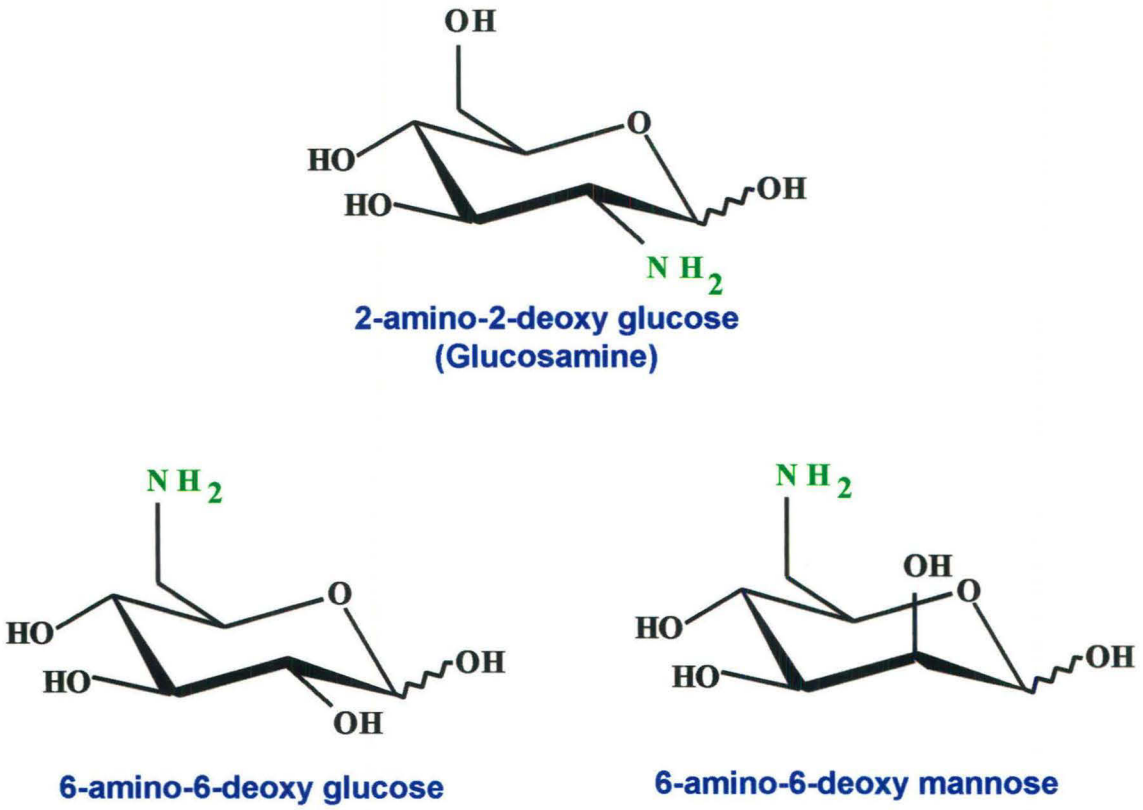


Figure 1. Aminosugar substrates tested for sortase catalyzed transpeptidation.

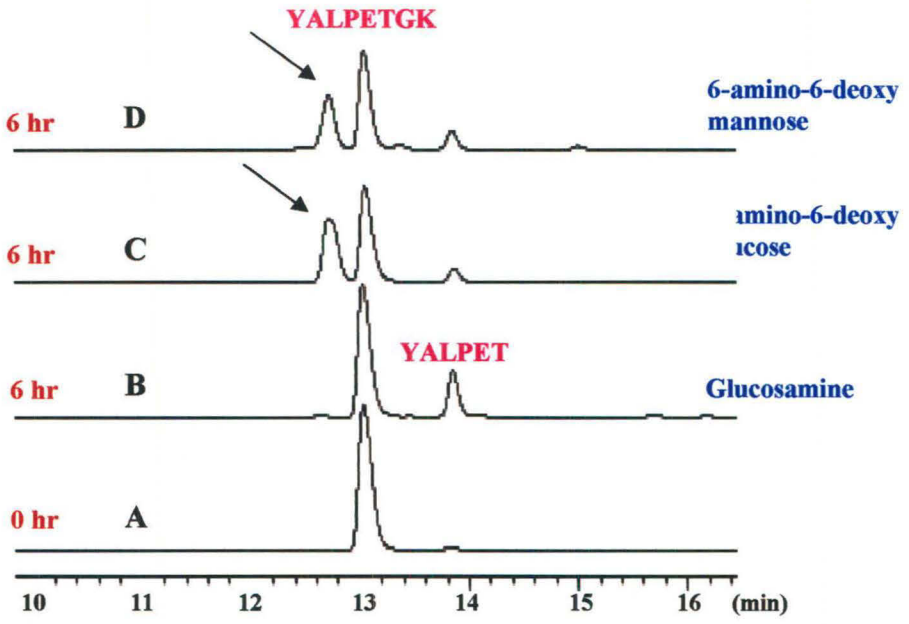


Figure 2. Sortase catalyzed ligation of YALPETGK to 6-aminohexoses. Panel **A** shows the 0hr HPLC profile; panel **B, C, D** show the 6hr profiles of the reaction of YALPETGK with glucosamine, 6-aminoglucose and 6-aminomannose respectively. Conjugates are highlighted by arrows. The HPLC runs were monitored at 210nm.

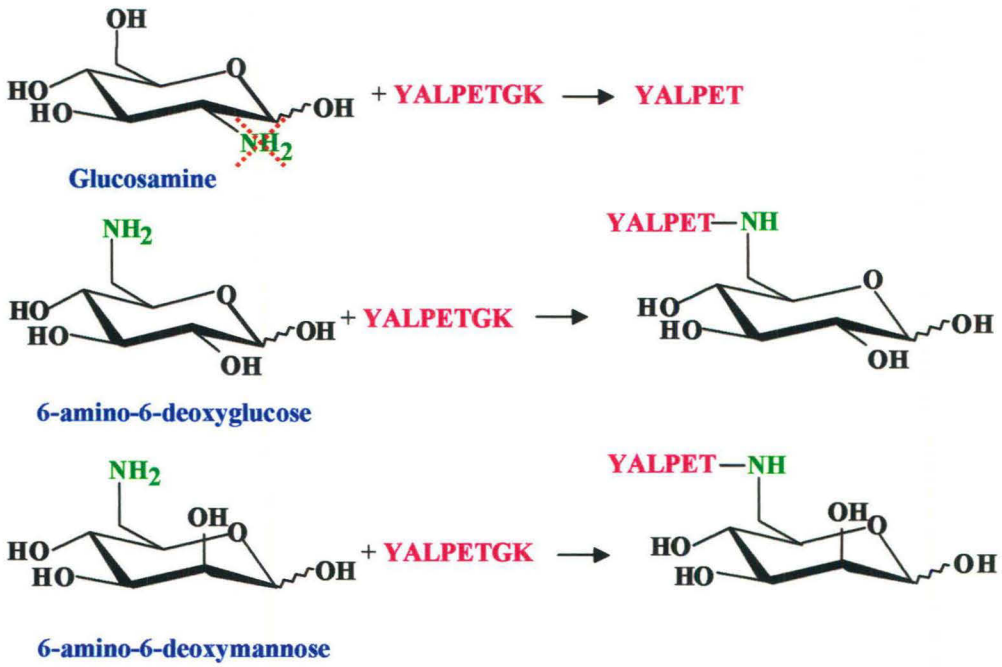


Figure 3. Diagrammatic representation of sortase catalyzed ligation of YALPETGK to 6-aminohexoses.

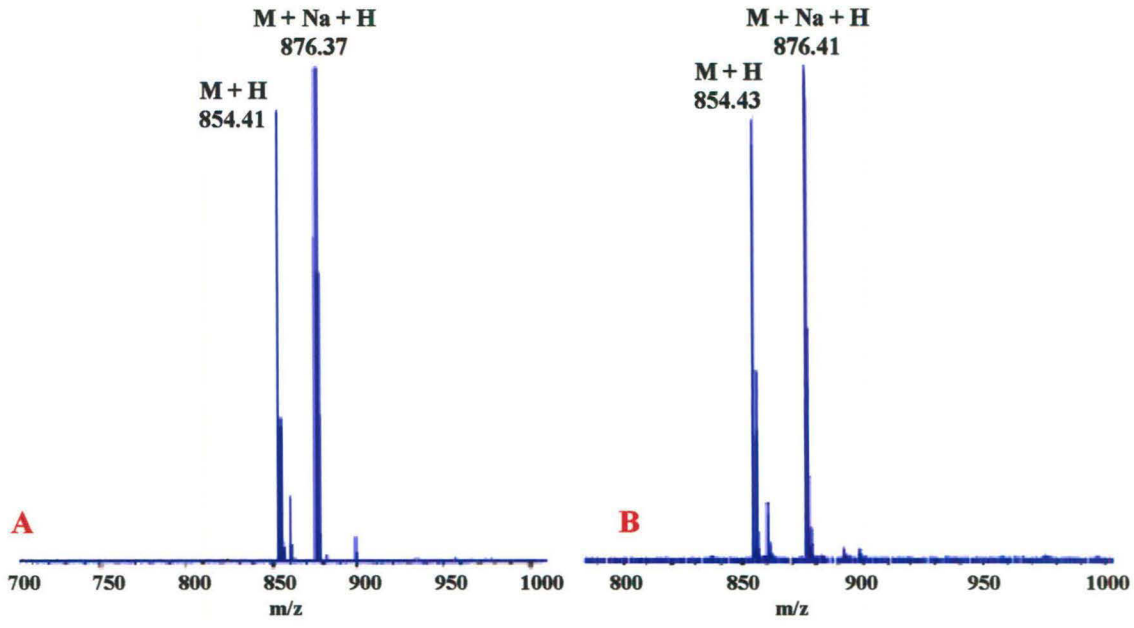


Figure 4. ESMS analysis of YALPET-6-deoxy-6-aminoglucose **(A)** and YALPET-6-deoxy-6-aminomannose **(B)**. The experimentally determined mass of 853.41 Da and 853.43 Da respectively is in accordance with the calculated mass of 853.41 Da for the peptide-sugar adduct.

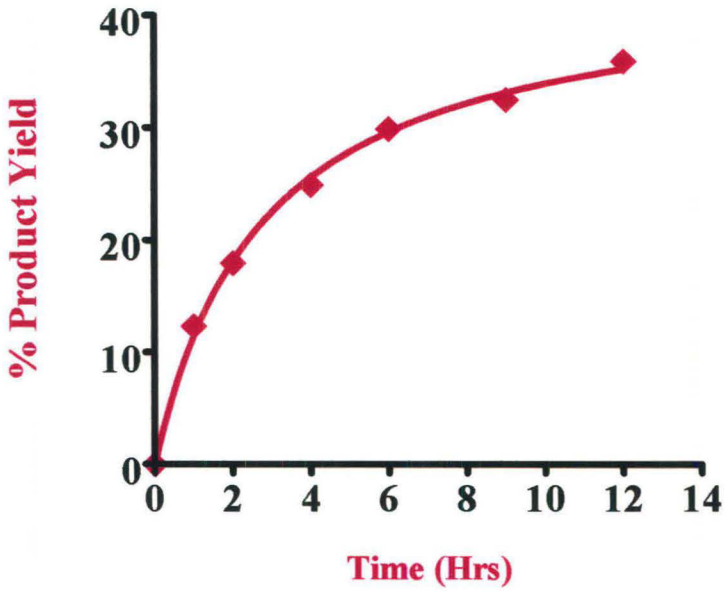


Figure 5. Time-course analysis of sortase catalyzed ligation of YALPETGK with 6-deoxy-6-amino mannose. The reaction reaches equilibrium between 6-8 hours.

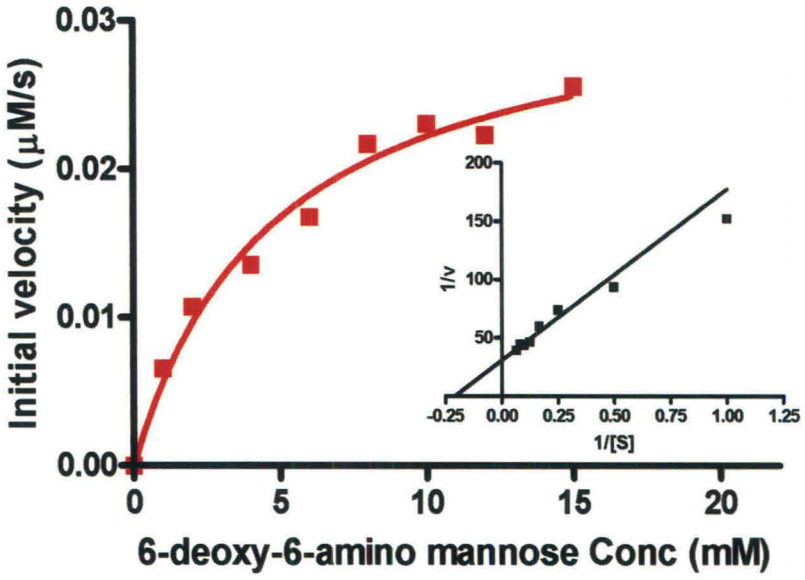


Figure 6. Determination of kinetic parameters for sortase (6-amino-6-deoxy mannose). Kinetic estimates of K_m= 4.855mM and K_{cat}= 0.003299 s^{-1} were obtained. Inset shows the Lineweaver-Burke plot.

Table 1. Characterization (mass spectrometry and yield) of sortase catalyzed ligation products formed from LPXTG containing peptides and aminohexose nucleophiles

Peptide	Aminohexose	Experimental (Da)	Calculated (Da)	Yield (%)
YALPETGK	6-deoxy-6-amino glucose	853.41	853.41	41.7
YALPETGK	6-deoxy-6-amino mannose	853.43	853.41	33.7
YALPMTGK	6-deoxy-6-amino glucose	855.62	855.41	54.2
YALPETGK	KanamycinA	1158.64	1158.56	42.4
YALPETGK	KanamycinB	1157.65	1157.58	38.6
YALPETGK	Tobramycin	1141.66	1141.58	73.5
YALPETGK	Ribostamycin	1128.62	1128.56	21.7
YALPETGK	Neomycin	1288.72	1288.63	29.1
YALPETGK	Paromomycin	1289.73	1289.62	17.7
YALPMTGK	KanamycinA	1160.76	1160.56	36.1
YALPMTGK	KanamycinB	1159.64	1159.57	21.7
YALPMTGK	Tobramycin	1143.66	1143.58	58.8
YALPMTGK	Ribostamycin	1130.61	1130.55	22.1
YALPMTGK	Neomycin	1290.73	1290.63	28.2
YALPETGK	Amikacin	A1: 1259.52 A2: 1259.56	1259.87	Total: 38.56 A1: 24.40 A2: 14.16
YALPETGK	ButirosinA	B1: 1229.54 B2: 1229.60	1230.37	Total: 36.90 B1: 18.48 B2: 18.41

Reaction conditions: peptide, 0.5mM; aminosugar/antibiotics,2.5mM; sortase,50μM; 37°C; pH 7.5. The reaction was carried out for 6hrs. The product yield is calculated based on HPLC peak areas.

sugars of the 6-amino or the 2, 6-diamino type besides containing several other amino functionalities[164,165]. The central scaffold of aminoglycoside antibiotics is the 2-deoxystreptamine (2-DOS) ring to which amino sugars are substituted at positions 4 and 6 (as in tobramycin and the kanamycins) or 4 and 5 (as in ribostamycin, neomycin, and paromomycin) (Figure 7, 8). Sortase mediated ligation of model YALPXTGK (X = E or M) peptide substrate to aminoglycoside antibiotics proceeded smoothly (Figure 9). Analysis of the reaction products by RP-HPLC followed by MALDI revealed the formation of specific conjugates between antibiotics and peptide in the yields varying from 35 to 70% for the kanamycin class, and about 18-30% for the ribostamycin class of antibiotics (Figure 11, Table 1). ESMS of the respective conjugates produced fragmentations that unambiguously showed occurrence of peptide ligation exclusively at a single 6-amino site in ring A of the kanamycins, tobramycin, and ribostamycin or ring D of paromomycin and neomycin (Figure 10, 12-15). Thus, conjugation of peptide substrates was limited to the 6-amino site in the antibiotics despite the presence of a plethora of amino groups, indicating rather strict specificity and selectivity for the sugar amino groups by sortase.

The kinetic analysis of the sortase catalyzed ligation of peptides to tobramycin and neomycin yielded K_m of 2.79mM and 2.57mM, and K_{cat} of 0.0068 s^{-1} and 0.0025 s^{-1} respectively (Figure 16, 17). The K_m and K_{cat} of the antibiotics are comparable or slightly better than that of the aminosugars.

Sortase catalyzed ligation of peptides to amikacin and butirosinA

Amikacin and butirosinA are aminoglycoside antibiotics substituted at the 4, 6 and 4, 5 position respectively. These antibiotics, of the kanamycin and ribostamycin class respectively, are unique in that they carry a 4-amino-2-hydroxybutyrate (A2HBA) side chain on the C-1 amino group of the central deoxystreptamine ring (Figure 18). Unlike the peptide ligation reactions with antibiotics described above, the HPLC chromatographic profile of sortase peptide substrates, YALPETGK, reacted with amikacin or butirosin B showed two product peaks (Figure 19). Interestingly, when analyzed by MALDI, both the product peaks of each antibiotic yielded identical mass (~1260 Da in the case of amikacin and ~1230 Da in the case of butirosinA) corresponding to a 1:1 ligation of the peptide to

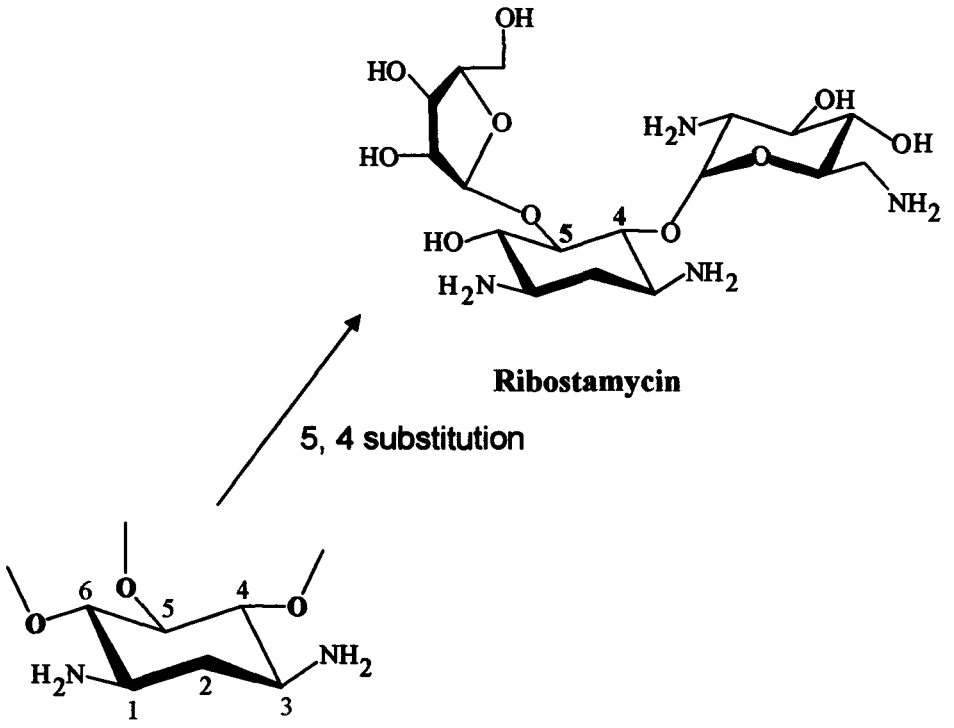


Figure 7. Structure of aminoglycosides. The core structure is formed by the 2-deoxystreptamine ring. The ribostamycin class are substituted by amino sugars at the 5, 4 position and the kanamycin class at the 4, 6 positions.

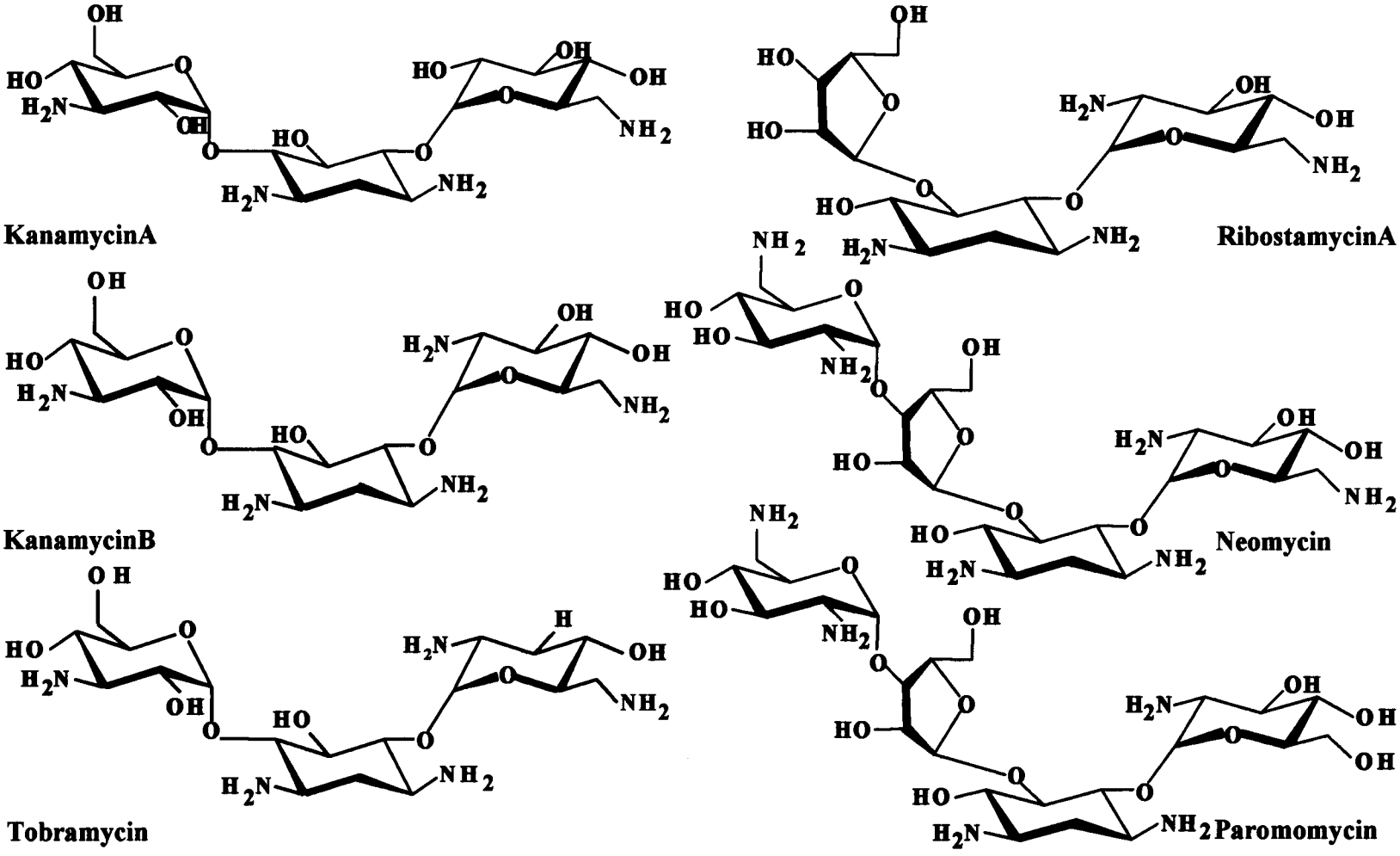


Figure 8. Structure of aminoglycosides used in the study. KanamycinA, B and Tobramycin are substituted by amino sugars at the 4, 6 positions whereas Ribostamycin, Neomycin and Paromomycin are substituted at the 5, 4 positions. Amine groups that are ligation sites for sortase are highlighted in green.

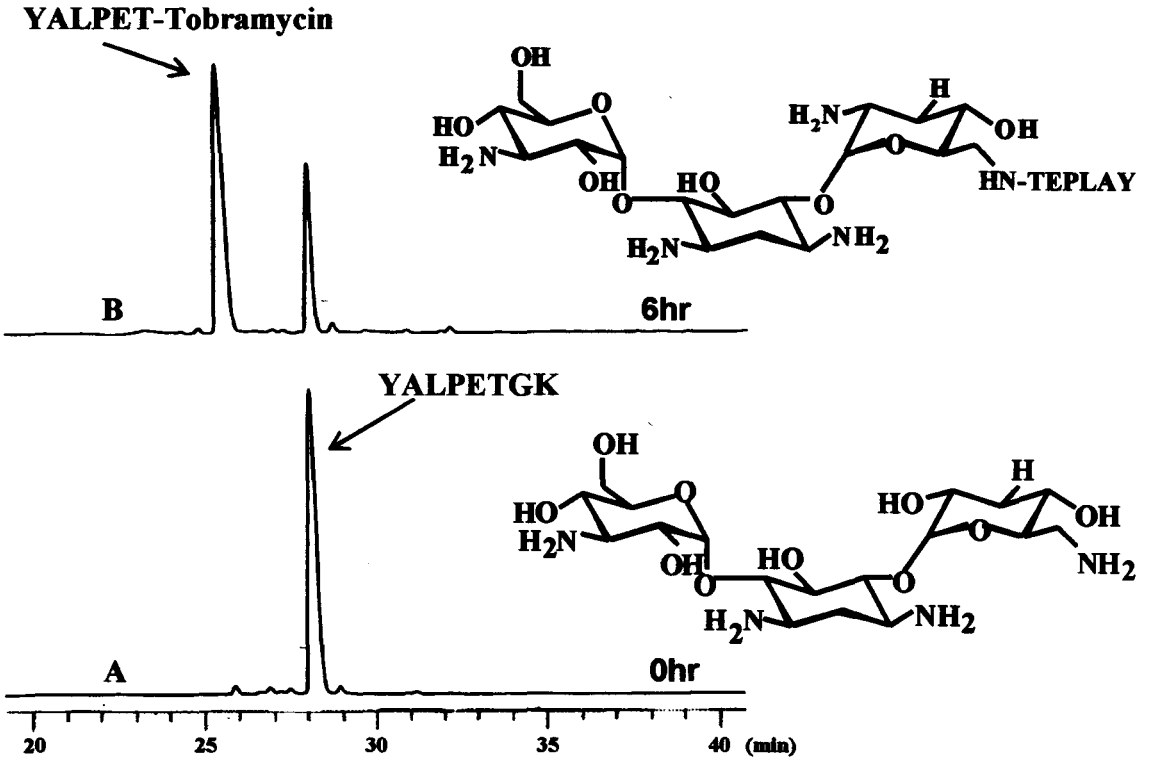


Figure 9. Representative RP-HPLC profile of sortase catalyzed ligation of YALPETGK with Tobramycin. Panel **A** and **B** are the 0h and 6h profiles respectively. The parent peptide and the peptide-tobramycin conjugate are marked by arrows. The HPLC runs were monitored at 210nm.

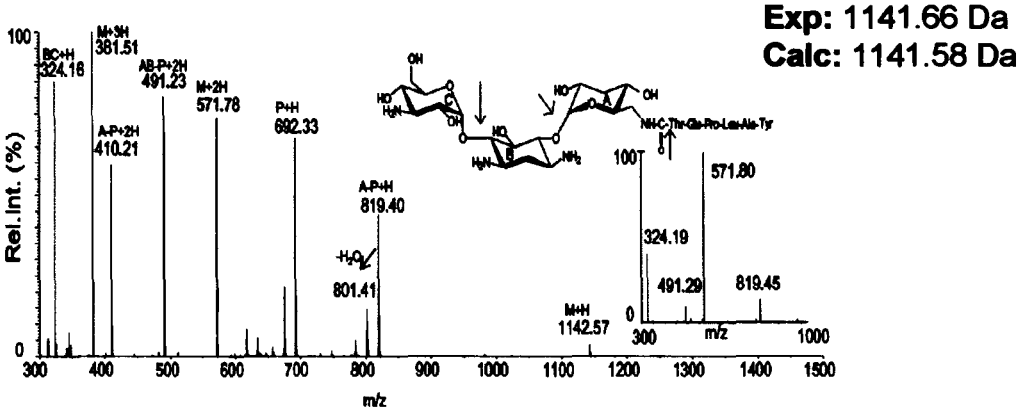


Figure 10. ESI-TOF mass characterization of YALPET-Tobramycin. The three rings of the antibiotic structure are labeled A, B and C and the peptide is labeled P. The ions carry the labels based on the part of the structure from which they are derived. For example, m/z 819.40 is labeled as A-P represents a structure in which ring A is attached to the peptide. Inset shows the MS/MS spectra of m/z 571.8

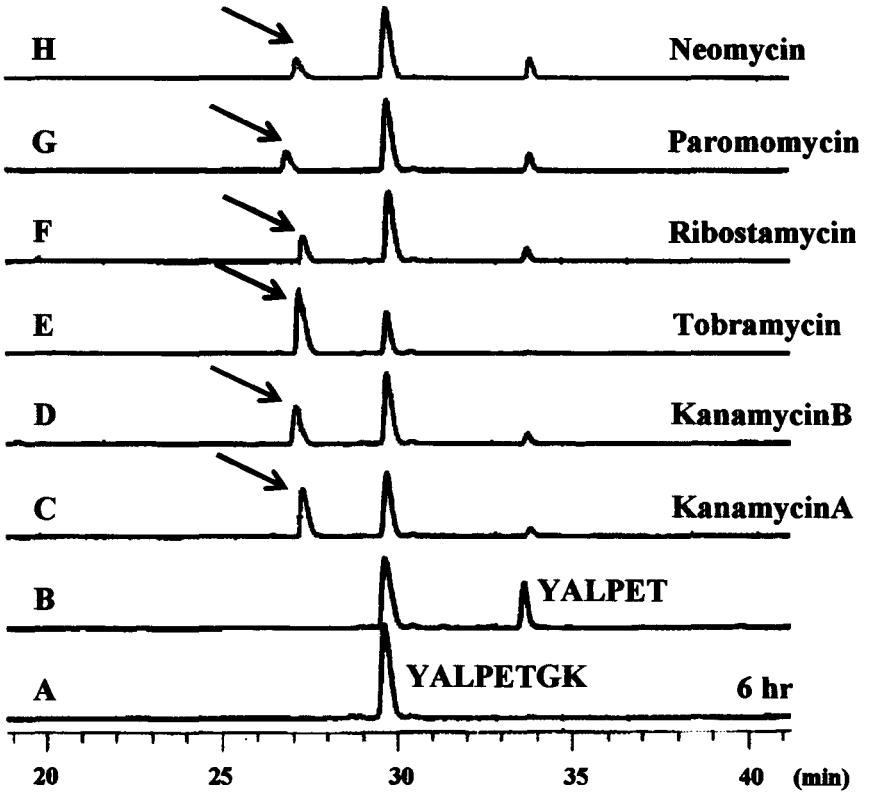


Figure 11. Sortase catalyzed conjugation of YALPETGK model peptide to aminoglycosides. RP-HPLC profiles **A** and **B** are peptide samples in absence and presence of sortase, respectively. Chromatograms **C-H** represent reactions of peptide with aminoglycoside antibiotics. The peptide-antibiotic conjugates are indicated by arrows. The HPLC runs were monitored at 210nm.

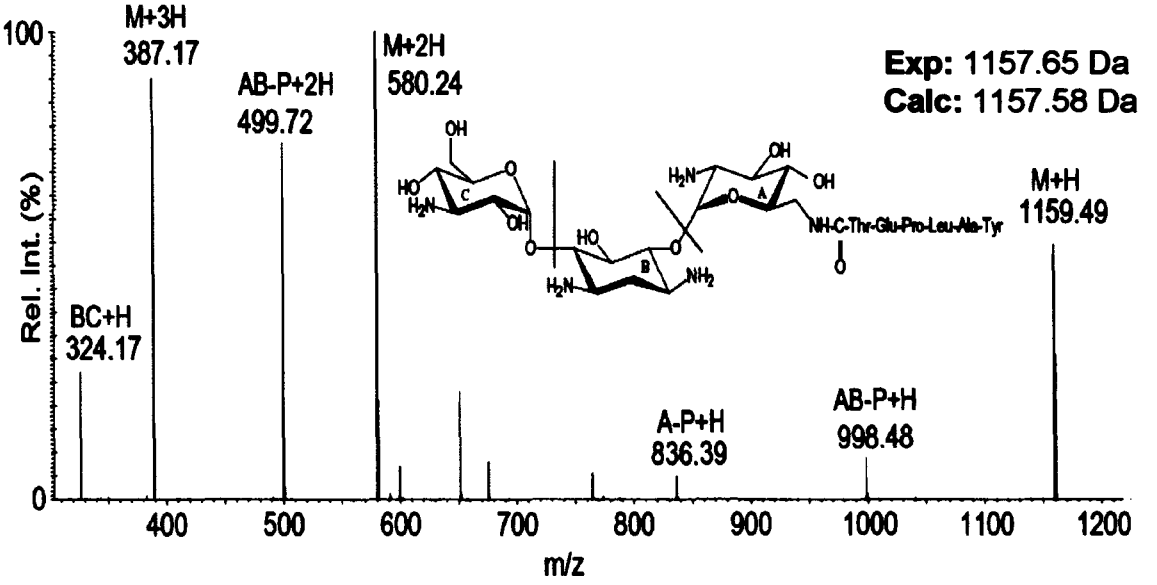


Figure 12. ESI-TOF mass characterization of YALPET-KanamycinB. The ion identification codes are as described in the legend to Figure 10.

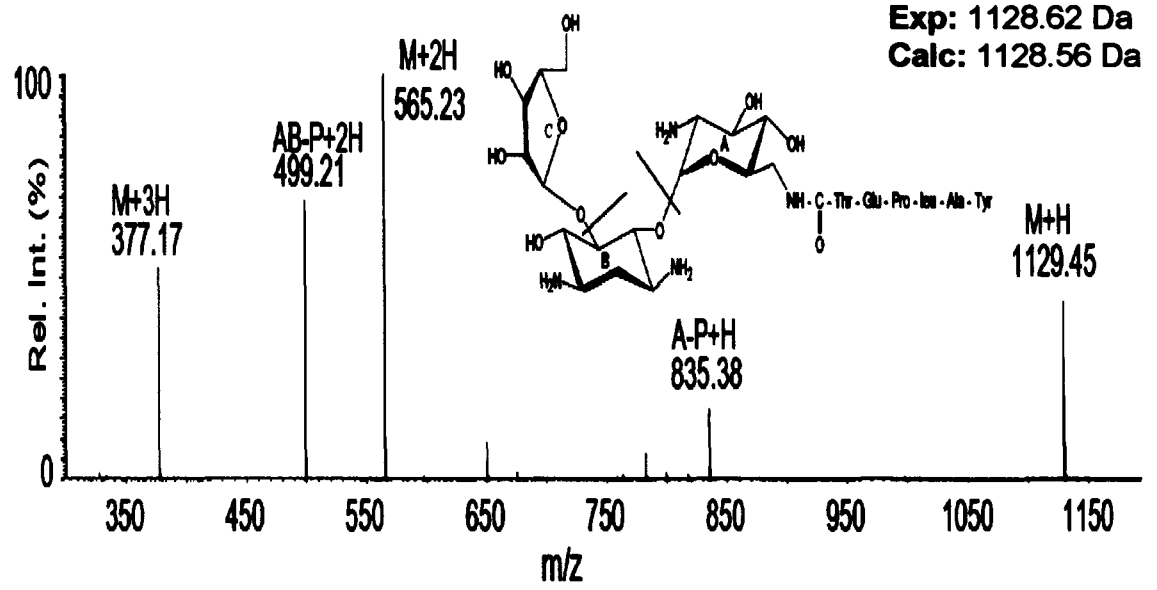


Figure 13. ESI-TOF mass characterization of YALPET-Ribostamycin. The ion identification codes are as described in the legend to Figure 10.

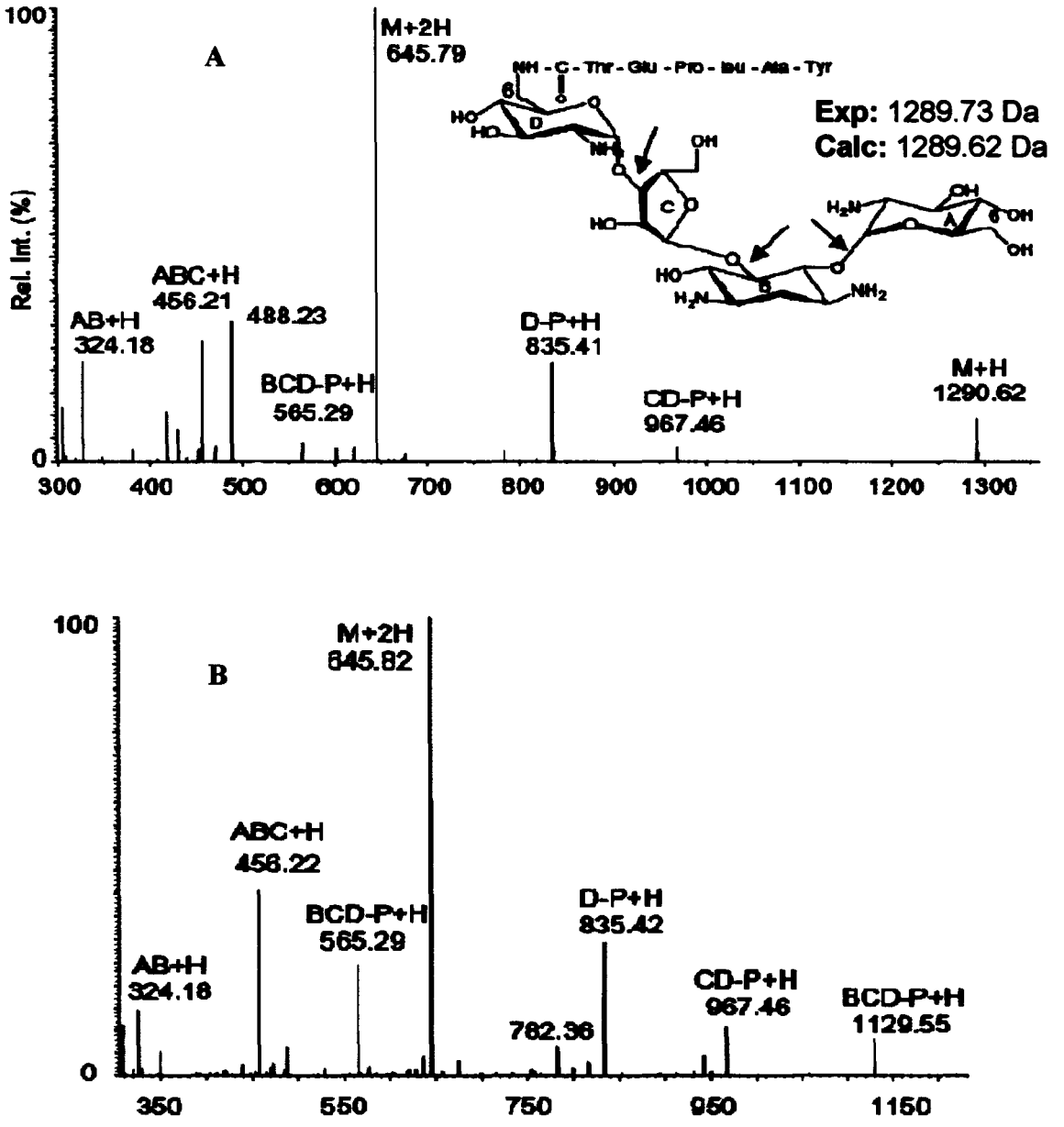


Figure 14. ESI-TOF mass characterization of YALPET-Paromomycin. **(A)** MS of YALPET-Paromomycin conjugate; **(B)** MS/MS spectra of m/z 645.8 of Paromomycin conjugate. The ion identification codes are as described in the legend to Figure 10.

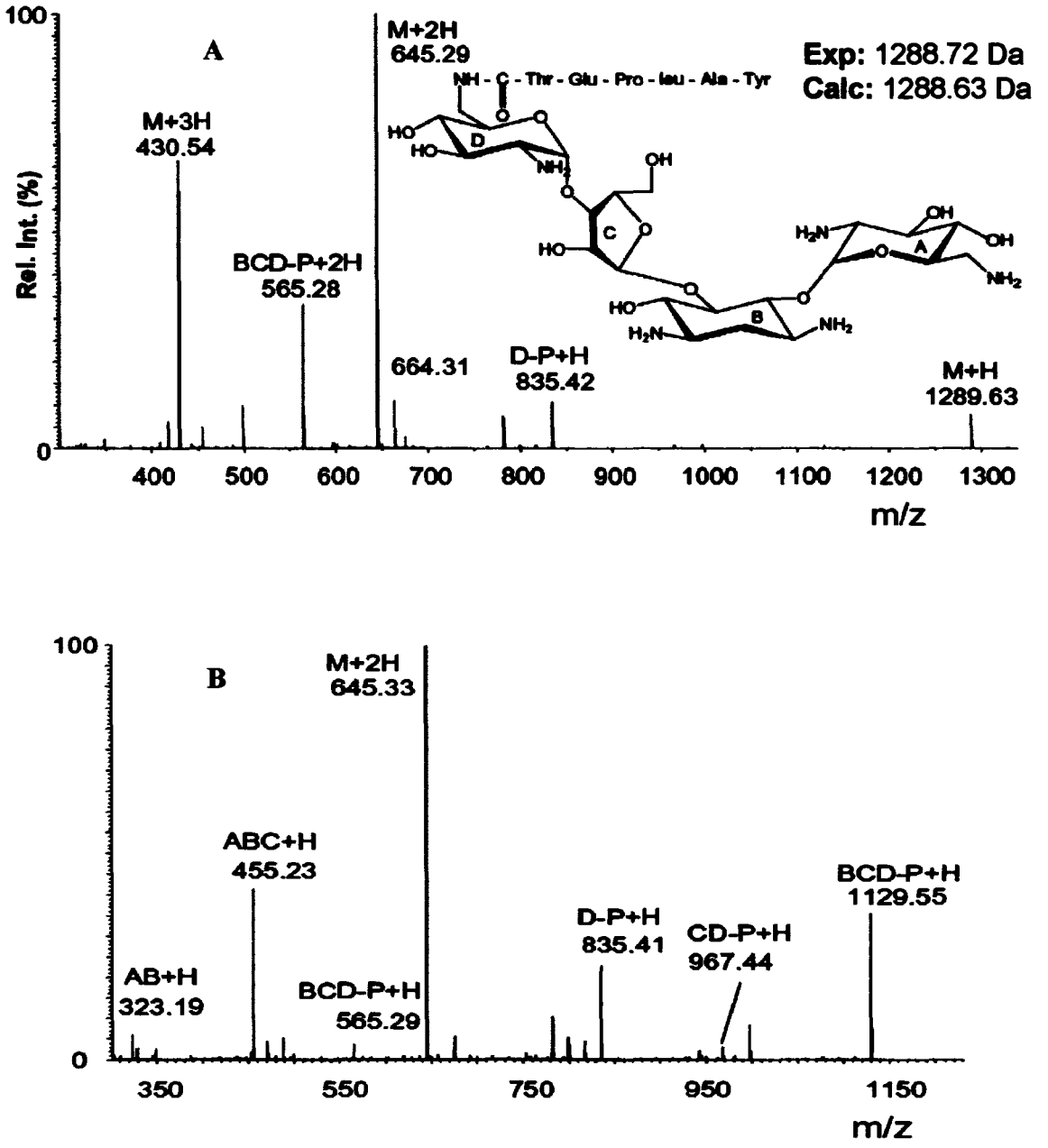


Figure 15. ESI-TOF mass characterization of YALPET-Neomycin. **(A)** MS of YALPET-Neomycin conjugate; **(B)** MS/MS spectra of m/z 645.3 of Neomycin conjugate. The ion identification codes are as described in the legend to Figure 10.

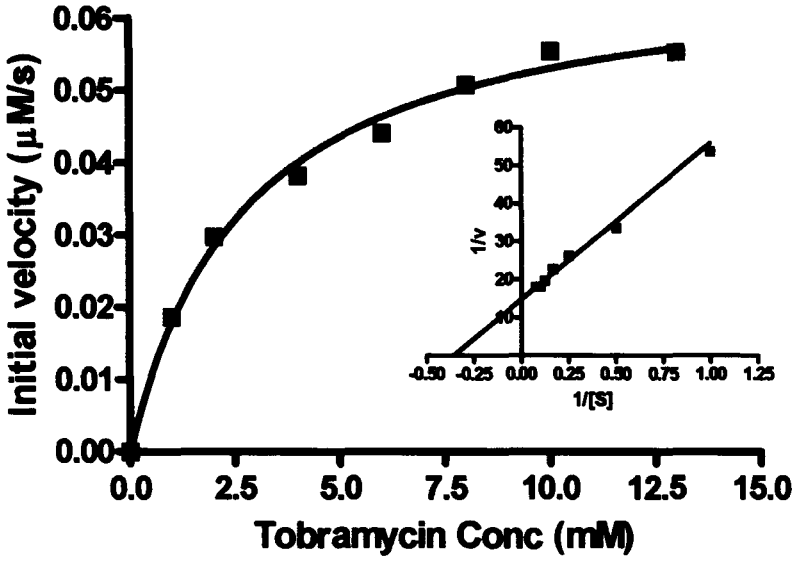


Figure 16. Determination of kinetic parameters for sortase (Tobramycin). Kinetic estimates of K_m=2.794mM and K_{cat}=0.0068s-1 were obtained. Inset shows the Lineweaver-Burke plot.

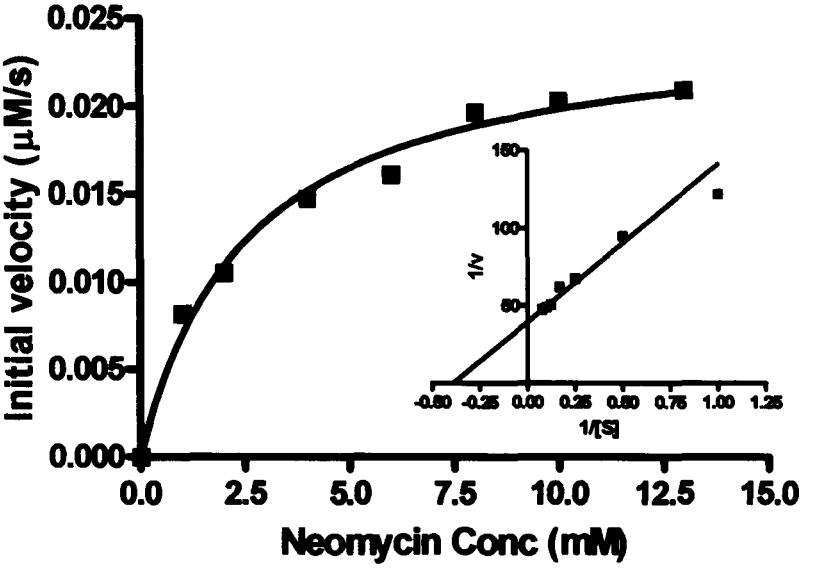


Figure 17. Determination of kinetic parameters for sortase (Neomycin). Kinetic estimates of K_m=2.571mM and K_{cat}=0.0025s-1 were obtained. Inset shows the Lineweaver-Burke plot.

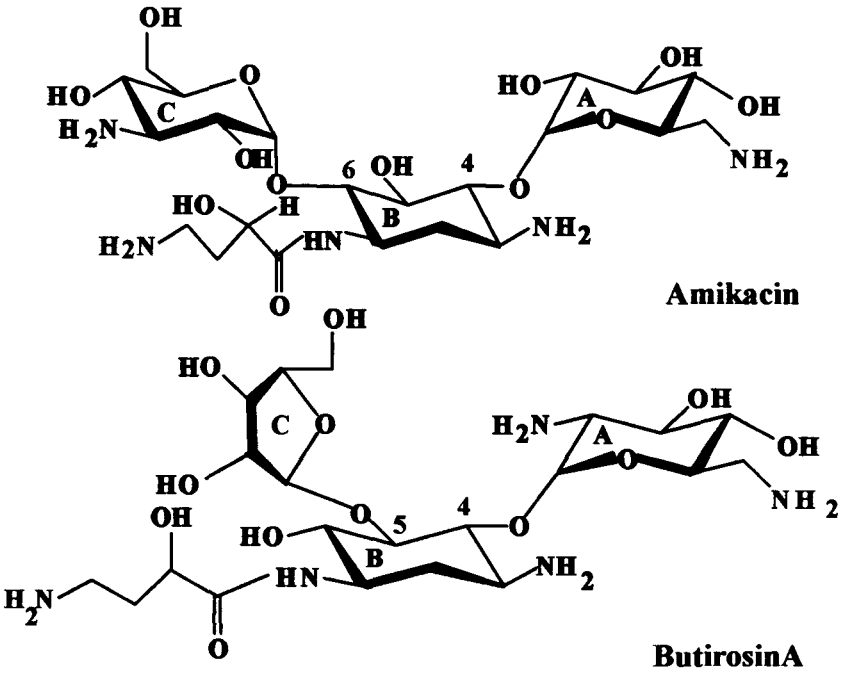


Figure 18. Structure of amikacin and butirosinA

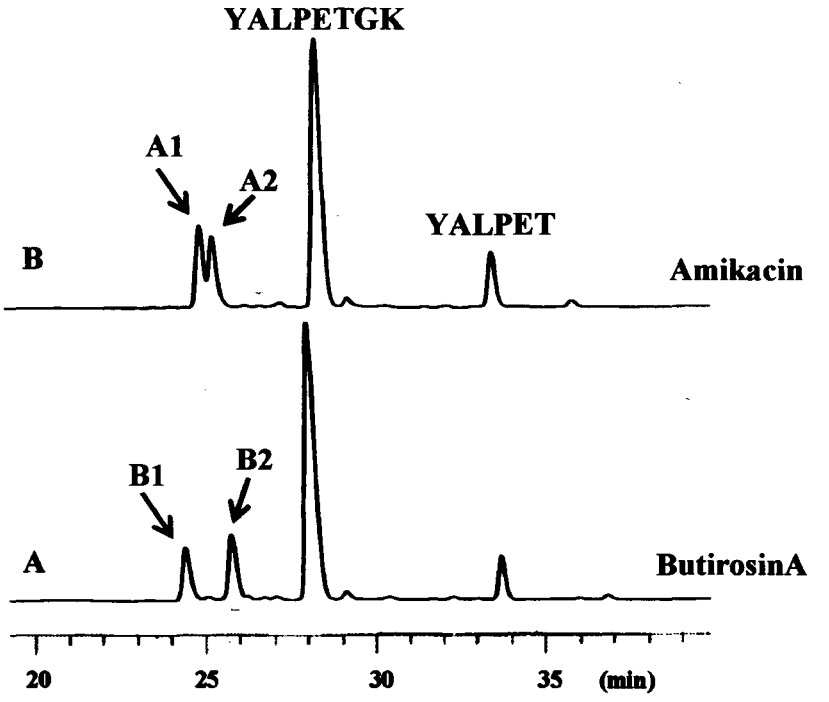


Figure 19. Sortase catalyzed conjugation of YALPETGK with amikacin and butirosinA. Panel **A** and **B** show the RP-HPLC analysis of the 6h reactions with butirosinA and amikacin respectively. The conjugates A1,A2 and B1, B2, are marked by arrows. The HPLC runs were monitored at 210nm.

antibiotics. The results suggested that each product represented a distinct site of peptide ligation. Given that the difference between their respective parent antibiotics (kanamycin and ribostamycin) was the presence of A2HBA group, the two products may represent conjugation of peptide either to the 6-amine in the A-ring or the amine of the A2HBA side chain of the 6-deoxystreptamine ring. Indeed MSMS analysis of product peaks revealed that A1 (in the case of amikacin) and B1 (in the case of butirosinA) originated from peptide ligation at the 6-amino site in the ring A (Figure 20, 21). Likewise, products A2 and B2 were formed as a result of peptide ligation involving amine of the A2HBA group attached to ring B (Figure 20, 21).

Sortase catalyzed ligation of LPETG peptides to aliphatic amines

To ascertain if the hydroxyamino acid moiety present in butirosinA and amikacin could independently act as nucleophiles, the ability of sortase to ligate YALPETGK peptide to 4-amino-2-hydroxy butyric acid (A2HBA) and 4-amino-3-hydroxy butyric acid (A3HBA) was determined. A3HBA used in this study was a recemic mixture whereas A2HBA was composed of the *l* (-) isomer only. HPLC analysis of the reaction mixture followed by ESMS analysis of the respective reaction product confirmed the formation of the expected transpeptidation product (Figure 22). The experimentally determined mass (793.58 Da and 793.21 Da, for A2HBA and A3HBA respectively) was in accord with the calculated mass of 793.82 Da for the peptide adducts (Table 2). The yield of product was in the range of 3-6% (Table 2).

The propensity of several aliphatic amines (6-amino hexanol, 6-amino-hexanoic acid, 7-amino-heptanoic acid, cadaverine, agmatine and spermine) to act as nucleophiles in the sortase catalyzed transpeptidation reaction was also evaluated. HPLC and ESMS assays (Table 2) showed the formation of transpeptidation product in each case *albeit* with varying yields (Figures 23, 24). The product yields, except for cadaverine (45%) and agmatine (27%), were in the range of 7 - 15 % (Table 2). The product yield for spermine could not be determined because an arbitrary amount of spermine was used in the reaction due to its extremely hygroscopic nature.

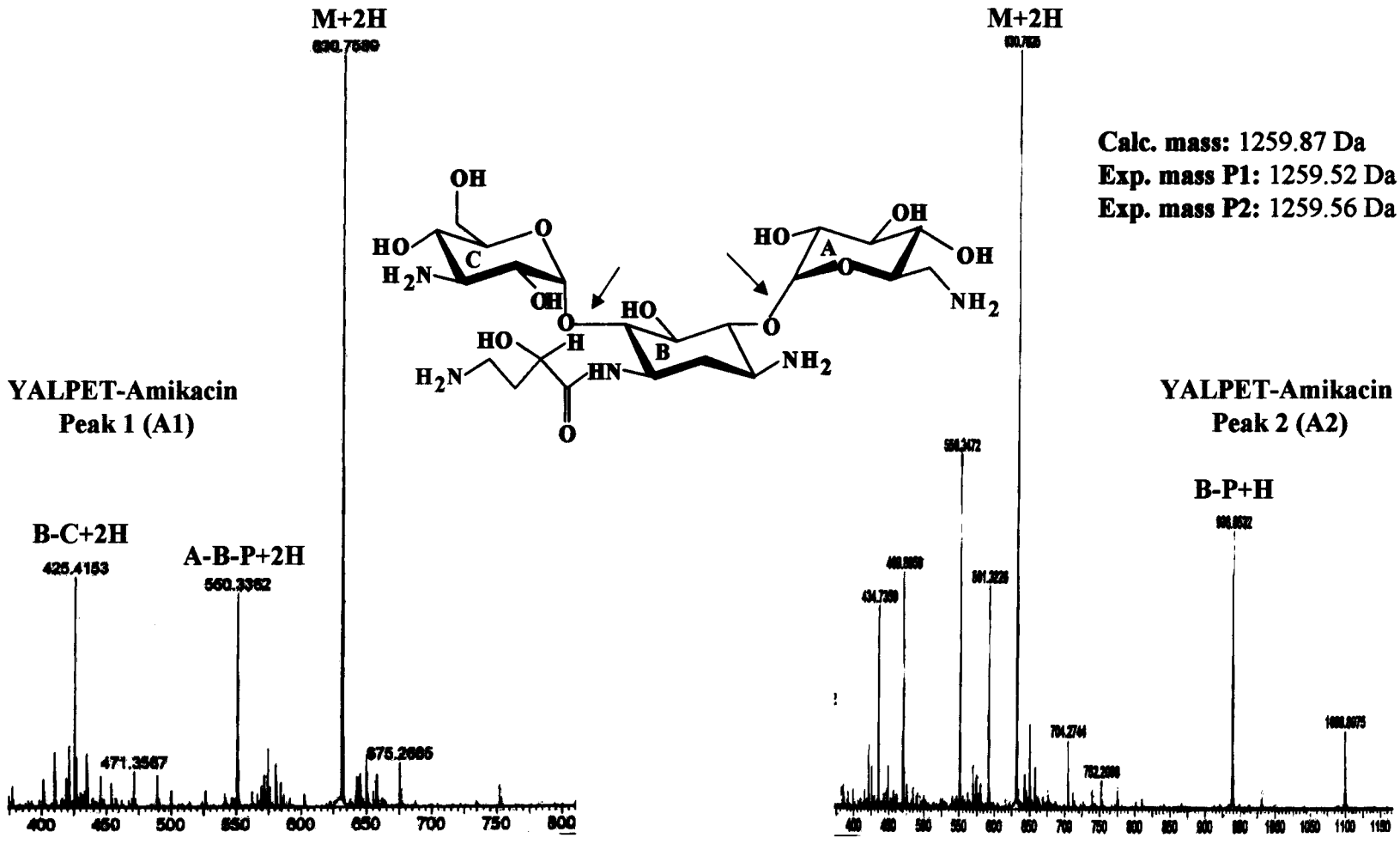


Figure 20. ESI-TOF mass characterization of YALPET-amikacin conjugate peaks 1 (A1) and 2 (A2).The three rings of the antibiotic are labeled A, B and C. The ion identification codes are as described in the legend to Figure 10. Fragment B-P in the mass spectra of A2 confirms the terminal amine of the side chain of ring B as the site of peptide ligation.

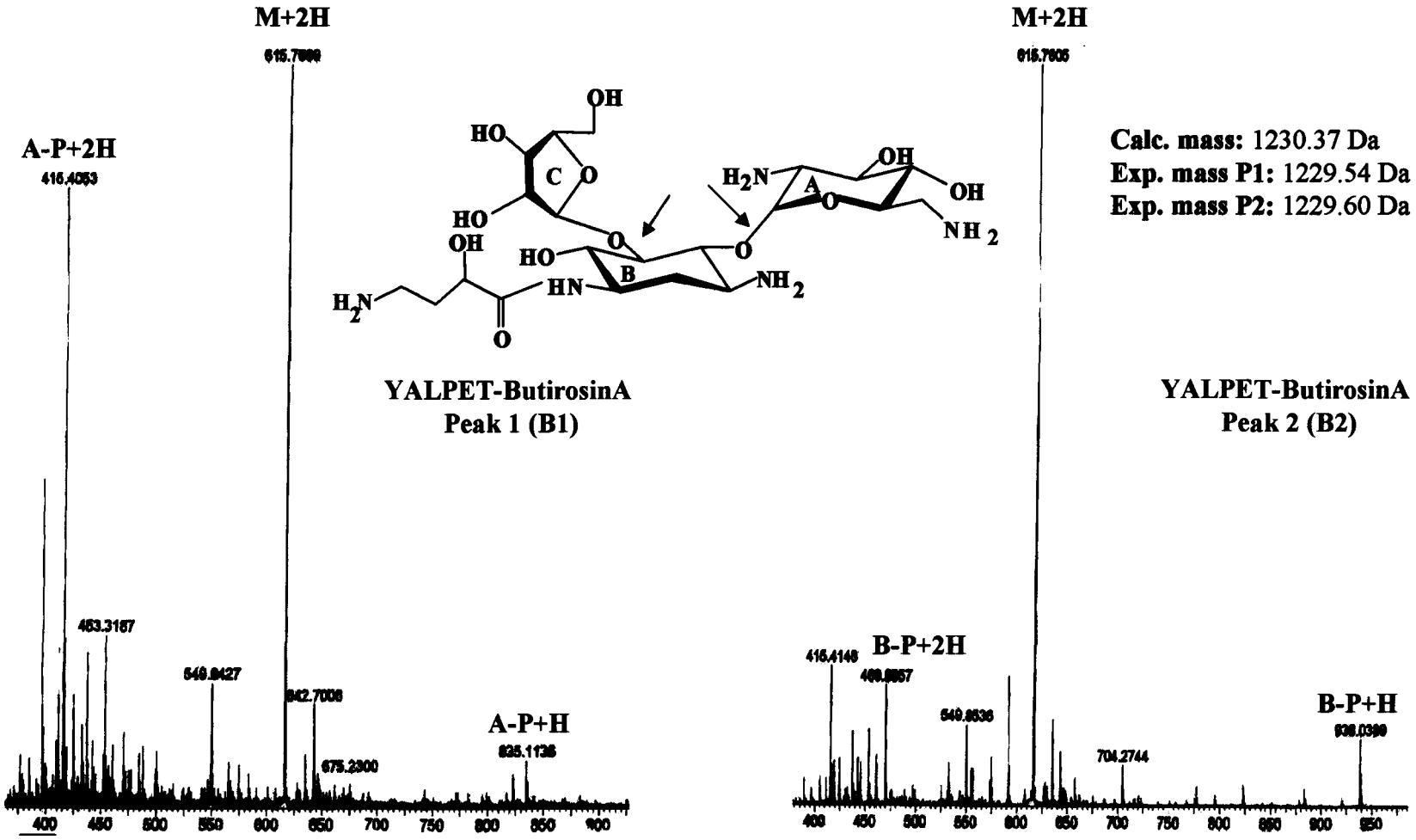


Figure 21. ESI-TOF mass characterization of YALPET-butirosinA conjugate peaks 1 (B1) and 2 (B2).The three rings of the antibiotic are labeled A, B and C. The ion identification codes are as described in the legend to Figure 10. Fragment B-P in the mass spectra of B2 confirms the terminal amine of the side chain of ring B as the site of peptide ligation.

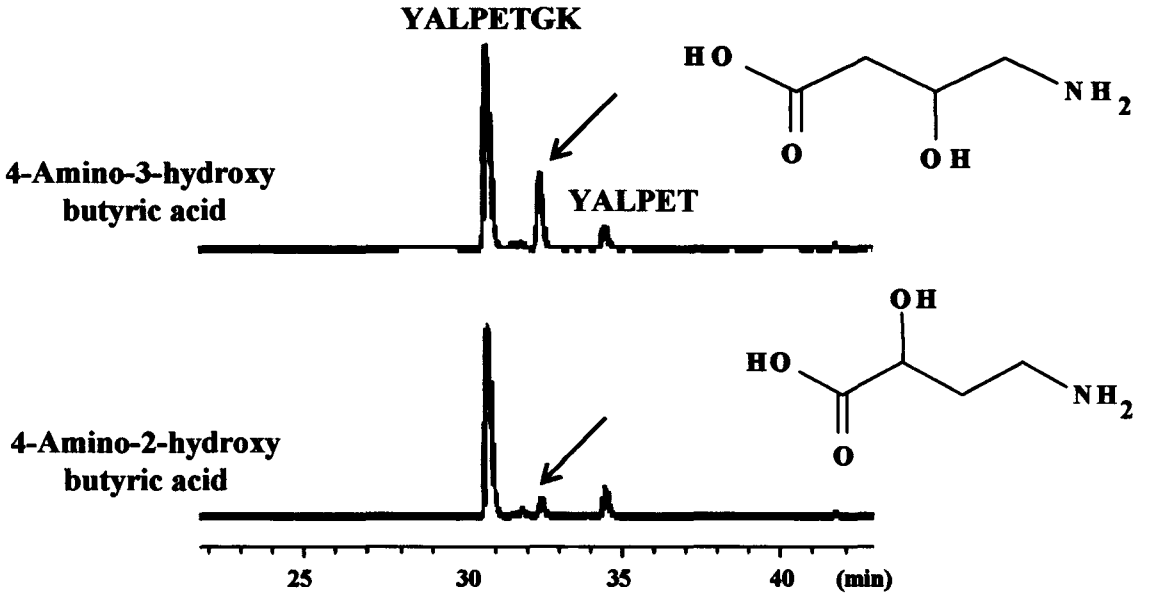


Figure 22. Sortase catalyzed conjugation of YALPETGK with 4A2HBA and 4A3HBA. Panel shows the the RP-HPLC analysis of the 6hr reactions with 4A2HBA and 4A3HBA. The conjugates are marked by arrows. The HPLC runs were monitored at 210nm.

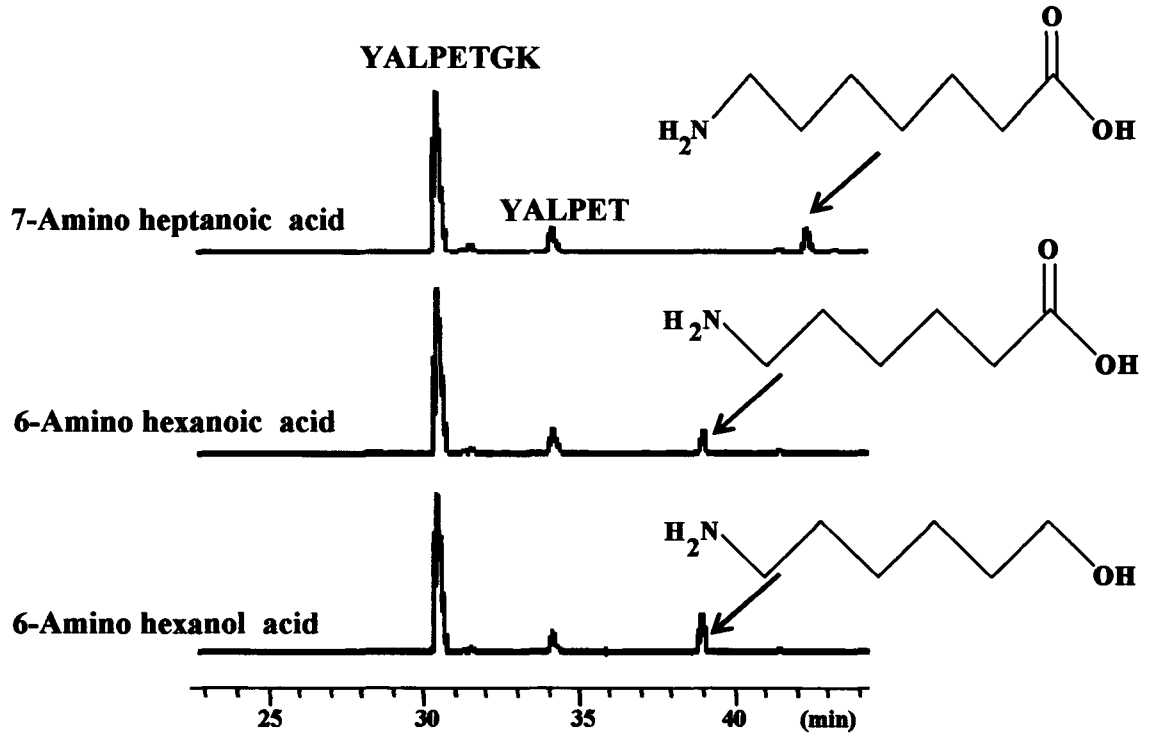


Figure 23. Sortase catalyzed conjugation of YALPETGK with 7-amino heptanoic acid, 6-amino hexanoic acid and 6-amino hexanol. The conjugates are marked by arrows. The HPLC runs were monitored at 210nm.

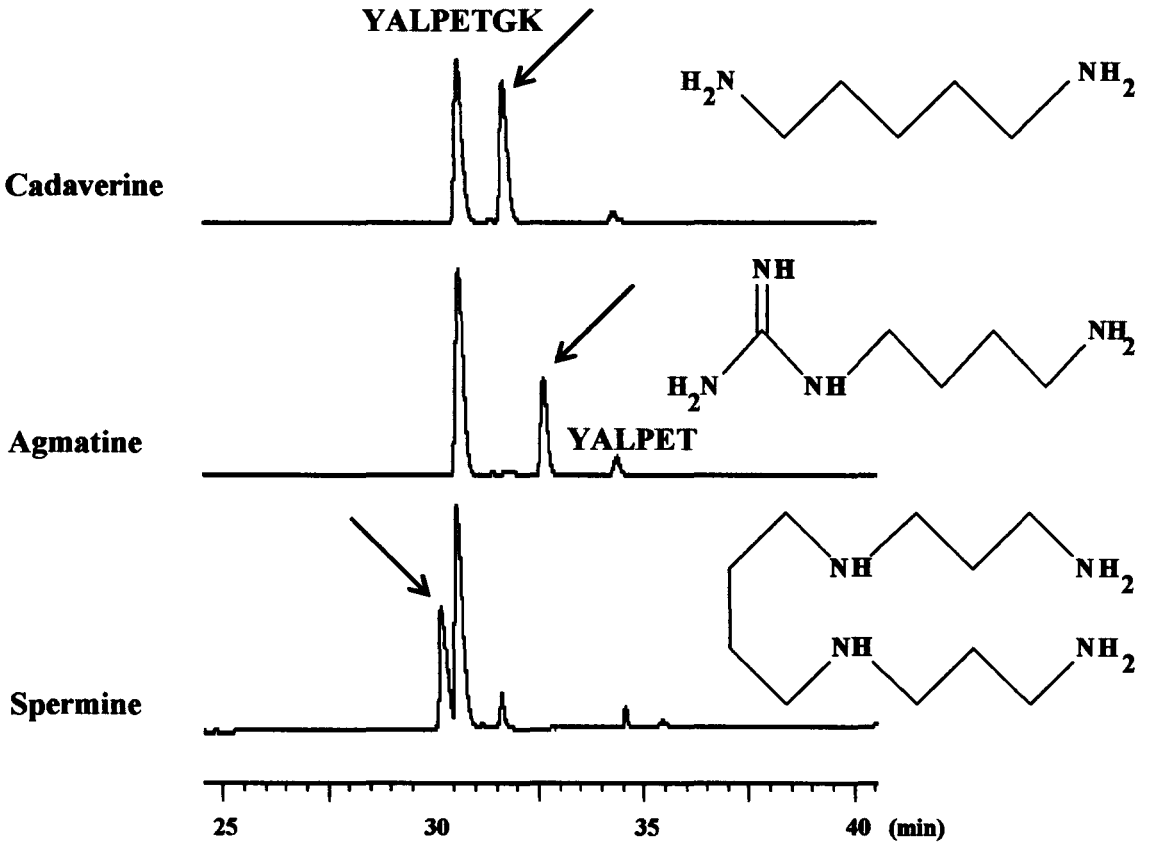


Figure 24. Sortase catalyzed conjugation of YALPETGK with spermine, agmatine and cadaverine. The conjugates are marked by arrows. The HPLC runs were monitored at 210nm.

Table 2. Characterization (mass spectrometry and yield) of sortase catalyzed ligation products formed from LPXTG containing peptides and hydroxylamines, amino alcohols, acids and polyamines.

Peptide	Amino Compound	Experimental (Da)	Calculated (Da)	Yield (%)
YALPETGK	4A2HBA	793.58	793.82	6.30
YALPETGK	4A3HBA	793.21	793.82	3.02
YALPETGK	7-amino heptanoic acid	819.40	819.90	9.28
YALPETGK	6-amino hexanoic acid	805.61	805.87	7.34
YALPETGK	6-amino hexanol	791.44	791.84	14.92
YALPETGK	Cadavarine	776.64	776.38	44.5
YALPETGK	Agmatine	804.62	804.20	26.74
YALPETGK	Spermine	876.50	877.11	*

Reaction conditions: peptide, 0.5mM; aminosugar/antibiotics,2.5mM; sortase,50μM; 37°C; pH 7.5. The reaction was carried out for 6hrs. The product yield is calculated based on HPLC peak areas.

* The yield for spermine could not be calculated (refer Results)

Sortase mediated ligation of aminoglycosides to recombinant proteins

Gene specific primers with engineered sites for restriction endonucleases, NheI in case of the upstream primer and XhoI in case of the downstream primer were used for PCR amplification of *mrp*. The PCR amplified DNA encoded for the complete Mrp protein (Residues 1-202; Figure 25 A, B, C).

The resulting amplicon of 609bp was cloned into the pGEMT-Easy vector. The presence of the insert was confirmed by cleavage with EcoRI, by PCR as well as by DNA sequencing using SP6 promoter primer in an automated sequencer (Figure 26, 27).

The *mrp* insert was sub-cloned into pET23b expression vector, transformed into *E. coli* strain BL21 (DE3) and cells were grown in the presence of IPTG to induce the expression of the 6-His tagged Mrp. The expression and purification of Mrp was monitored by SDS-PAGE analysis (Figure 28). The results presented in Figure 28 (lane 3) demonstrate the induced expression of the protein in the cell lysate. Cells were lysed by sonication and the recombinant protein present in the supernatant was purified to homogeneity from the cell lysate using Ni-NTA agarose resin. Figure 28 (lane 4) shows the purified protein migrating as a single band *albeit* higher than its apparent molecular weight. RP-HPLC analysis showed the recombinant protein eluting as a single peak (Figure 29).The identity of the recombinant protein was established by MALDI-TOF mass analysis (Figure 30A). The experimentally determined mass (23856 ± 20 Da) was in agreement with the calculated value of 23872 Da. The integrity of the primary sequence was also confirmed by Edman degradation. The amino acid sequence was in agreement with that of the NCBI database (Figure 30B).

Sortase mediated ligation of Mrp to aminoglycoside antibiotics

The ability of sortase to ligate recombinant Mrp to aminoglycosides was investigated. Sortase was incubated along with Mrp and tobramycin for a period of 6hrs. When the reaction was analyzed on RP-HPLC, it was discovered that the Mrp_{1-170}-tobramycin conjugate does not resolve from the cleaved Mrp_{1-170} and Mrp (refer Figure 31; Mrp, P2, P3). Nonetheless, MALDI analysis of the unresolved peak (Figure 32; P3) confirmed the formation of the specific Mrp_{1-170}-tobramycin conjugate. In an attempt to resolve the conjugate, the heterogenous peak was re-

A. Nucleotide sequence of *S. aureus mrp* (NC_002758)

gtgcaaaatt atcgaaaagt aagtaatcgt aataaggctg atgcattaaa agctataact
gctttaaaat tacaaatgga tgaagaatta aaaacagcac gcactaatgc tgatgttgat
gcagttttaa aacgatttaa tgttgcatta agcgatatag aagcagtaat tactgaaaaa
gaaaatagct tactgcgaat tgataacatt gctcaacaaa catatgcgaa attcaaagcg
atcgcaacac cagaacaatt agctaaagta aaagtattaa ttgatcaata tgttgcagat
ggcaatagaa tgattgatga agatgcgaca ttaaatgaca tcaaacaaca cacgcaattc
attgttgatg aaattttagc aattaaatta ccagctgaag cgacgaaagt atcaccaaaa
gaaattcagc cagctccaaa agtttgtacg cctattaaaa aagaagagac acatgaatcg
cgcaaagttg aaaaagaact tccaaataca ggttctgaag gaatggattt accattgaaa
gaatttgcac tgattacagg tgcggctttg ttagctagaa gacgtactaa aaacgaaaaa
gaatcataa

B. Protein sequence of *S. aureus* Mrp (NP_372281)

mqnyrkvsnrnkadalkaitalklqmdeelktartnadvdavlkrfnvalsdieavit
ekensllridniaqqtyakfkaiatpeqlakvkvlidqyvadgnrmidedatlndikq
htqfivdeilaiklpaeatkvspkeiqpapkvctpikkeethesrkvekelpntgseg
mdlplkefalitgaallarrrtknekes

C. PCR amplification of *S. aureus mrp*

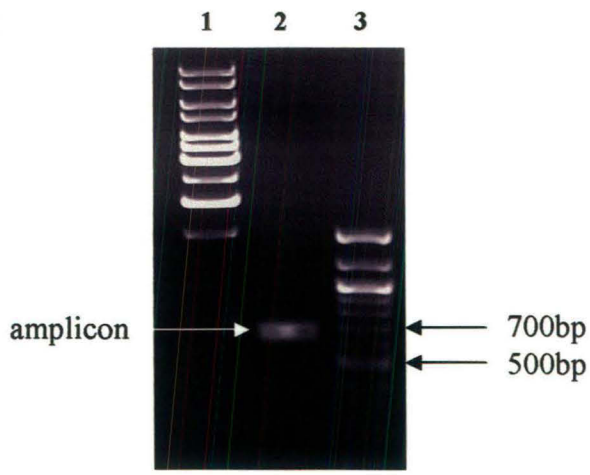


Figure 25. **(A)** The complete nucleotide sequence of *Staphylococcus aureus mrp* as obtained from the NCBI database. **(B)** The complete protein sequence of Mrp; the LPNTG motif is highlighted in red **(C)** PCR amplification of *mrp* from *S. aureus* genomic DNA. Lane 1, 3: 1kb and 100bp DNA ladders, the 500bp and 700bp bands are marked by arrows; lane 2: PCR amplified *mrp*

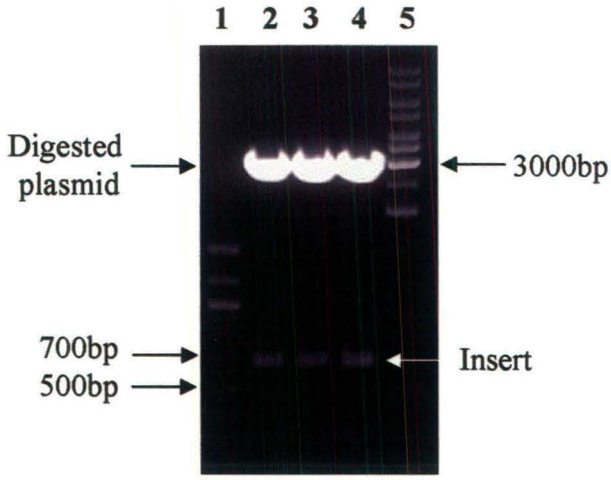


Figure 26. Restriction digest confirmation of pGEMT:*mrp* clones. Plasmids isolated from candidate recombinant clones were digested with EcoRI to check for the presence of insert. Lane 1, 5:100kb DNA ladder, 1kb ladder; lane 2,3,4: EcoRI digest of 3 candidate clones The insert can be seen lower than the DNA ladder band corresponding to 700bp (marked by arrow). The digested plasmid can be seen alongside the 3kb band.

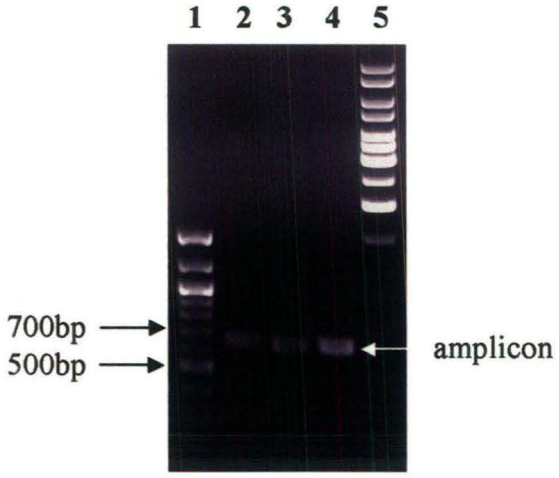


Figure 27. PCR confirmation of pGEMT:*mrp* clones. Plasmids isolated from the candidate recombinant clones were used for PCR amplification using the same primer combination and PCR parameters as was used for the original PCR. Amplicon of desired size confirms presence of insert. Lane 1,5: 100bp DNA ladder, 1kb ladder; lane 2,3,4: PCR amplification of 3 candidate clones. The amplicon can be seen just below the 700bp marker (marked by arrow).

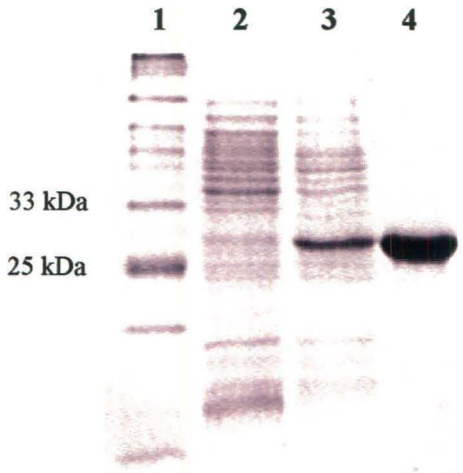


Figure 28. Expression and purification of *S. aureus* Mrp. The samples were subjected to 15% SDS-PAGE followed by Commassie brilliant blue staining. Lane 1:protein marker; lane 2:uninduced cell extract; lane 3:IPTG induced cell extract; lane 4:purified Mrp protein. Note that the recombinant protein runs higher than its molecular weight.

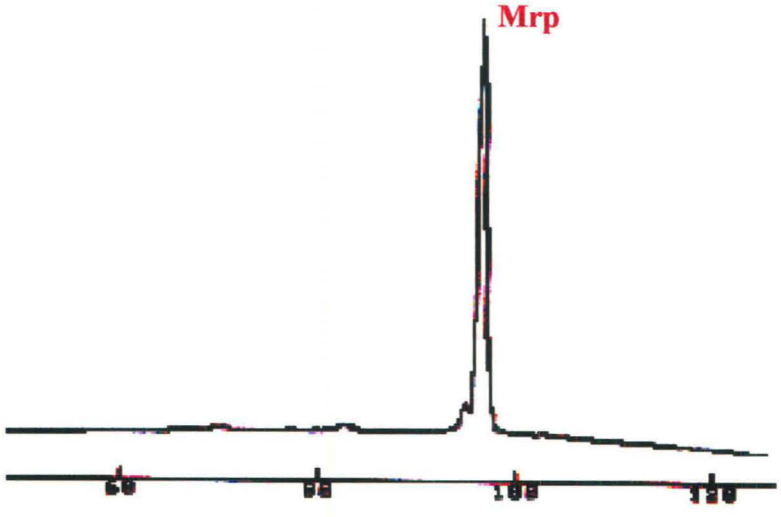
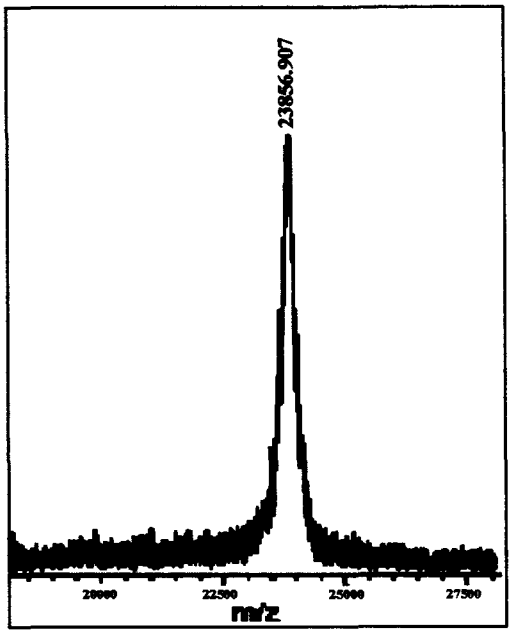


Figure 29. RP-HPLC analysis of recombinant Mrp.

A



B

AAcid #	AAcid ID	R.Time (min)	C.Time (min)	Pmol (raw)	Pmol (-back)	Pmol (+lag)	AAcid ID
1	A	7.70	7.70	580.95	549.72	549.72	Ala
2	S	4.95	4.90	98.17	86.49	86.49	Ser
3	V	12.74	12.68	499.88	490.52	501.55	Val
4	Q	5.22	5.17	192.10	187.72	196.81	Gln
5	N	4.36	4.31	162.58	155.47	162.99	Asn
6	Y	9.63	9.60	519.22	517.62	517.62	Tyr
7	R	9.47	9.43	817.68	816.83	816.83	Arg
8	K	15.95	15.93	573.69	522.49	539.43	Lys
9	V	12.69	12.68	464.19	433.53	453.50	Val
10	S	4.93	4.90	66.47	59.49	59.49	Ser
11	N	4.38	4.31	132.56	128.05	128.05	Asn
12	R	9.23	9.43	733.51	710.22	744.59	Arg
13	N	4.19	4.31	132.04	128.40	128.40	Asn
14	K	15.74	15.93	510.25	416.02	426.95	Lys
15	A	7.52	7.70	420.90	409.36	409.36	Ala
16	D	3.70	3.80	107.99	105.89	105.89	Asp
17	A	7.53	7.70	429.05	420.33	420.33	Ala
18	L	15.99	16.21	413.40	329.73	345.65	Leu
19	K	15.73	15.93	464.86	334.77	336.23	Lys
20	A	7.50	7.70	416.50	412.00	412.00	Ala

Figure 30. (A) MALDI-TOF analysis of recombinant Mrp. The experimentally determined value of 23856 ± 20 Da was in accordance with the theoretically calculated value of 23872 Da. **(B)** N-terminal protein sequencing of Mrp. 20 residues were sequenced and were identical to that of the NCBI database. Note the absence of the starting Met residue. The N-terminus Met may have been removed by *E.coli* aminopeptidase M.

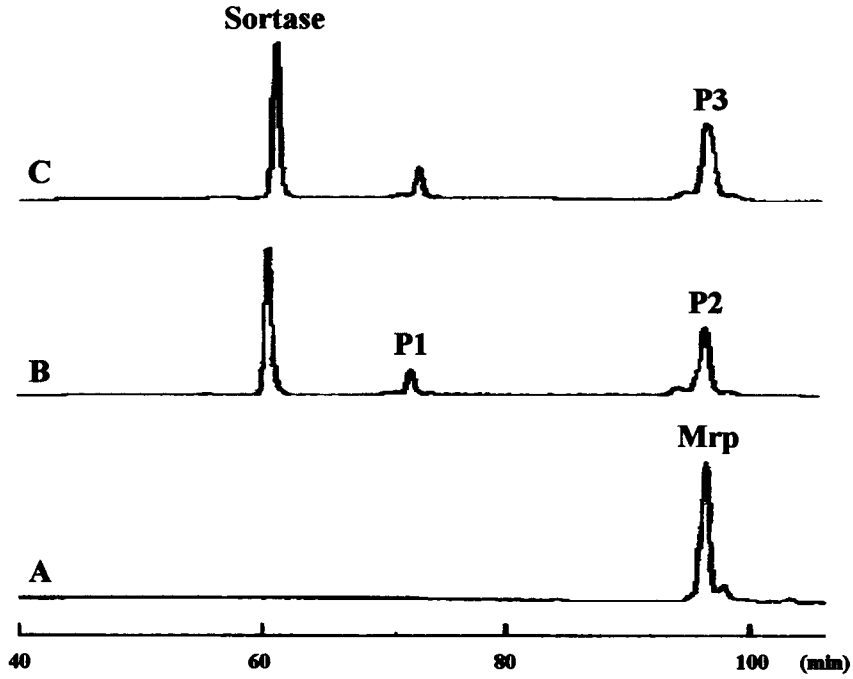


Figure 31. RP-HPLC of sortase catalyzed conjugation of Mrp to tobramycin. Profiles **A** and **B** are Mrp samples in absence and presence of sortase respectively. Chromatogram **C** represents reaction of Mrp with tobramycin. A distinct peak corresponding to the Mrp_{1-170}-tobramycin conjugate was not observed hence, identity of HPLC peaks labeled P1, P2 and P3 was investigated by ESMS analysis. All HPLC runs were monitored at 210nm.

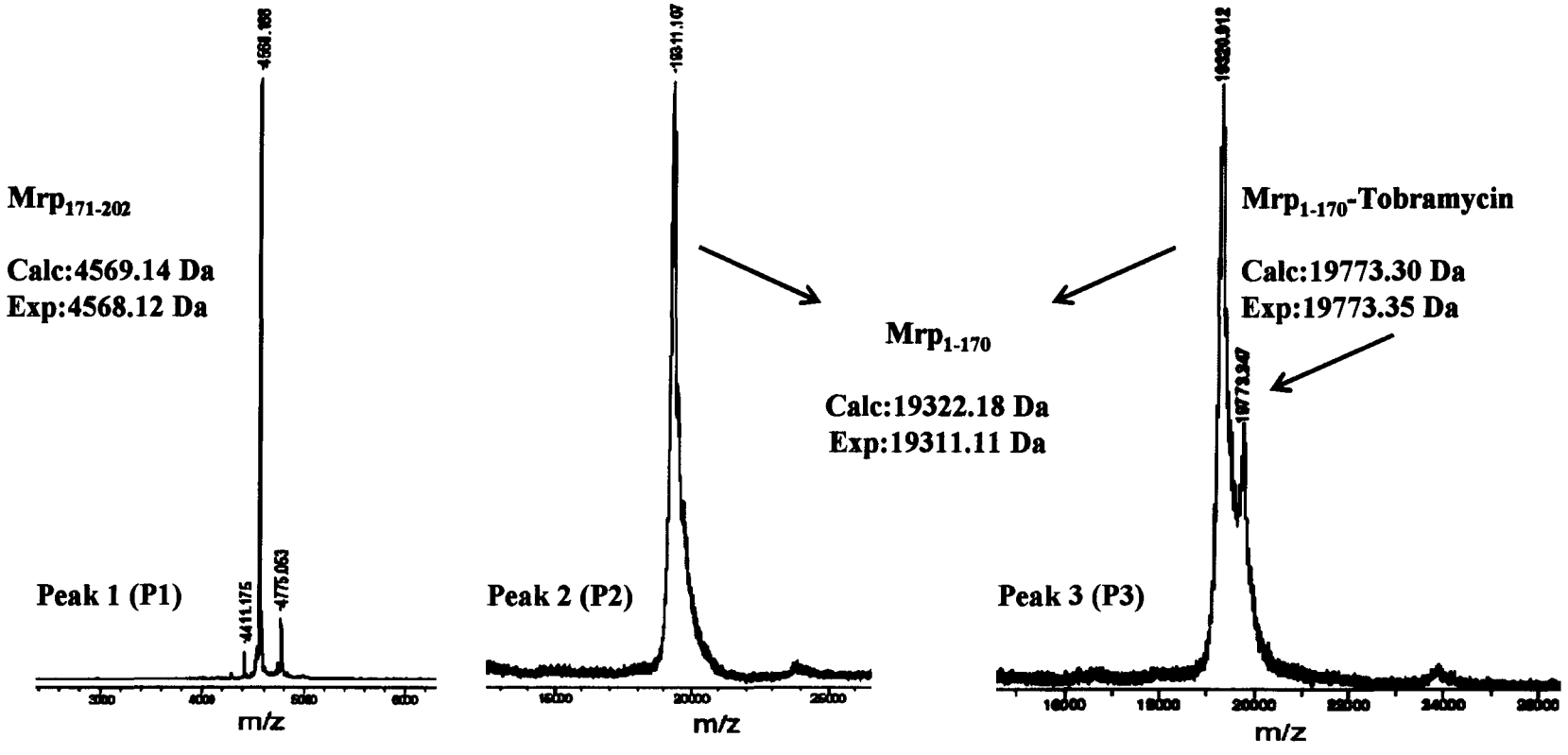


Figure 32. MALDI-TOF characterization of Mrp_{1-170}-tobramycin peaks 1 (P1), 2 (P2) and 3 (P3). Mass of P1 and P2 corresponded to that of cleaved Mrp (residues 171-202) and MrpLPNT (residues 1-170) respectively. ESMS analysis of P3 revealed the presence of the Mrp_{1-170}-tobramycin conjugate.

analyzed on 15% SDS-PAGE where distinct bands corresponding to Mrp (Mrp_{1-202}), cleaved Mrp (Mrp_{1-170}) and Mrp_{1-170}-tobramycin conjugate could be observed (Figure 33, lane 4). Hence, sortase mediated ligation of Mrp to other aminoglycosides (neomycin, kanamycinA, ribostamycin and paromomycin) were analyzed similarly i.e., RP-HPLC followed by analysis of the heterogenous peak by SDS-PAGE (Figure 34). The yield of the conjugates was estimated to be in the range of 25-50%.

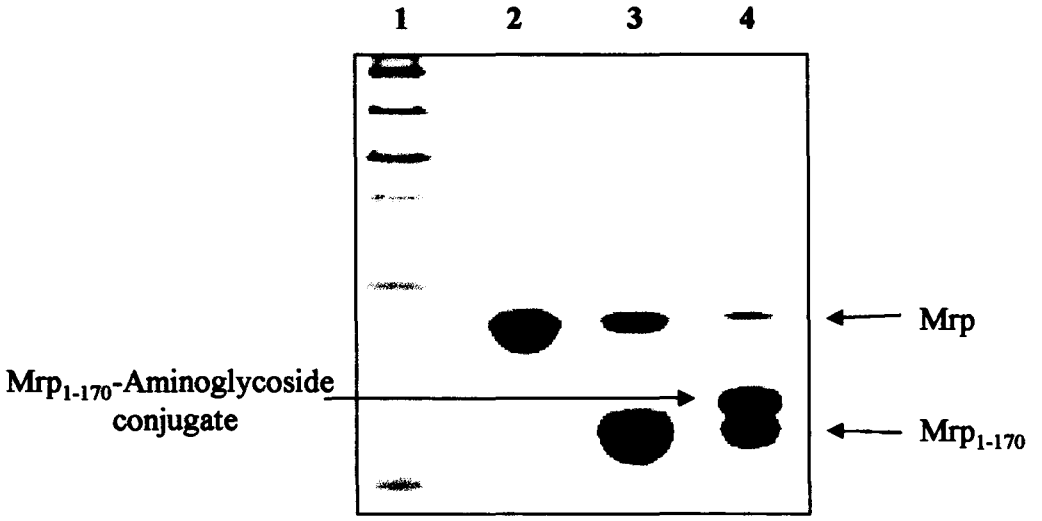


Figure 33. RP-HPLC of sortase catalyzed conjugation of Mrp to tobramycin. The heterogenous peaks, P2 and P3 (refer Figure 31), were analyzed by SDS-PAGE. Lane 1: protein marker; lane 2: Mrp only; lane 3: peak P2 (Mrp + sortase); lane 4 (Mrp + tobramycin + sortase). Distinct bands corresponding to Mrp, Mrp_{1-170} and Mrp_{1-170}-tobramycin can be observed (marked by arrows).

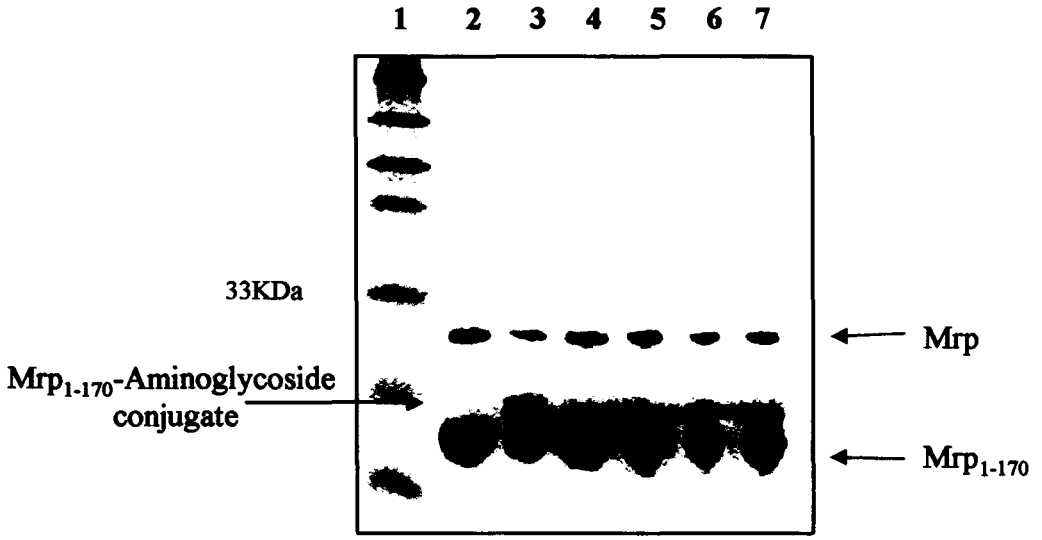


Figure 34. SDS-PAGE analysis of sortase catalyzed conjugation of Mrp to various aminoglycosides. Lane 1 is the protein marker; lane 2 is the reaction of Mrp only with Sortase, lane 3, 4, 5, 6, 7 are sortase catalyzed Mrp conjugation reactions with tobramycin, neomycin, kanamycinA, ribostamycin and paromomycin respectively. In each reaction a distinct band corresponding to the Mrp_{1-170}-aminoglycoside conjugate can be observed.

Discussion

The *in-vitro* transpeptidation reaction catalyzed by sortase entails cleavage of the Thr-Gly peptide bond of an LPXTG containing protein and its subsequent transfer to the terminal amino group of a pentaglycine moiety. It has been reported that pentaglycine *per se* is not essential for sortase activity as even a single glycine molecule can function as a nucleophile *albeit* with about 10-fold less efficiency[156]. This observation hinted at relaxed substrate specificity for sortase and provided the motivation for exploring aminosugars as amine donors in transpeptidation reaction. Accordingly, aminohexoses such as 6-deoxy-6-aminoglucose/mannose containing an amine at the C-6 were chosen with the intuition that $-CH_2-NH_2$ group may bear some structural resemblance to the glycine molecule. Indeed, transpeptidation reaction of a model YALPETGK substrate to the aforementioned aminohexoses proceeded smoothly with yields of about 30-40%. By contrast, the substrate peptide was hydrolyzed to YALPET without the formation of a peptide-sugar adduct when glucosamine was used a substrate suggesting that the $-NH_2$ group present in glucosamine could not act as a nucleophile in the sortase catalyzed transpeptidation reaction. These results suggest that the selectivity of this reaction is sufficiently broad to accommodate 6-deoxy-6-aminohexoses as a replacement for glycine, but not so broad as to allow attachment of peptides to the more common 2-deoxy-2-aminohexoses.

To test if sortase could be used to 'tag' larger molecules appended with an aminohexose moiety, sortase catalyzed reactions with aminoglycoside antibiotics were carried out. Aminoglycosides are a class of therapeutically important antibiotics that are widely used to treat bacterial infections. They are built up by a variety of amino sugars of the 6-amino or the 2, 6-diamino type besides containing several other amino functionalities[164, 165]. The purpose of using aminoglycosides as sortase substrates was two-fold: (i) to determine if sortase mediated transpeptidation involving 6-aminohexoses and LPXTG substrates could be extrapolated to larger, more complex structures; (ii) as aminoglycosides contain a large number of $-NH_2$ groups, specificity of the reaction towards the $-CH_2-NH_2$ group could be determined. Experimental evidence demonstrated that in spite of a large number of other amines being present, sortase transferred peptide substrates, in a selective

manner, to 6-aminohexose moiety. Nonetheless, the amine in the A2HBA side-chain of amikacin and butirosin was also amidated in addition to the 6-amino site. Curiously, only mono-substituted adducts were obtained. The reason for the absence of di-substituted product is not readily apparent but may be due to steric hindrance as the mono-substituted aminoglycoside molecule may be too large to fit back into the enzyme active site. The total yield for formation of peptide-amikacin/butirosinA adducts were comparable at ~40% each. The yield for the formation of the peptide adducts (for amikacin) at the $-CH_2-NH_2$ moiety was marginally higher as compared to the A2HBA group (24% and 14% respectively); for butirosinA the yields were identical at 18% each. The A2HBA or A3HBA were independently capable of acting as nucleophiles but were not as effective as when present in the context of aminoglycosides. The yield of the reaction for A2HBA (6.3%) was much lower than that obtained with amikacin and butirosinA. Thus the results show that although A2HBA is capable of acting as a nucleophile by itself, the reaction is more robust when A2HBA is part of amikacin or butirosinA. This is probably because aminoglycosides provide better proximity and orientation of the amine nucleophile for catalytic action.

Transpeptidation reaction also occurred when amine containing compounds such as 6-amino-hexanol, 6-amino-hexanoic acid, and 7-amino-heptanoic acid were used as nucleophiles but the yields were relatively poor. However, polyamines like cadaverine, agmatine and spermine produced yields that were comparable to that obtained with aminosugars (44% and 27% for cadaverine and agmatine respectively). The reason as to why polyamines are good sortase substrates is not clear.

To investigate if the reaction could be extrapolated to full length recombinant proteins as well, the *Staphylococcus aureus* multidrug resistance protein (Mrp) was cloned and expressed. Mrp, a 202 residue protein, is a natural sortase substrate. It has a typical C-terminal CWS signal containing an LPNTG motif. When the reaction mixture of Mrp, tobramycin and sortase was analyzed by RP-HPLC it was observed that the Mrp_{1-170}-tobramycin conjugate did not separate from the cleaved Mrp_{1-170} and the Mrp. ESMS analysis, however, confirmed the formation of the desired conjugate. However, the heterogenous peak obtained from the RP-HPLC run could be resolved by SDS-PAGE electrophoresis. On the SDS-PAGE the three

components migrated as distinct bands and the Mrp_{1-170}-tobramycin conjugate could be clearly observed. The same reaction was performed with other aminoglycosides like neomycin, kanamycinA, ribostamycin and paromomycin; in each case the formation of the Mrp_{1-170}-aminoglycoside conjugate could be observed. The yield of the conjugates varied between 25-50%.

The results so far demonstrate that the staphylococcal transpeptidase sortase has more relaxed substrate specificity than was assumed till now. It is capable of catalyzing transpeptidation reactions with a plethora of compounds containing the $-CH_2-NH_2$ group, including aminosugars (like 6-amino hexoses and aminoglycosides) and aliphatic amines (like 6-amino hexanol, 6-amino heptanoic acid etc.). Such an enzymatic activity for sortase was hitherto unknown. It provides a simple and straightforward method for covalent ligation of a prefabricated sugar containing a 6-aminohexose tag to synthetic peptides and expressed proteins encoded with a C-terminal LPXTG sortase recognition sequence. As aminoglycosides are known for their antibacterial as well as RNA-binding capabilities, this provides a platform for the synthesis of aminoglycoside-peptide conjugates with potential bioactivity.

Chapter III

Sortase Mediated Construction of Bioconjugates

Introduction

Development of new methods for linking sugars to peptides or proteins is an active area of research because natural glycopeptides or glycoconjugates play important roles in biology and medicine and are indispensable tools for probing several biological processes[166-169]. However, despite dramatic progress in synthetic carbohydrate and peptide chemistry in recent years, glycoconjugate synthesis involving sugar and polypeptide remains a formidable task. This is principally because synthetic protocols are quite demanding and involve multiple reaction steps with requirements of extensive protection of reactive functionalities. This problem may be in part or completely obviated through the intermediary of enzymes. The ability of sortase to ligate LPXTG containing peptides and proteins to aminosugars can be a useful tool in the construction of neoglycoconjugates.

The results described in chapter 2 demonstrated the ability of staphylococcal sortase for site specific ligation of model LPXTG peptide substrates to therapeutically important aminoglycoside antibiotics. In this chapter, sortase-mediated facile ligation of peptides to antibiotics has been applied to the construction of bioconjugates comprising aminoglycosides and peptide sequences derived from Rev/Tat proteins of HIV-1. The RNA binding ability of the conjugates has been evaluated.

Materials and methods

Materials

All Fmoc amino acids and resins were purchased from Novabiochem. DCM and DMF were obtained from Merck. Custom made RRE and TAR RNA were synthesized by Sigma. [γ-^{32}P]ATP was obtained from Perkin Elmer. G50 Probe Quant columns were obtained from GE Healthcare. All antibiotics and chemicals were obtained from Sigma unless otherwise mentioned.

Methods

Synthesis of Rev/Tat based peptide substrates

The Rev based peptides, TRQARRNRRRRWRERQR-GGG-LPETGK (RevG3LPETGK), TRQARRNRRRRWRERQR (Rev Native) and TRQARRNRRRRWRERQR-GGG-LPET (RevG3LPET) and the Tat based peptide RKKRRQRRR-YGGG-LPETGK (TatG3LPETGK) were synthesized as has been described previously.

Sortase catalyzed ligation of Rev/Tat based peptides to aminoglycosides

The Rev and Tat based peptide-aminoglycoside conjugates were prepared as per the peptide-aminoglycoside ligation procedure described previously. The peptides were purified by RP-HPLC. Peptide concentration was determined spectrophotmetrically using molar extinction coefficients, $\varepsilon_{280,\ 1cm}$ = 5690, for the Rev related peptides, and $\varepsilon_{280,\ 1\ cm}$ = 1280, in case of Tat. Working dilutions of 8, 1, 0.25, 0.05 and 0.005 pmol/µl were made in DEPC treated water and stored at -20°C till use.

Electrophoretic mobility gel shift assay (EMSA)

The ability of Rev and Tat based peptides and conjugates to bind to RRE and TAR, respectively, was evaluated by electrophoretic mobility gel shift assays[170]. The

47 nucleotide (nt) RRE IIB sequence and 31nt TAR stem-loop sequence used in the RNA binding study are given below:

RRE IIB sequence
5' GGC UGG UAU GGG CGC AGC GUC AAU GAC GCU GAC GGU ACA GGC CAG CC 3'

TAR stem-loop sequence:
5' GGC CAG AUC UGA GCC UGG GAG CUC UCU GGC C 3'

Both the RNAs were custom synthesized by Sigma. The RNA was dissolved in DEPC treated water (DEPC was added to water to a final concentration of 0.1% (v/v), kept for 14-16 hours at room temperature and autoclaved) to give a stock solution of 10pmol/μl and sub-diluted to a working stock of 2pmol/μl.

All solutions were made in DEPC treated water and stored at room temperature unless otherwise mentioned. All chemicals were weighed on oven-baked (200°C, 4hrs) aluminum foil using baked, metal spatulas. Glassware used was similarly baked in the oven and only sterile, disposable plastic ware was used. Gloves were worn for all procedures.

The following buffers were prepared:
0.5M HEPES-NaOH; pH 7.5, 0.5M EDTA, 1M DTT and 10X TBE were prepared and autoclaved to maintain sterility.
2M KCl, 1M $MgCl_2$ and 50% glycerol were treated overnight with 0.1% DEPC and autoclaved.
40% acrylamide-bis acrylamide (38:2, w/w) was made in DEPC treated water.
Gel loading buffer (30% glycerol, 10mM HEPES-NaOH, 0.5% BPP and xylene cyanol) and 10 X reaction buffer (100mM HEPES-NaOH, 1M KCl, 10mM $MgCl_2$, 5mM EDTA and 10mM DTT) were aliquoted into small fractions and stored at -20°C till use.

Radiolabelling RRE IIB and TAR RNA

The RRE IIB and TAR RNAs were end-labeled with [γ-^{32}P]ATP using polynucleotide kinase (PNK).

Protocol for Radiolabeling:

Constituent	Volume (μl)
DEPC treated water	7
PNK buffer (10X)	2
TAR/ RRE RNA (2 pmol/ μl)	3 (6 pmol)
RNase Inhibitor	1 (40U)
[γ- ^{32}P]ATP (10μCi/ μl)	5 (50μCi)
PNK (10,000U/ ml)	2 (10U)

The labeling reaction was set up in in a total volume of 20μl as given above, incubated in a water bath preheated at 37°C for 30min. This was followed by heat denaturation at 95°C for 2min for inactivation of PNK. The labeled RNA was then separated from the unreacted radionucleotide by gel filtration chromatography using G50 probe quant spin columns from GE Healthcare. The radiolabeled material was diluted appropriately to obtain a RNA stock solution of 20 fmol/μl.

Protocol for gel shift assay reaction

The radiolabeled RNA mix was made as follows:

Constituent	Volume (μl)
DEPC treated water	123
10X reaction buffer	100 (2X)
50% glycerol	200 (20%)
tRNA (5mg/ml)	5 (50μg/ml)
RNase inhibitor	2 (80U)
Labeled RNA	50 (2 fmol/μl)

10µl of the master-mix was incubated with varying concentrations of Rev/Tat peptide or aminoglycoside conjugates for 30min at 4°C. The reactions were analyzed on a 10% native PAGE with 0.5X TBE buffer for 2hrs at 125V at 4°C. Subsequent to autoradiography, the gel bands corresponding to bound and free RNA were excised and the radioactivity associated with each band was determined using a liquid scintillation counter.

Results

Choice and design of bioconjugates

Replication of the human immunodeficiency virus-1 (HIV-1) in the host cell cytoplasm is critically dependent upon two viral regulatory proteins: the regulator-of-virion-expression protein (Rev) and the *trans*-activator of transcription protein (Tat) (Figure 1)[171, 172]. Rev is required for the efficient cytoplasmic export of the RNAs encoding viral structural proteins whereas Tat stimulates transcriptional elongation from the viral 5'- long terminal repeat (LTR). Both proteins exert their effects by binding to structured RNA elements.

Rev, a 116 amino acid phosphoprotein, binds to a *cis*-acting RNA regulatory element termed the Rev response element (RRE) through a 17 residue arginine rich RNA binding domain (residues 34-50, TRQARRNRRRRWRERQR; RBD) (Figure 2A)[173]. The RRE is a highly structured 234 nucleotide (nt) RNA; the high-affinity Rev binding site is localized to the bulge structure in stem-loop IIB (Figure 2B)[174-177]. The Rev interaction with RRE activates the cytoplasmic export of partially spliced and unspliced HIV-1 RNA which serves as open reading frame for viral structural proteins like Gag, Pol and Env[178-180].

Tat is an 86 residue protein that interacts with the trans activation response RNA (TAR) via a 9 residue stretch (residues 49-57, RKKRRQRRR) (Figure 2C)[181]. The TAR element is a 59nt RNA sequence located near the 5'-end of the viral genome[182-184]. The high-affinity Tat binding site has been localized to an unusual stem-loop structure containing a three-nucleotide bulge and a six-nucleotide loop (Figure 2D)[182, 183, 185, 186]. The formation of the Tat-TAR complex increases the processivity of the host RNA polymerase II transcription complex resulting in the production of longer viral RNAs and in increased viral gene expression[181, 187].

The Rev-RRE and Tat-TAR interactions were targeted for the development of bioconugates as it is known that disruption of these interactions impairs viral replication. The RRE sequence is an especially attractive target for therapeutic intervention as this viral sequence serves both as the high affinity Rev binding site and as part of the open reading frame for the envelope (env) protein. Due to this dual role, a very low rate of mutation is observed in RRE and it is hypothesized that

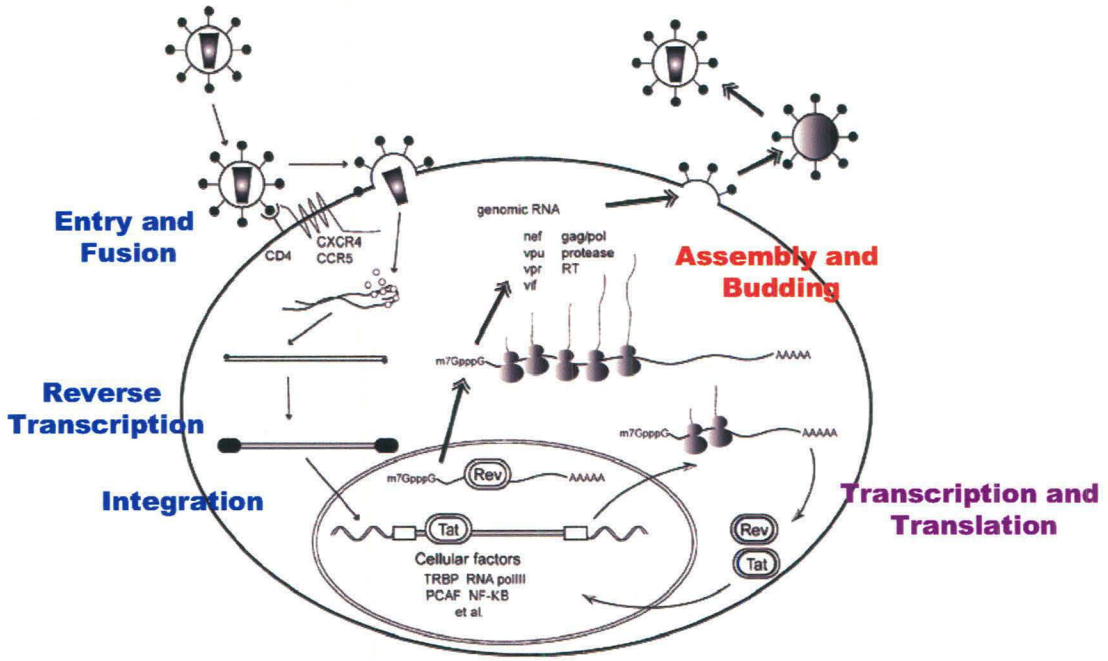


Figure 1. Diagrammatic representation of the HIV-1 life cycle showing the three major stages of replication: Entry and fusion, reverse transcription and integration (Stage I); Transcription and translation (Stage II) ; Assembly and budding (Stage III). Tat and Rev are important for stage II of the viral life cycle.

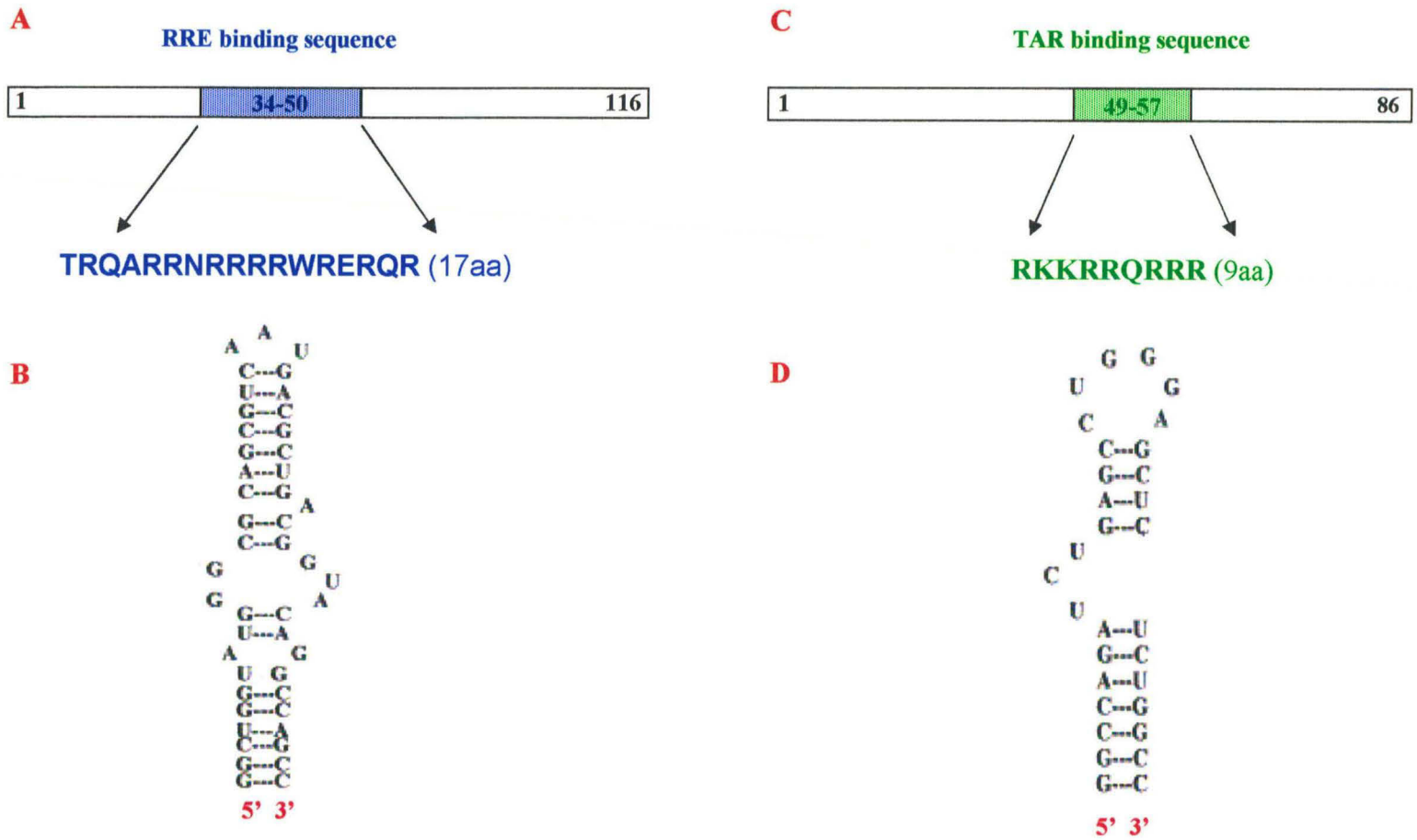


Figure 2. (A) 116 residue Rev with the sequence required for interacting with RRE RNA in blue **(B)** Sequence and secondary structure of the RRE IIB RNA used in study. RRE IIB spans the minimal sequences that are required for *in-vitro* binding of Rev-derived peptides. **(C)** Sequence of the Tat peptide (49-57) that recognizes TAR with high affinity and specificity. **(D)** Sequence and secondary structure of the high-affinity Tat binding region of TAR RNA used in study.

drugs targeted against the Rev-RRE interaction are unlikely to lead to the evolution of resistant variants[188-190].

Aminoglycoside antibiotics have been shown to inhibit the binding of the HIV-1 Rev and Tat proteins to their RNA counterparts, RRE and TAR, by directly binding to the later[188, 191, 192]. Three aminoglycosides, neomycin, tobramycin and ribostamycin, were chosen for conjugation to Rev or Tat derived peptides. The idea was to see if individual RNA binding propensity of peptide and antibiotics respectively is enhanced upon conjugation.

Preparation of the Rev/Tat peptide-antibiotics conjugates

Peptides corresponding to the 17 residue Rev RBD (RevG3LPETGK) and the 9 residue Tat RBD (TatG3LPETGK), appended with an LPETG motif and triglycine spacer, were synthesized. Additional Rev based peptides composed of only the Rev RBD sequence (Rev Native) and the Rev RBD appended with the truncated LPET motif and triglycine spacer (RevG3LPET) were also synthesized. The last two peptides were synthesized as controls; to evaluate if addition of the triglycine spacer and the LPETG motif has any impact on the ability of the peptide to bind to RRE RNA. The peptides were synthesized and purified as described previously (Figure 3). The identity of each peptide was verified by mass spectrometry (Table 1).

The experimentally determined masses for RevG3LPETGK, Rev Native, RevG3LPET and TatG3LPET were 3234.24 Da (calc: 3234.64 Da), 2436.40 Da (calc: 2436.39 Da), 3047.59 Da (calc: 3047.68 Da) and 2298.89 Da (calc: 2299.67 Da) respectively.

The peptide-antibiotics conjugates comprising RevG3LPETGK and TatG3LPETGK and neomycin, tobramycin and ribostamycin were prepared using the same procedure as described earlier (Figure 4A, B). The conjugates were purified by RP-HPLC (Figure 5) their identity was verified by mass spectroscopy as described earlier. The experimentally determined mass of the purified conjugates were in accord with their calculated mass (Table 1).

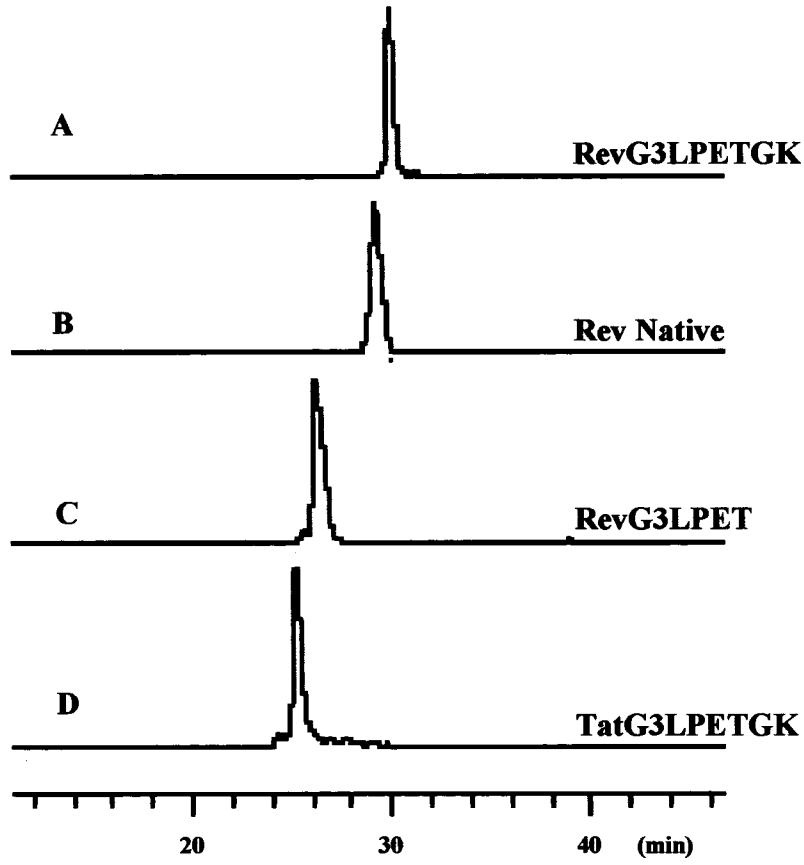


Figure 3. The elution profiles of purified **(A)** RevG3LPETGK; **(B)** Rev Native; **(C)** RevG3LPET and **(D)** TatG3LPETGK.

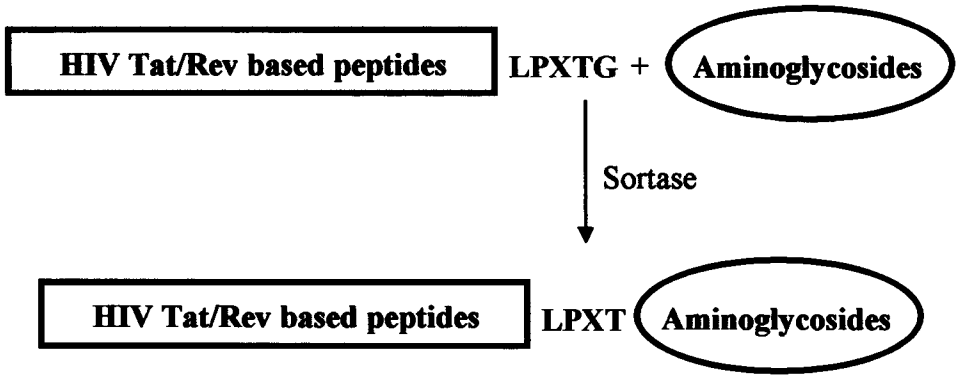


Figure 4A. Diagrammatic representation of the construction of Rev/Tat based peptide-aminoglycoside conjugates

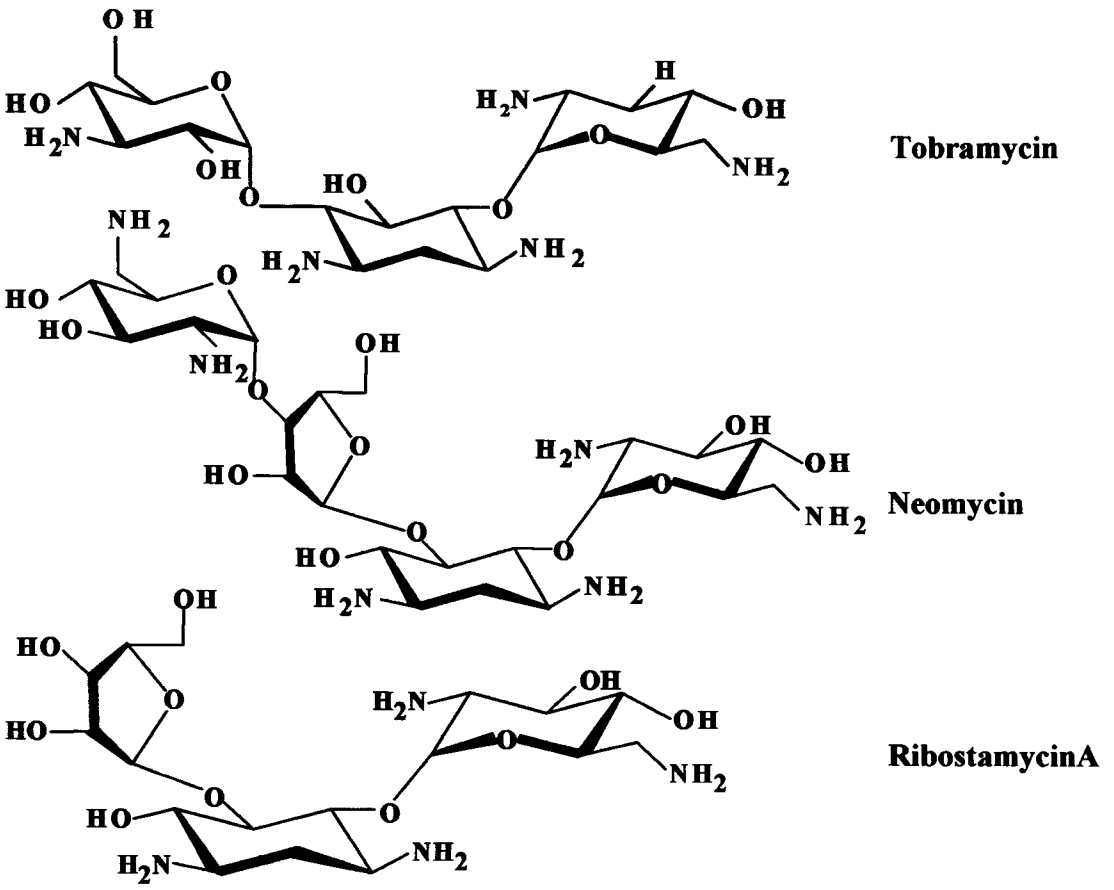


Figure 4B. The three aminoglycosides chosen for the study. Amine groups that serve as ligation sites for Sortase catalyzed transpeptidation are highlighted in green.

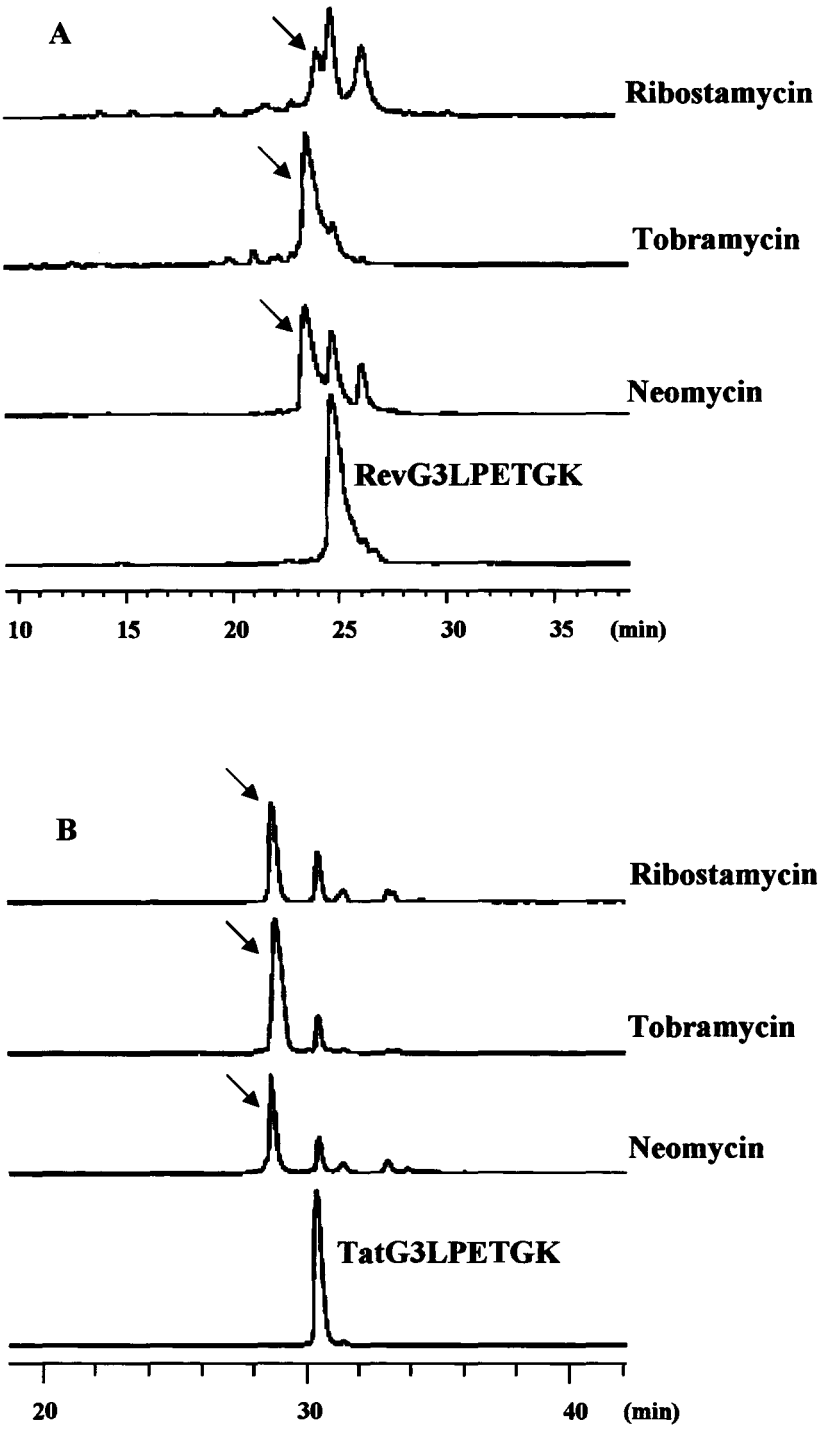


Figure 5. RP-HPLC analysis of sortase catalyzed ligation of neomycin, tobramycin and ribostamycin with **(A)** RevG3LPETGK; and **(B)** TatLPETGK. The conjugates are marked by arrows.

Table 1: ESMS characterization of Rev/Tat based peptides and peptide-aminoglycoside conjugates.

Peptide/Conjugate	Calculated mass (Da)	Experimental mass (Da)
RevG3LPETGK	3234.64	3234.24
Rev Native	2436.39	2436.40
RevG3LPET	3047.68	3047.59
TatG3LPETGK	2299.67	2298.89
RevG3LPET-Neomycin	3646.05	3646.55
RevG3LPET-Tobramycin	3498.39	3498.30
RevG3LPET-Ribostamycin	3484.15	3483.25
TatG3LPET-Neomycin	2711.09	2711.09
TatG3LPET-Tobramycin	2563.96	2563.32
TatG3LPET-Ribostamycin	2550.92	2551.14

Evaluation of RNA binding of conjugates by EMSA

The ability of Rev/Tat based peptide-aminoglycoside conjugates to bind to their respective RNAs was evaluated by EMSA as described in Materials and Methods (Figure 6A-13A). Following autoradiography, the radioactivity associated with the bound and free RNA fractions was estimated using liquid scintillation counting. K_d values, defined as the concentration of peptide/conjugate that displayed 50% binding to RRE/TAR RNA were estimated using the Graphpad Prism software assuming a 1:1 stoichiometry of binding. The K_d values are summarized in Table 2 and representative binding curves are shown in Figures 6B-13B.

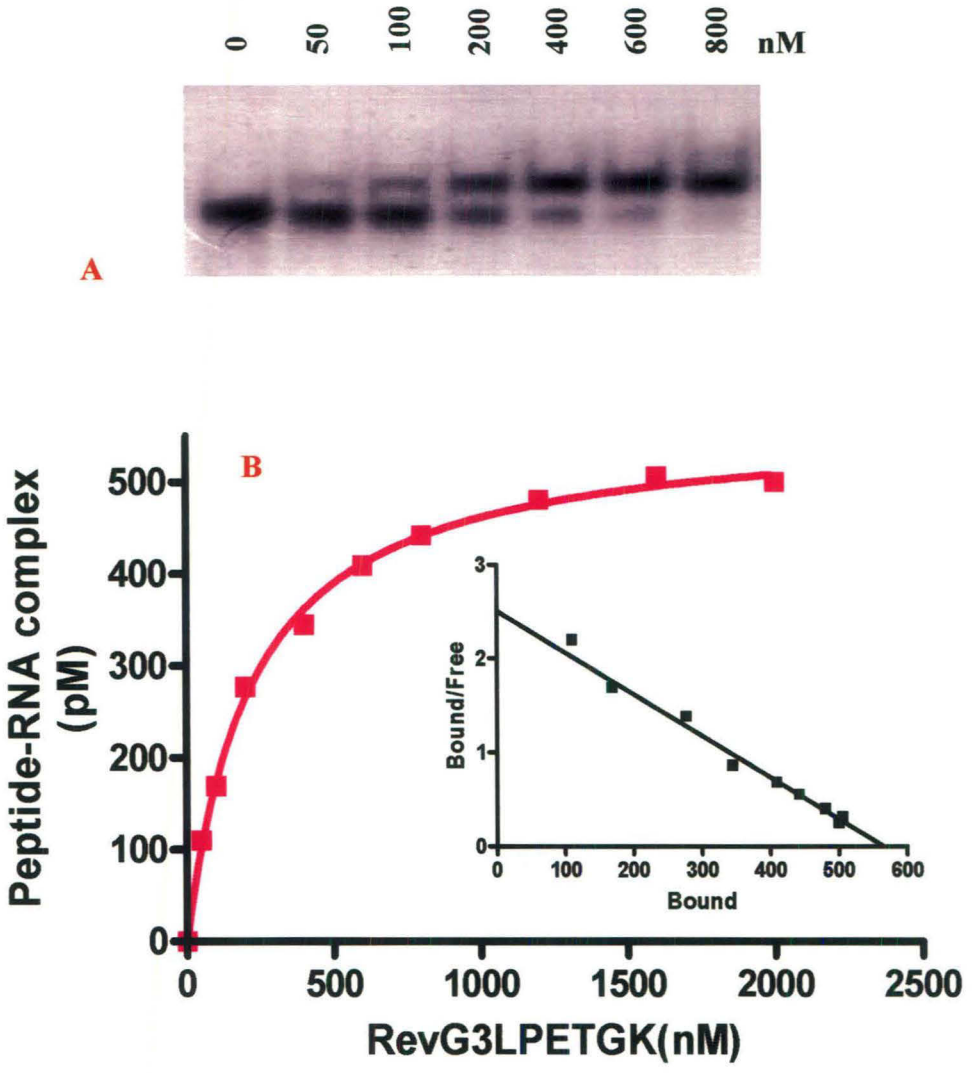


Figure 6. Electrophoretic mobility gel shift assays of RRE IIB RNA with RevG3LPETGK. Panel **(A)** shows a representative EMSA; panel **(B)** shows the analysis of the peptide-RNA binding. Inset shows the Scatchard plot. A dissociation constant of 216.68 ± 32.68 nM was estimated.

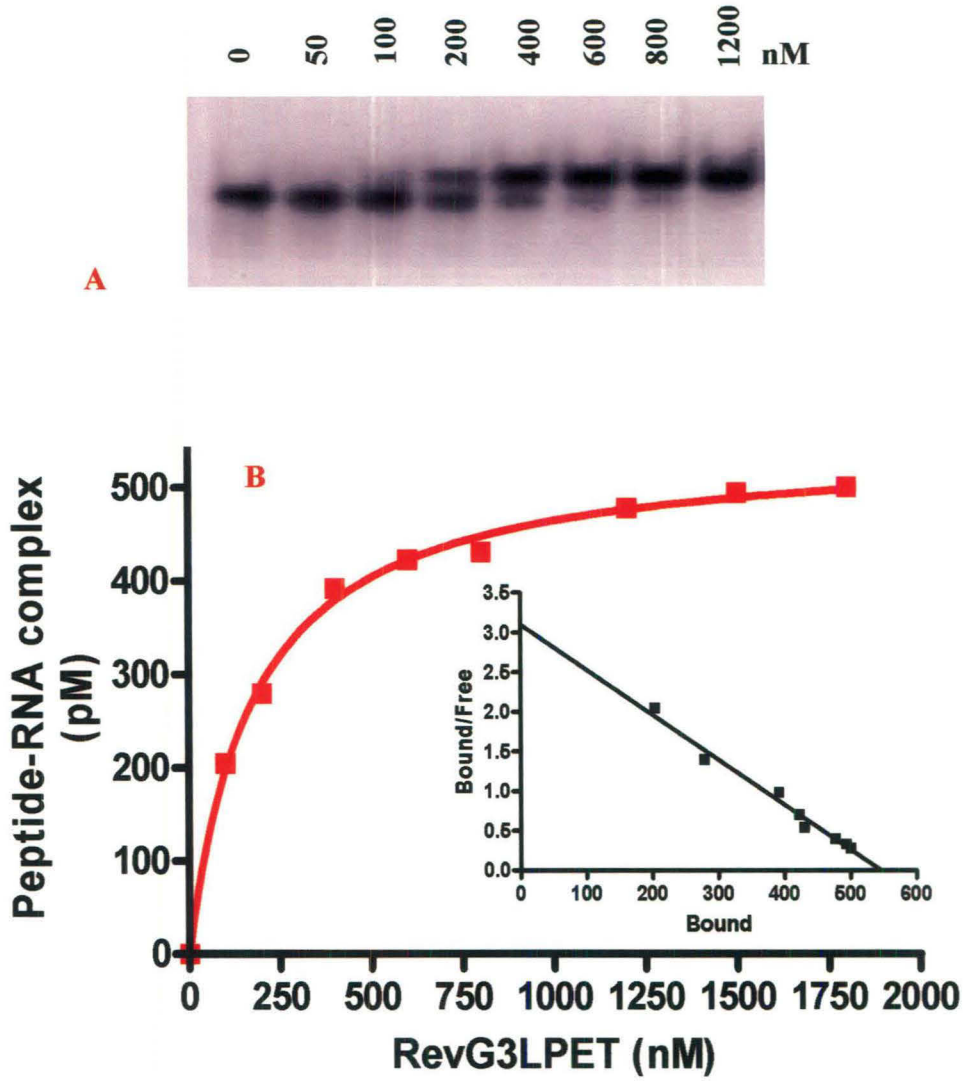


Figure 7. Electrophoretic mobility gel shift assays of RRE IIB RNA with RevG3LPET. Panel **(A)** shows a representative EMSA; panel **(B)** shows the analysis of the peptide-RNA binding. Inset shows the Scatchard plot. A dissociation constant of 172.45 ± 5.16 nM was estimated.

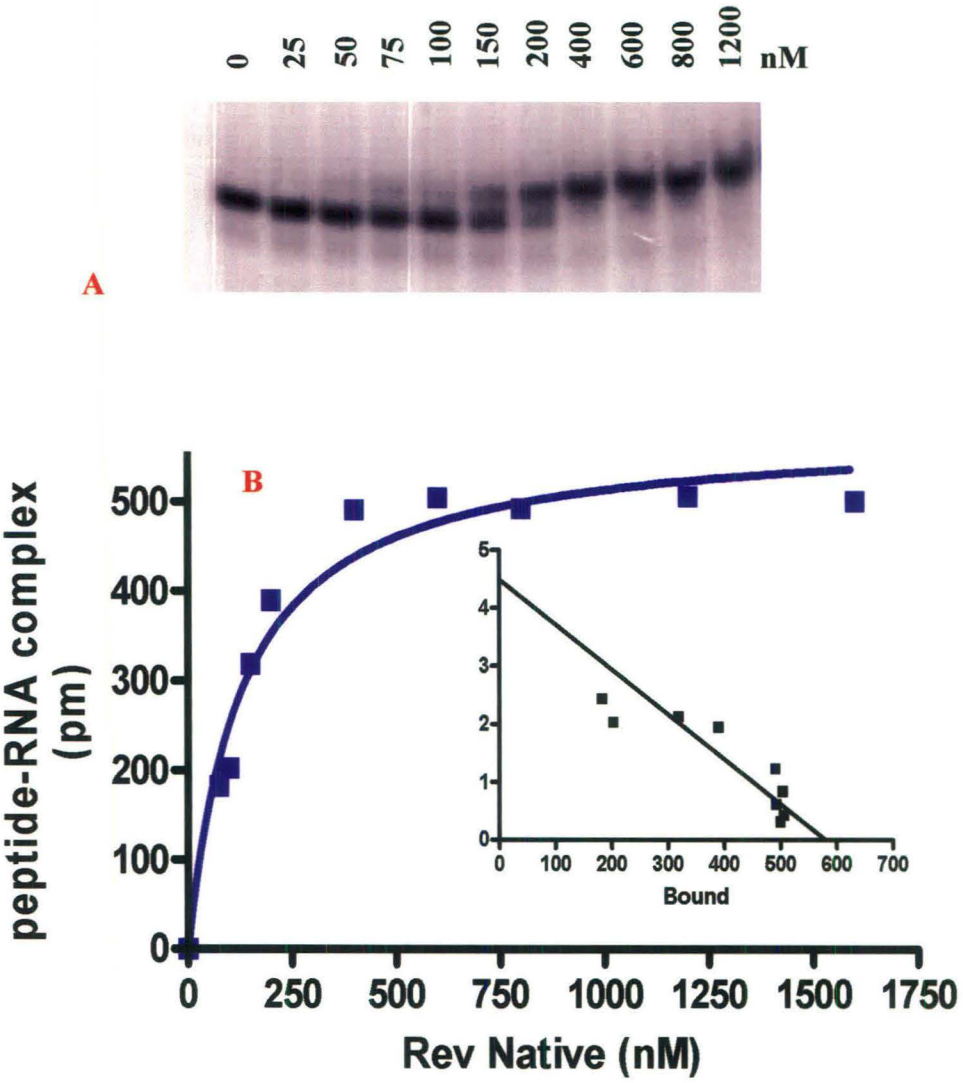


Figure 8. Electrophoretic mobility gel shift assays of RRE IIB RNA with Rev Native. Panel **(A)** shows a representative EMSA; panel **(B)** shows the analysis of the peptide-RNA binding. Inset shows the Scatchard plot. A dissociation constant of 128.37 ± 24.71 nM was estimated.

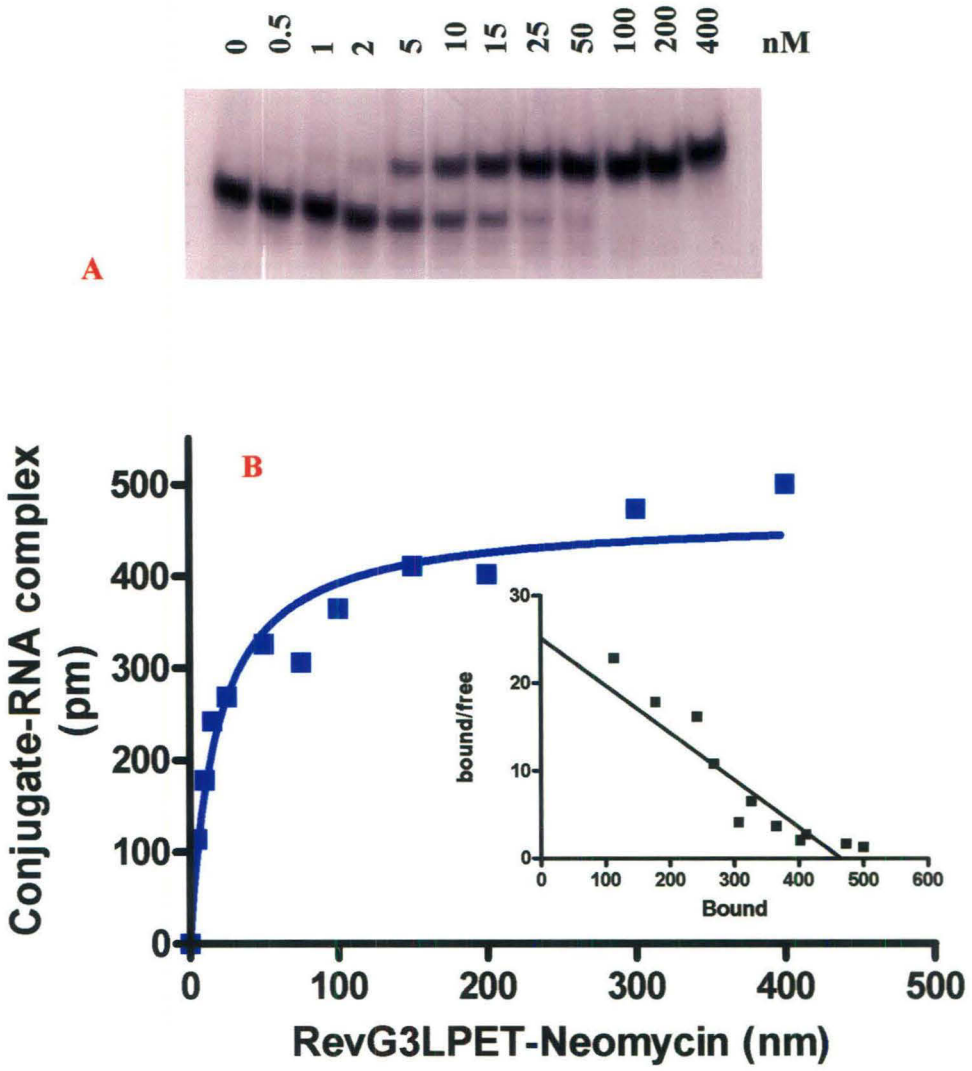


Figure 9. Electrophoretic mobility gel shift assays of RRE IIB RNA with RevG3LPET-Neomycin. Panel **(A)** shows a representative EMSA; panel **(B)** shows the analysis of the peptide-RNA binding. Inset shows the Scatchard plot. A dissociation constant of 9.20 ± 1.88 nM was estimated.

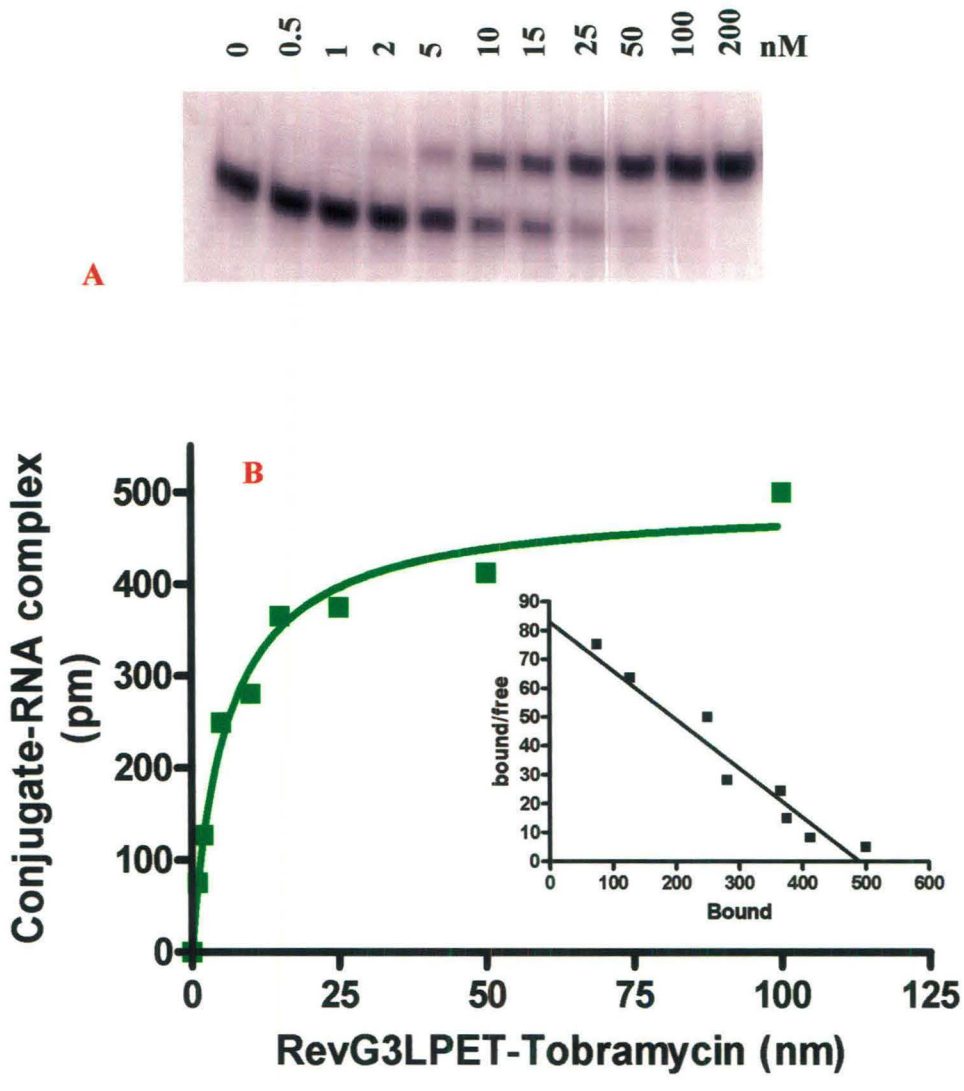


Figure 10. Electrophoretic mobility gel shift assays of RRE IIB RNA with RevG3LPET-Tobramycin. Panel **(A)** shows a representative EMSA; panel **(B)** shows the analysis of the peptide-RNA binding. Inset shows the Scatchard plot. A dissociation constant of 9.8 ± 5.50 nM was estimated.

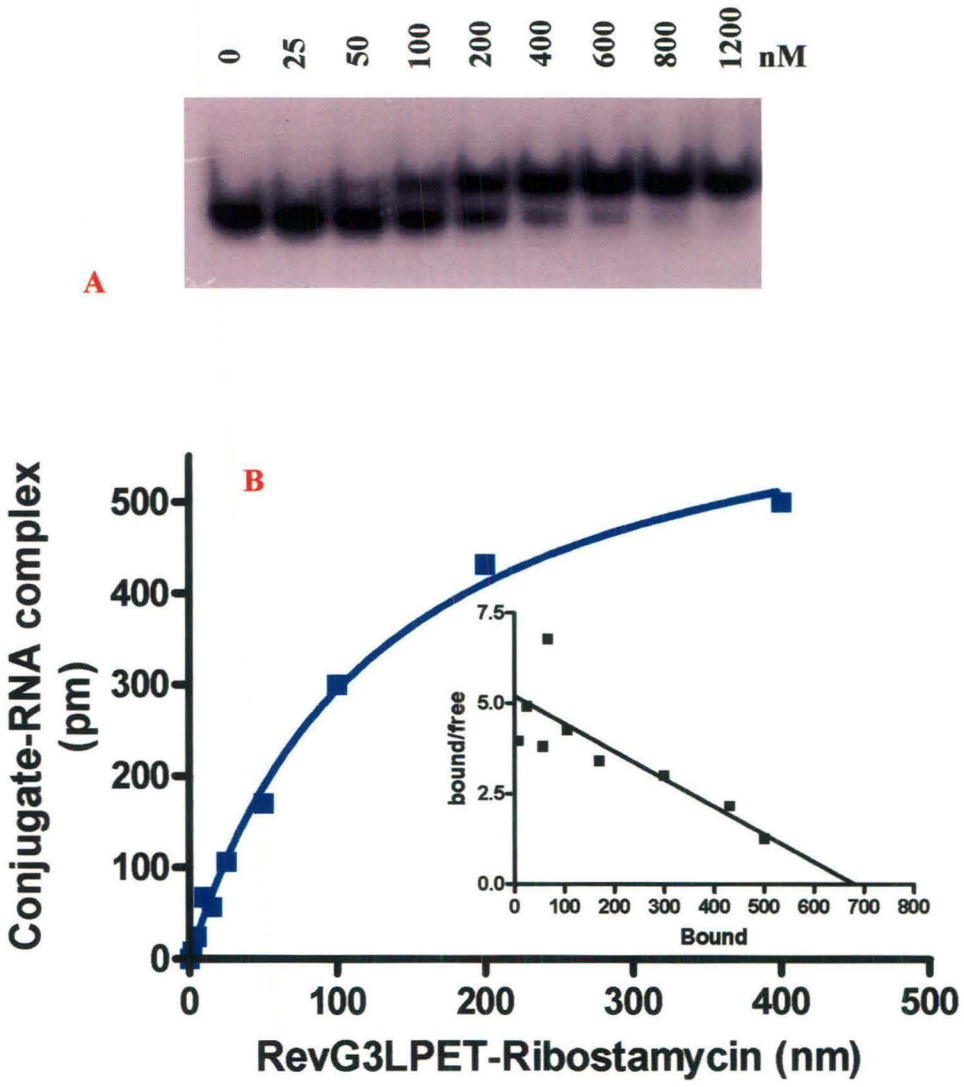


Figure 11. Electrophoretic mobility gel shift assays of RRE IIB RNA with RevG3LPET-Ribostamycin. Panel **(A)** shows a representative EMSA; panel **(B)** shows the analysis of the peptide-RNA binding. Inset shows the Scatchard plot. A dissociation constant of 147.6 ± 25.76 nM was estimated.

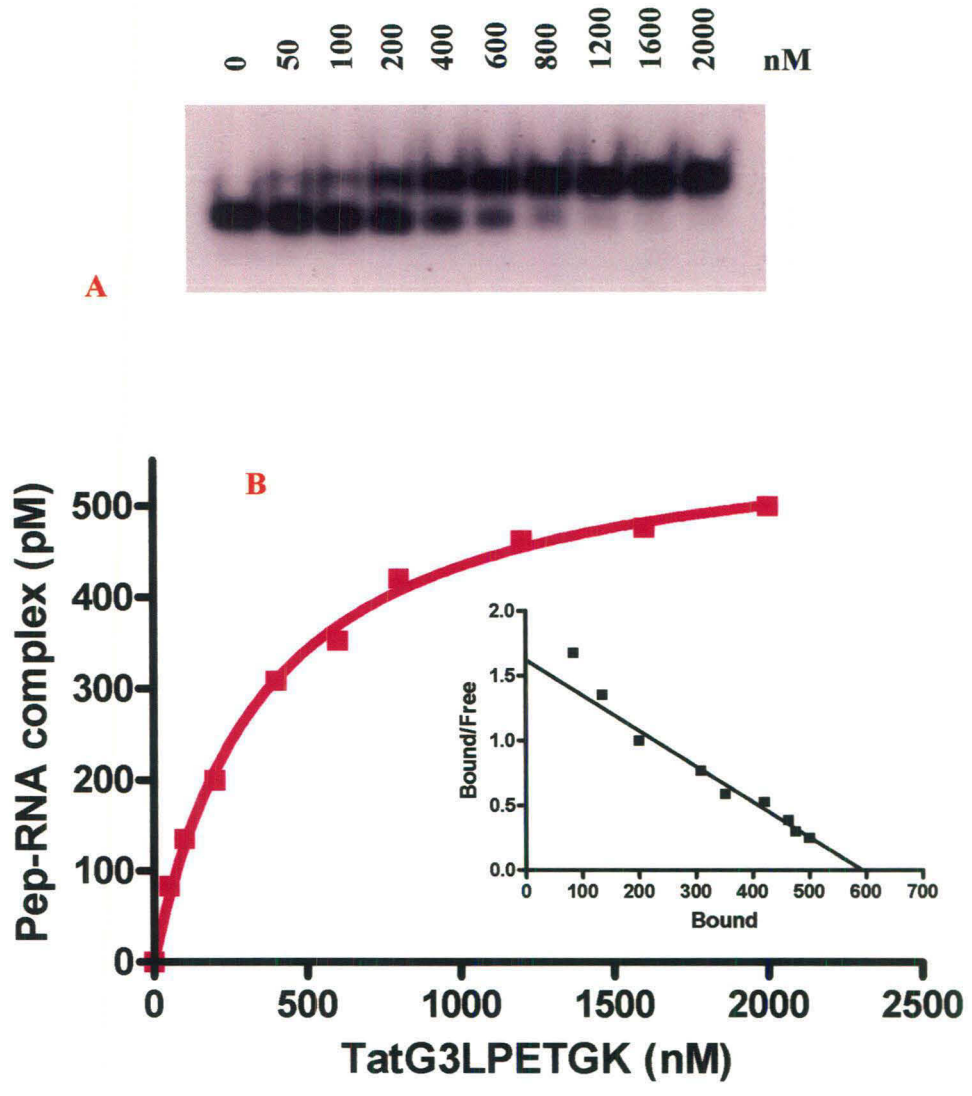


Figure 12. Electrophoretic mobility gel shift assays of TAR RNA with TatG3LPETGK. Panel **(A)** shows a representative EMSA; panel **(B)** shows the analysis of the peptide-RNA binding. Inset shows the Scatchard plot. A dissociation constant of 348.27 ± 26.13 was estimated.

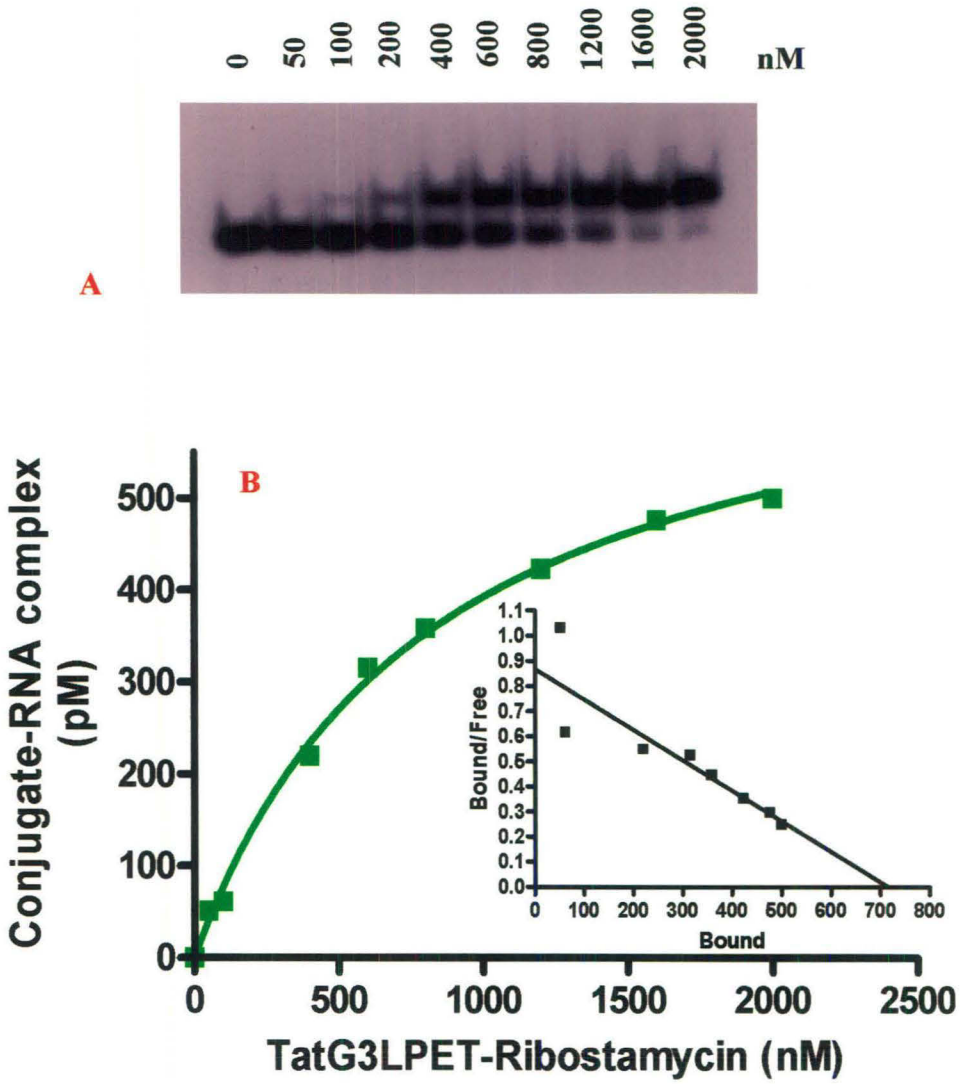


Figure 13. Electrophoretic mobility gel shift assays of TAR RNA with TatG3LPET-Ribostamycin. Panel **(A)** shows a representative EMSA; panel **(B)** shows the analysis of the conjugate-RNA binding. Inset shows the Scatchard plot. A dissociation constant of 775.67 ± 205.30 was estimated.

Table 2: Dissociation constants (K_d) for Rev/Tat based peptides and conjugates with RRE IIB and TAR RNA as determined by EMSA.

Peptide/Conjugate	RNA	K_d (nM)
RevG3LPETGK	RRE IIB	216.68 ± 32.68
RevG3LPET	RRE IIB	172.45 ± 5.16
Rev Native	RRE IIB	128.37 ± 24.71
RevG3LPET-Neomycin	RRE IIB	9.20 ± 1.88
RevG3LPET-Tobramycin	RRE IIB	9.8 ± 5.50
RevG3LPET-Ribostamycin	RRE IIB	147.6 ± 25.76
TatG3LPETGK	TAR	348.27 ± 26.13
TatG3LPET-Neomycin	TAR	684.8 ± 0.85
TatG3LPET-Tobramycin	TAR	802.6
TatG3LPET-Ribostamycin	TAR	775.67 ± 205.30

Discussion

Currently available therapies for treatment of AIDS are insufficient due to their inability to cure the disease and the propensity of the virus to become resistant to the treatment over time. The need for alternative chemotherapeutic strategies has led to a search for and development of agents that interrupt critical events in the HIV life cycle. A current direction in this field is the development of drugs that interfere with the action of the Rev and Tat proteins. Rev and Tat are regulatory proteins that are crucial for HIV-1 replication. Rev is required for the efficient cytoplasmic expression of the mRNAs encoding the viral structural proteins whereas Tat stimulates transcriptional elongation by increasing the processivity of host RNA polymerase II. Both proteins exert their effect by binding to structured RNA sequences, RRE in case of Rev and TAR in case of Tat. The RRE sequence is an especially attractive target for therapeutic intervention as this viral sequence serves both as the high affinity Rev binding site and as part of the open reading frame for the envelope (env) protein. Due to this dual role, drugs targeted against the Rev-RRE interaction are unlikely to lead to the evolution of resistant variants[188-190].

The RNA binding propensity of aminoglycoside antibiotics is widely acknowledged. Since Moazed and Noller's [193] landmark report localizing the binding site of aminoglycosides to the decoding region on the 16S rRNA, numerous discoveries have established aminoglycosides as universal or promiscuous RNA binders. These antibiotics have been demonstrated to bind to a large number of structurally and functionally unrelated RNAs like the hammerhead ribozyme, the hepatitis delta virus ribozyme, RNase P, transfer RNAs as well as the group I introns[194-198]. They have also been shown to inhibit the binding of the HIV-1 Rev and Tat proteins to their RNA counterparts, RRE and TAR, by directly binding to the later[188, 191, 192]. While these naturally occurring antibiotics are unlikely to display the necessary selectivity to become therapeutically relevant anti-viral drugs by themselves, they provide a foundation for the development of more potent and selective binders.

Several reports on aminoglycosides and aminoglycoside-based Rev/Tat inhibitors have been published. Zapp *et al* were the first to demonstrate the potential of aminoglycosides antibiotics as potential anti-HIV drugs by reporting their ability

to block the binding of Rev with RRE[188]. They reported that aminoglycosides like neomycin, tobramycin and lividomycinA had significant inhibitory potential whereas several others, like kanamycinA, B, ribostamycin, butirosinA and neamine, were inhibitory only at extremely high concentrations (> 100μM). They particularly highlighted the ability of neomycin to inhibit Rev binding at concentrations as low as 1μM. Their attempts at localizing the binding of neomycin indicated that the binding sites for Rev and aminoglycosides overlapped. Thus, inhibition by aminoglycosides is most likely due to the mutually exclusive interaction with common RRE nucleotides especially the GGG region of RRE IIB. Luedtke *et al* designed aminoglycoside dimers and aminoglycoside conjugates with 9-aminoacridine as potential RRE binders[199]. Although both types of compounds exhibited high affinity binding to RRE RNA, a high degree of specificity for the RRE target was lacking. Park *et al* used a Ugi-type one-pot reaction to synthesize a library of neomycin mimetics several of whom were found to be better RRE binders than the parent aminoglycoside[200]. Aminoglycosides are also known to inhibit Tat-TAR interactions albeit at much higher concentrations. Wang *et al* have reported that *in-vitro* inhibition of the interaction by neomycin occurs only when concentrations are as high as 1mM[191]. Lapidot and coworkers synthesized conjugates of L-arginine with aminoglycosides as potential Tat-TAR inhibitors. These compounds were found to bind TAR with K_d values in the range of 200nM-2.5μM as compared to the native Tat which bound TAR in the range of 6-15nM. The neomycin-hexaarginine conjugate bound TAR in the range of 130nM whereas the R52 peptide, which is a Tat-derived peptide, bound TAR at 50nM. They also demonstrated the efficacy of these conjugates in inhibiting HIV replication *in-vivo*[201-205]. Riguet and coworkers reported the conjugation of neamine to a TAR targeted PNA sequence[206]. The resulting conjugate exhibited TAR cleavage *in vitro* and anti-HIV activity in infected cells. Lee and coworkers reported TAR and RRE binding of several neomycin conjugates containing chloramphenicol and linezolid[207]. The conjugates displayed approximately 10-fold higher affinity to TAR compared to the parent aminoglycoside. Hamasaki *et al* reported that neamine derivatives containing pyrene and arginine bound TAR and RRE with equivalent binding affinities as Tat and Rev peptide respectively[208].

A large number of the aminoglycoside derivatives discussed above are capable of high affinity binding to the RRE or TAR RNA site but lack the degree of specificity required to be of therapeutic use. Introduction of a moiety specific for TAR or RRE into the aminoglycoside scaffold could help endowing the conjugate with the desired ability to discriminate between RNAs. Peptides derived from Rev and Tat ('decoy' peptides) have been shown to interfere with Tat-TAR and Rev-RRE interactions leading to inhibition of viral replication[209-215].

Sortase mediated ligation was employed to create conjugates comprising of an aminoglycoside ligated with Rev/Tat based peptides. The objective was to perceive if combining two separate Rev/Tat binding modules could have a synergistic effect on the RNA binding propensity of the conjugate. The Rev/Tat peptides were synthesized with a sortase recognition motif (LPETG) and a triglycine spacer. Three aminoglycosides, namely neomycin, tobramycin and ribostamycin, were employed for this study. Neomycin and tobramycin belong to the 4, 5 and 4, 6 substituted class and were chosen for this study as they are known to be the most effective inhibitors of Rev/RRE and Tat/TAR interactions. Ribostamycin also belongs to the 4, 5 substituted class of aminoglycosides but its ability to inhibit the HIV-1 protein-RNA interaction has been reported to be much less as compared to neomycin or tobramycin[188]. To assess the impact of the addition of the LPETG motif and the triglycine linker on the ability of the peptides to bind RNA, two additional peptides were synthesized as controls. Rev Native is composed of only the 17 residue RBD whereas RevG3LPET contains the triglycine spacer and a truncated LPET motif. The ability of these conjugates to bind to their respective RNAs was evaluated by electrophoretic mobility gel shift assays using the protocol outlined by Mill *et al*[170].

The gel shift assays with the RevG3LPETGK, RevG3LPET and Rev Native yielded comparable RRE binding affinities [(K_d) of 216.68 ± 32.68 nM, 172.45 ± 5.16 nM and 128.37 ± 24.71 nM] respectively which compares well with RRE binding affinities of the Rev peptide (50-100nM)[170, 216, 217] reported in other studies. The binding of RevG3LPET-ribostamycin to RRE RNA (K_d ~147.6 ± 25.76 nM) was similar to that of the Rev Native peptide. In contrast, RevG3LPET-Neomycin and RevG3LPET-Tobramycin conjugates showed about 10-12 fold improvement of RRE RNA binding (K_d values of 9-10nM).

A similar enhancement of RNA binding affinity, however, was not observed for the Tat-aminoglycoside conjugates for TAR sequence. The K_d of the TatG3LPET-Neomycin, tobramycin and ribostamycin conjugates (684.8 ± 0.85 nM, 802.6 nM and 775.67 ± 205.30 nM respectively) were not significantly different from that of the Tat G3LPETGK peptide (K_d: 348.27 ± 26.13 nM).

The structural and mechanistic basis for the several fold improvement of RNA binding in the case of the Rev conjugates merits further study. Nonetheless, the results show that sortase can be a useful tool for synthesis of improved RNA binders as has been demonstrated here in the case of the HIV-1 RRE RNA.

Summary

Development of new methods for linking sugars to peptides or proteins is an active area of research because natural glycopeptides or neoglycoconjugates play important roles in biology and medicine and are indispensable tools for probing several biological processes. However, glycoconjugate synthesis remains a formidable task as synthetic protocols are quite demanding and involve multiple reaction steps with requirements of extensive protection of reactive functionalities. Given the current ease with which peptides are assembled by solid phase methodology and proteins obtained from expression systems, the availability of enzymes capable of covalently linking a pre-synthesized sugar and a polypeptide would greatly facilitate the synthesis of glycoconjugates.

The membrane-anchored sortase enzymes present in gram positive bacteria are a group of cysteine transpeptidases that catalyze the anchoring of several cell surface proteins to the cell wall peptidoglycan. The prototypical sortase, sortase A of *Staphylococcus aureus,* recognizes a LPXTG like sequence motif located near the C-terminus of the target proteins, cleaves at Thr-Gly peptide bond, and catalyzes the formation of a new peptide bond between threonyl carboxyl and amino group of the pentaglycine cross-bridges of peptidoglycan. The transpeptidation reaction proceeds in two steps without the aid of any extraneous molecule; the active site cysteine residue first attacks the target LPXTG substrate forming an acyl-enzyme intermediate which in the second step is resolved by the nucleophillic attack of the amino group of the terminal glycine residue of the peptidoglycan. In the absence of a suitable amino nucleophile, the LPXTG peptide substrate is slowly hydrolyzed.

Sortase mediated transpeptidation reaction involving LPXTG and aminoglycine containing polypeptides proceeds smoothly *in vitro* and the feasibility of using sortase in semi-synthetic strategies has been tested in some cases. It has been used to synthesize peptide-peptide, protein-peptide and protein nucleic acid–peptide conjugates that would have been difficult to assemble by purely chemical or genetic means. Interestingly, ligation of LPXTG substrates can occur even with polypeptides containing a single glycine residue at the amino terminus. This observation of relaxed specificity for the amine nucleophile prompted further exploration of sortase-mediated ligation of polypeptides to amino sugars with a view to develop an enzymatic approach for synthesis of glycoconjugates with potential bioactivity.

$SrtA_{\Delta 59}$ sequence corresponding to amino acids 60-204 was amplified from the *Staphylococcus aureus Mu50* genomic DNA and cloned into pET23b vector. Purification of protein was carried out as per standard Ni-NTA procedure. The identity of recombinant sortase was established by Electrospray (ES) mass spectrometry (MS) and Edman degradation. 6-aminohexoses were considered as potential sugar substrates with the idea that the $-CH_2-NH_2$ moiety present in these sugars might mimic some elements of the glycine structure. Accordingly, the potential of sortase to ligate 6-amino-6-deoxyglucose and 6-amino-6-deoxymannose to the YALPETGK peptide substrate was tested. HPLC assays followed by MALDI-TOF or ESMS analysis revealed the formation of the respective YALPET-sugar adducts suggesting that the above amino sugars acted as nucleophiles in the transamidation reaction. To further probe the specificity requirements as well as to see if 6-aminohexoses can serve as recognition tags for peptide-sugar ligations, the ability of sortase to ligate peptides the aminoglycoside class of therapeutically important antibiotics was examined. Sortase mediated ligation of YALPETGK to aminoglycoside proceeded as expected. Analyses of the reaction products by RP-HPLC followed by MALDI-TOF revealed the formation of specific conjugates between antibiotics and YALPETGK in the yields varying from 35-70% for the kanamycin class, and 18-30% for the ribostamycin class of antibiotics. ESMS of the respective conjugates produced fragmentations that unambiguously showed occurrence of peptide ligation exclusively at a single 6-amino site in ring A of the kanamycins, tobramycin and ribostamycin or ring D of paromomycin and neomycin. Thus, conjugation of peptide substrates was limited to the 6-amino site in the antibiotics despite the presence of a plethora of amino groups, indicating rather strict specificity and selectivity for the sugar amino groups by sortase.

Transpeptidation of LPXTG peptide substrates to aminoglycosides amikacin and butirosinA that belong to the kanamycin and ribostamycin class respectively but contain an additional side chain with a terminal amino group was also investigated. Both these antibiotics produced peptide-aminoglycoside adduct at the amino group present in the butyryl side chain in addition to that at the 6-amino site. Interestingly, the butyryl side chains, 4-amino-2-deoxy butyric acid and 4-amino-3-hydroxy butyric acid, independently acted as substrates in the sortase catalyzed transpeptidation reaction. Likewise, 6-amino hexanoic acid, 7-amino heptanoic acid,

6-aminohexanol, spermine, cadaverine and agmatine also formed adducts with LPXTG peptides.

The feasibility of using sortase for site-specific conjugation of sugars to proteins was also tested. For this a protein (Mrp, NP_372281) from *Staphylococcus aureus* nested with a LPNTG sequence motif in its carboxy terminal region was expressed. Incubation of Mrp with sortase in the presence of tobramycin produced specific conjugates (Mrp_{1-170}-Tobramycin). Similar conjugates were also obtained with other aminoglycoside antibiotics.

The conjugation of biologically relevant peptides to aminoglycoside antibiotics was also explored. Towards this, peptides derived from Tat and Rev proteins of HIV were considered because these proteins play important roles in viral replication through their interactions with structured viral RNA targets; TAR in case of Tat and RRE in case of Rev. Aminoglycosides and peptides derived from Tat and Rev have been independently shown to interfere with Tat-TAR and Rev-RRE interactions leading to inhibition of viral replication. Using sortase-catalyzed transpeptidation, conjugates of neomycin, tobramycin and ribostamycin with Rev and Tat sequences appended with an LPETG motif were prepared. The ability of these conjugates to bind to their respective RNA was evaluated by electrophoretic mobility gel shift assays. The gel shift assays for the Neomycin/tobramycin-Rev conjugates yielded a RRE binding affinity that was about 10-12 fold higher as compared with the native Rev peptide. In contrast, the binding affinities of the Tat-aminoglycoside conjugates were found to be more or less similar to the Tat peptide. The overall results suggested the utility of using sortase-catalyzed transpeptidation for the generation of tight RNA binders with potential bioactivity.

In summary, the work presented in this thesis demonstrates that *Staphylococcus aureus* sortase can transfer peptide substrates to oligosaccharides appended with a 6-deoxy-6-aminohexose moiety in a selective manner as that of an oligoglycine sequence. This reaction provides a simple and straightforward method for covalent ligation of a prefabricated sugar containing a 6-aminohexose tag to synthetic peptides and expressed proteins encoded with a C-terminal LPXTG sortase recognition sequence. The reaction can be exploited for the facile assembly of aminoglycoside antibiotic-peptide conjugates with improved RNA binding properties. In addition, the results show that sortase can transfer LPXTG containing

peptides/proteins to a plethora of substrates such as α-amino hydroxy acids, long chain aliphatic amines and aminoalcohols/acids. Sortase appears to be a highly promiscuous enzyme endowed with tremendous synthetic capability.

Bibliography

1. Patti, J.M., *et al.* (1994) MSCRAMM-mediated adherence of microorganisms to host tissues. *Annu Rev Microbiol* 48, 585-617

2. Foster, T.J., and Hook, M. (1998) Surface protein adhesins of *Staphylococcus aureus*. *Trends Microbiol* 6, 484-488

3. Schneewind, O., *et al.* (1995) Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. *Science* 268, 103-106

4. Hancock, I.C. (1997) Bacterial cell surface carbohydrates: structure and assembly. *Biochem Soc Trans* 25, 183-187

5. Ghuysen, J.M., and Strominger, J.L. (1963) Structure of the Cell Wall of *Staphylococcus aureus*, Strain Copenhagen. II. Separation and Structure of Disaccharides. *Biochemistry* 2, 1119-1125

6. Ghuysen, J.M., and Strominger, J.L. (1963) Structure of the Cell Wall of *Staphylococcus aureus*, Strain Copenhagen. I. Preparation of Fragments by Enzymatic Hydrolysis. *Biochemistry* 2, 1110-1119

7. Ghuysen, J.M., *et al.* (1965) Structure of the Cell Wall of *Staphylococcus aureus*, Strain Copenhagen. IV. The Teichoic Acid-Glycopeptide Complex. *Biochemistry* 4, 474-485

8. Petit, J.F., *et al.* (1966) Peptide cross-links in bacterial cell wall peptidoglycans studied with specific endopeptidases from *Streptomyces albus* G. *Biochemistry* 5, 2764-2776

9. Tipper, D.J., *et al.* (1967) Structure of the cell wall of *Staphylococcus aureus*, strain Copenhagen. VII. Mode of action of the bacteriolytic peptidase from Myxobacter and the isolation of intact cell wall polysaccharides. *Biochemistry* 6, 906-920

10. Tipper, D.J. (1969) Structures of the cell wall peptidoglycans of *Staphylococcus epidermidis* Texas 26 and *Staphylococcus aureus* Copenhagen. II. Structure of neutral and basic peptides from hydrolysis with the Myxobacter al-1 peptidase. *Biochemistry* 8, 2192-2202

11. Tipper, D.J., and Berman, M.F. (1969) Structures of the cell wall peptidoglycans of *Staphylococcus epidermidis* Texas 26 and *Staphylococcus aureus* Copenhagen. I. Chain length and average sequence of cross-bridge peptides. *Biochemistry* 8, 2183-2192

12. Tipper, D.J., and Strominger, J.L. (1968) Biosynthesis of the peptidoglycan of bacterial cell walls. XII. Inhibition of cross-linking by penicillins and cephalosporins: studies in *Staphylococcus aureus* in vivo. *J Biol Chem* 243, 3169-3179

13. Tipper, D.J., and Strominger, J.L. (1965) Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc Natl Acad Sci U S A* 54, 1133-1141

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

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

16. Strominger, J.L. (1968) Penicillin-sensitive enzymatic reactions in bacterial cell wall synthesis. *Harvey Lect* 64, 179-213

17. Strominger, J.L., *et al.* (1967) Peptidoglycan transpeptidase and D-alanine carboxypeptidase: penicillin-sensitive enzymatic reactions. *Fed Proc* 26, 9-22

18. Gunetileke, K.G., and Anwar, R.A. (1966) Biosynthesis of uridine diphospho-N-acetyl muramic acid. *J Biol Chem* 241, 5740-5743

19. Ito, E., and Strominger, J.L. (1960) Enzymatic synthesis of the peptide in a uridine nucleotide from *Staphylococcus aureus*. *J Biol Chem* 235, PC5-6

20. Ito, E., and Strominger, J.L. (1964) Enzymatic Synthesis of the Peptide in Bacterial Uridine Nucleotides. III. Purification and Properties of L-Lysin-Adding Enzyme. *J Biol Chem* 239, 210-214

21. Mizuno, Y., *et al.* (1973) Purification and properties of uridine diphosphate N-acetylmuramate: L-alanine ligase. *J Biochem* 74, 525-538

22. Nathenson, S.G., *et al.* (1964) Enzymatic Synthesis of the Peptide in Bacterial Uridine Nucleotides. IV. Purification and Properties of D-Glutamic Acid-Adding Enzyme. *J Biol Chem* 239, 1773-1776

23. Ito, E., and Strominger, J.L. (1960) Enzymatic addition of lysine to a uridine nucleotide. *J Biol Chem* 235, PC7-8

24. Higashi, Y., *et al.* (1970) Biosynthesis of the peptidoglycan of bacterial cell walls. XIX. Isoprenoid alcohol phosphokinase. *J Biol Chem* 245, 3683-3690

25. Higashi, Y., and Strominger, J.L. (1970) Biosynthesis of the peptidoglycan of bacterial cell walls. XX. Identification of phosphatidylglycerol and cardiolipin as cofactors for isoprenoid alcohol phosphokinase. *J Biol Chem* 245, 3691-3696

26. Higashi, Y., *et al.* (1970) Biosynthesis of the peptidoglycan of bacterial cell walls. XXI. Isolation of free C55-isoprenoid alcohol and of lipid intermediates in peptidoglycan synthesis from *Staphylococcus aureus*. *J Biol Chem* 245, 3697-3702

27. Anderson, J.S., *et al.* (1967) Biosythesis of the peptidoglycan of bacterial cell walls. II. Phospholipid carriers in the reaction sequence. *J Biol Chem* 242, 3180-3190

28. Katz, W., *et al.* (1967) Biosynthesis of the peptidoglycan of bacterial cell walls. IV. Incorporation of glycine in *Micrococcus lysodeikticus*. *J Biol Chem* 242, 3207-3217

29. Ghuysen, J.M. (1991) Serine beta-lactamases and penicillin-binding proteins. *Annu Rev Microbiol* 45, 37-67

30. Nakagawa, J., *et al.* (1984) Functional biosynthesis of cell wall peptidoglycan by polymorphic bifunctional polypeptides. Penicillin-binding protein 1Bs of *Escherichia coli* with activities of transglycosylase and transpeptidase. *J Biol Chem* 259, 13937-13946

31. Rossi, M.S., *et al.* (2003) Haemophore-mediated signal transduction across the bacterial cell envelope in *Serratia marcescens*: the inducer and the transported substrate are different molecules. *Mol Microbiol* 48, 1467-1480

32. Van Bambeke, F., *et al.* (2000) Antibiotic efflux pumps. *Biochem Pharmacol* 60, 457-470

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

34. Flock, J.I., *et al.* (1987) Cloning and expression of the gene for a fibronectin-binding protein from *Staphylococcus aureus*. *EMBO J* 6, 2351-2357

35. Chen, C.C., and Cleary, P.P. (1990) Complete nucleotide sequence of the streptococcal C5a peptidase gene of *Streptococcus pyogenes*. *J Biol Chem* 265, 3161-3167

36. Fox, E.N. (1974) M proteins of group A streptococci. *Bacteriol Rev* 38, 57-86

37. Lancefield, R.C. (1962) Current knowledge of type-specific M antigens of group A streptococci. *J Immunol* 89, 307-313

38. Moks, T., *et al.* (1986) Staphylococcal protein A consists of five IgG-binding domains. *Eur J Biochem* 156, 637-643

39. Heermann, R., and Jung, K. (2004) Structural features and mechanisms for sensing high osmolarity in microorganisms. *Curr Opin Microbiol* 7, 168-174

40. Hazelbauer, G.L., *et al.* (2008) Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem Sci* 33, 9-19

41. Lopez, R., *et al.* (2000) Biological roles of two new murein hydrolases of *Streptococcus pneumoniae* representing examples of module shuffling. *Res Microbiol* 151, 437-443

42. Holtje, J.V., and Tomasz, A. (1975) Specific recognition of choline residues in the cell wall teichoic acid by the N-acetylmuramyl-L-alanine amidase of Pneumococcus. *J Biol Chem* 250, 6072-6076

43. Kocks, C., *et al.* (1992) *L. monocytogenes*-induced actin assembly requires the actA gene product, a surface protein. *Cell* 68, 521-531

44. Wang, P.Z., and Novick, R.P. (1987) Nucleotide sequence and expression of the beta-lactamase gene from *Staphylococcus aureus* plasmid pI258 in *Escherichia coli*, *Bacillus subtilis*, and *Staphylococcus aureus*. *J Bacteriol* 169, 1763-1766

45. Sutcliffe, I.C., and Harrington, D.J. (2002) Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. *Microbiology* 148, 2065-2077

46. Braun, L., *et al.* (1997) InlB: an invasion protein of *Listeria monocytogenes* with a novel type of surface association. *Mol Microbiol* 25, 285-294

47. Jonquieres, R., *et al.* (1999) Interaction between the protein InlB of *Listeria monocytogenes* and lipoteichoic acid: a novel mechanism of protein association at the surface of gram-positive bacteria. *Mol Microbiol* 34, 902-914

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

49. Mazmanian, S.K., *et al.* (1999) *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science* 285, 760-763

50. Mazmanian, S.K., *et al.* (2000) *Staphylococcus aureus* sortase mutants defective in the display of surface proteins and in the pathogenesis of animal infections. *Proc Natl Acad Sci U S A* 97, 5510-5515

51. Mazmanian, S.K., *et al.* (2001) Sortase-catalysed anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. *Mol Microbiol* 40, 1049-1057

52. Ilangovan, U., *et al.* (2001) Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* 98, 6056-6061

53. Marraffini, L.A., *et al.* (2006) Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. *Microbiol Mol Biol Rev* 70, 192-221

54. Pallen, M.J., *et al.* (2003) Genomic analysis of secretion systems. *Curr Opin Microbiol* 6, 519-527

55. Pallen, M.J., *et al.* (2001) An embarrassment of sortases - a richness of substrates? *Trends Microbiol* 9, 97-102

56. Comfort, D., and Clubb, R.T. (2004) A comparative genome analysis identifies distinct sorting pathways in gram-positive bacteria. *Infect Immun* 72, 2710-2722

57. Dramsi, S., *et al.* (2005) Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria. *Res Microbiol* 156, 289-297

58. Bierne, H., *et al.* (2004) Sortase B, a new class of sortase in *Listeria monocytogenes*. *J Bacteriol* 186, 1972-1982

59. Andrade, M.A., *et al.* (2002) NEAT: a domain duplicated in genes near the components of a putative Fe3+ siderophore transporter from Gram-positive pathogenic bacteria. *Genome Biol* 3, RESEARCH0047

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

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

62. Scott, J.R., and Zahner, D. (2006) Pili with strong attachments: Gram-positive bacteria do it differently. *Mol Microbiol* 62, 320-330

63. Telford, J.L., *et al.* (2006) Pili in gram-positive pathogens. *Nat Rev Microbiol* 4, 509-519

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

65. Yeung, M.K., *et al.* (1998) Identification of a gene involved in assembly of *Actinomyces naeslundii* T14V type 2 fimbriae. *Infect Immun* 66, 1482-1491

66. Elliot, M.A., *et al.* (2003) The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in *Streptomyces coelicolor*. *Genes Dev* 17, 1727-1740

67. Marraffini, L.A., and Schneewind, O. (2006) Targeting proteins to the cell wall of sporulating *Bacillus anthracis*. *Mol Microbiol* 62, 1402-1417

68. Paterson, G.K., and Mitchell, T.J. (2004) The biology of Gram-positive sortase enzymes. *Trends Microbiol* 12, 89-95

69. Schneewind, O., *et al.* (1992) Sorting of protein A to the staphylococcal cell wall. *Cell* 70, 267-281

70. Uhlen, M., *et al.* (1984) Complete sequence of the staphylococcal gene encoding protein A. A gene evolved through multiple duplications. *J Biol Chem* 259, 1695-1702

71. Schneewind, O., *et al.* (1993) Cell wall sorting signals in surface proteins of gram-positive bacteria. *EMBO J* 12, 4803-4811

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

73. Pritz, S., *et al.* (2007) Synthesis of biologically active peptide nucleic acid-peptide conjugates by sortase-mediated ligation. *J Org Chem* 72, 3909-3912

74. Ilangovan, U., *et al.* (2001) Assignment of the 1H, 13C and 15N signals of Sortase. *J Biomol NMR* 19, 379-380

75. Zong, Y., *et al.* (2004) Crystal structures of *Staphylococcus aureus* sortase A and its substrate complex. *J Biol Chem* 279, 31383-31389

76. Liew, C.K., *et al.* (2004) Localization and mutagenesis of the sorting signal binding site on sortase A from *Staphylococcus aureus*. *FEBS Lett* 571, 221-226

77. Naik, M.T., *et al.* (2006) *Staphylococcus aureus* Sortase A transpeptidase. Calcium promotes sorting signal binding by altering the mobility and structure of an active site loop. *J Biol Chem* 281, 1817-1826

78. Lu, C., *et al.* (2007) *Staphylococcus aureus* sortase A exists as a dimeric protein in vitro. *Biochemistry* 46, 9346-9354

79. Zhu, J., *et al.* (2008) Single mutation on the surface of *Staphylococcus aureus* Sortase A can disrupt its dimerization. *Biochemistry* 47, 1667-1674

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

81. Mazmanian, S.K., *et al.* (2002) An iron-regulated sortase anchors a class of surface protein during *Staphylococcus aureus* pathogenesis. *Proc Natl Acad Sci U S A* 99, 2293-2298

82. Skaar, E.P., *et al.* (2006) *Bacillus anthracis* IsdG, a heme-degrading monooxygenase. *J Bacteriol* 188, 1071-1080

83. Torres, V.J., *et al.* (2006) *Staphylococcus aureus* IsdB is a hemoglobin receptor required for heme iron utilization. *J Bacteriol* 188, 8421-8429

84. Escolar, L., *et al.* (1999) Opening the iron box: transcriptional metalloregulation by the Fur protein. *J Bacteriol* 181, 6223-6229

85. Horsburgh, M.J., *et al.* (2001) In *Staphylococcus aureus*, fur is an interactive regulator with PerR, contributes to virulence, and Is necessary for oxidative stress resistance through positive regulation of catalase and iron homeostasis. *J Bacteriol* 183, 468-475

86. Mazmanian, S.K., *et al.* (2003) Passage of heme-iron across the envelope of *Staphylococcus aureus*. *Science* 299, 906-909

87. Zong, Y., *et al.* (2004) The structure of sortase B, a cysteine transpeptidase that tethers surface protein to the *Staphylococcus aureus* cell wall. *Structure* 12, 105-112

88. Zhang, R., *et al.* (2004) Structures of sortase B from *Staphylococcus aureus* and *Bacillus anthracis* reveal catalytic amino acid triad in the active site. *Structure* 12, 1147-1156

89. Brook, I. (2002) Microbiology of polymicrobial abscesses and implications for therapy. *J Antimicrob Chemother* 50, 805-810

90. Roche, F.M., *et al.* (2003) Characterization of novel LPXTG-containing proteins of *Staphylococcus aureus* identified from genome sequences. *Microbiology* 149, 643-654

91. Dziewanowska, K., *et al.* (1999) Fibronectin binding protein and host cell tyrosine kinase are required for internalization of *Staphylococcus aureus* by epithelial cells. *Infect Immun* 67, 4673-4678

92. Patti, J.M., *et al.* (1992) Molecular characterization and expression of a gene encoding a *Staphylococcus aureus* collagen adhesin. *J Biol Chem* 267, 4766-4772

93. Switalski, L.M., *et al.* (1993) Collagen mediates adhesion of *Streptococcus mutans* to human dentin. *Infect Immun* 61, 4119-4125

94. Switalski, L.M., *et al.* (1993) A collagen receptor on *Staphylococcus aureus* strains isolated from patients with septic arthritis mediates adhesion to cartilage. *Mol Microbiol* 7, 99-107

95. McDevitt, D., *et al.* (1994) Molecular characterization of the clumping factor (fibrinogen receptor) of *Staphylococcus aureus*. *Mol Microbiol* 11, 237-248

96. Ni Eidhin, D., *et al.* (1998) Clumping factor B (ClfB), a new surface-located fibrinogen-binding adhesin of *Staphylococcus aureus*. *Mol Microbiol* 30, 245-257

97. Josefsson, E., *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, 31145-31152

98. Josefsson, E., *et al.* (1998) Three new members of the serine-aspartate repeat protein multigene family of *Staphylococcus aureus*. *Microbiology* 144 (Pt 12), 3387-3395

99. Hartleib, J., *et al.* (2000) Protein A is the von Willebrand factor binding protein on *Staphylococcus aureus*. *Blood* 96, 2149-2156

100. Gomez, M.I., *et al.* (2004) *Staphylococcus aureus* protein A induces airway epithelial inflammatory responses by activating TNFR1. *Nat Med* 10, 842-848

101. Gomez, M.I., *et al.* (2006) *Staphylococcus aureus* protein A activates TNFR1 signaling through conserved IgG binding domains. *J Biol Chem* 281, 20190-20196

102. Dryla, A., *et al.* (2003) Identification of a novel iron regulated staphylococcal surface protein with haptoglobin-haemoglobin binding activity. *Mol Microbiol* 49, 37-53

103. Skaar, E.P., and Schneewind, O. (2004) Iron-regulated surface determinants (Isd) of *Staphylococcus aureus*: stealing iron from heme. *Microbes Infect* 6, 390-397

104. Jonsson, I.M., *et al.* (2002) On the role of *Staphylococcus aureus* sortase and sortase-catalyzed surface protein anchoring in murine septic arthritis. *J Infect Dis* 185, 1417-1424

105. Weiss, W.J., *et al.* (2004) Effect of srtA and srtB gene expression on the virulence of *Staphylococcus aureus* in animal models of infection. *J Antimicrob Chemother* 53, 480-486

106. Bubeck Wardenburg, J., *et al.* (2007) Surface proteins and exotoxins are required for the pathogenesis of *Staphylococcus aureus* pneumonia. *Infect Immun* 75, 1040-1044

107. Stranger-Jones, Y.K., *et al.* (2006) Vaccine assembly from surface proteins of *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* 103, 16942-16947

108. Gaudreau, M.C., *et al.* (2007) Protective immune responses to a multi-gene DNA vaccine against *Staphylococcus aureus*. *Vaccine* 25, 814-824

109. Jonsson, I.M., *et al.* (2003) The role of *Staphylococcus aureus* sortase A and sortase B in murine arthritis. *Microbes Infect* 5, 775-780

110. Bierne, H., *et al.* (2002) Inactivation of the srtA gene in *Listeria monocytogenes* inhibits anchoring of surface proteins and affects virulence. *Mol Microbiol* 43, 869-881

111. Garandeau, C., *et al.* (2002) The sortase SrtA of *Listeria monocytogenes* is involved in processing of internalin and in virulence. *Infect Immun* 70, 1382-1390

112. Kharat, A.S., and Tomasz, A. (2003) Inactivation of the srtA gene affects localization of surface proteins and decreases adhesion of *Streptococcus pneumoniae* to human pharyngeal cells in vitro. *Infect Immun* 71, 2758-2765

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

114. Chen, S., *et al.* (2005) Sortase A contributes to pneumococcal nasopharyngeal colonization in the chinchilla model. *FEMS Microbiol Lett* 253, 151-154

115. Lalioui, L., *et al.* (2005) The SrtA Sortase of *Streptococcus agalactiae* is required for cell wall anchoring of proteins containing the LPXTG motif, for adhesion to epithelial cells, and for colonization of the mouse intestine. *Infect Immun* 73, 3342-3350

116. Levesque, C.M., *et al.* (2005) Involvement of sortase anchoring of cell wall proteins in biofilm formation by *Streptococcus mutans*. *Infect Immun* 73, 3773-3777

117. Sabet, C., *et al.* (2005) LPXTG protein InlJ, a newly identified internalin involved in *Listeria monocytogenes* virulence. *Infect Immun* 73, 6912-6922

118. Yamaguchi, M., *et al.* (2006) Role of *Streptococcus sanguinis* sortase A in bacterial colonization. *Microbes Infect* 8, 2791-2796

119. Sauer, F.G., *et al.* (2000) Bacterial pili: molecular mechanisms of pathogenesis. *Curr Opin Microbiol* 3, 65-72

120. Pichichero, M.E., *et al.* (1982) Do pili play a role in pathogenicity of *Haemophilus influenzae* type B? *Lancet* 2, 960-962

121. Oligino, L., and Fives-Taylor, P. (1993) Overexpression and purification of a fimbria-associated adhesin of *Streptococcus parasanguis*. *Infect Immun* 61, 1016-1022

122. Rudel, T., *et al.* (1995) Pilus biogenesis and epithelial cell adherence of *Neisseria gonorrhoeae* pilC double knock-out mutants. *Mol Microbiol* 17, 1057-1071

123. Soto, G.E., and Hultgren, S.J. (1999) Bacterial adhesins: common themes and variations in architecture and assembly. *J Bacteriol* 181, 1059-1071

124. Smyth, C.J., *et al.* (1996) Fimbrial adhesins: similarities and variations in structure and biogenesis. *FEMS Immunol Med Microbiol* 16, 127-139

125. Hultgren, S.J., *et al.* (1991) Chaperone-assisted assembly and molecular architecture of adhesive pili. *Annu Rev Microbiol* 45, 383-415

126. Swierczynski, A., and Ton-That, H. (2006) Type III pilus of corynebacteria: Pilus length is determined by the level of its major pilin subunit. *J Bacteriol* 188, 6318-6325

127. Mandlik, A., *et al.* (2007) Corynebacterium diphtheriae employs specific minor pilins to target human pharyngeal epithelial cells. *Mol Microbiol* 64, 111-124

128. Gaspar, A.H., and Ton-That, H. (2006) Assembly of distinct pilus structures on the surface of *Corynebacterium diphtheriae*. *J Bacteriol* 188, 1526-1533

129. Mandlik, A., *et al.* (2008) Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends Microbiol* 16, 33-40

130. Ton-That, H., *et al.* (2004) Sortases and pilin elements involved in pilus assembly of *Corynebacterium diphtheriae*. *Mol Microbiol* 53, 251-261

131. Rosini, R., *et al.* (2006) Identification of novel genomic islands coding for antigenic pilus-like structures in *Streptococcus agalactiae*. *Mol Microbiol* 61, 126-141

132. Kim, S.W., *et al.* (2002) Inhibition of the bacterial surface protein anchoring transpeptidase sortase by medicinal plants. *Biosci Biotechnol Biochem* 66, 2751-2754

133. Kim, S.H., *et al.* (2003) Inhibition of sortase, a bacterial surface protein anchoring transpeptidase, by beta-sitosterol-3-O-glucopyranoside from *Fritillaria verticillata*. *Biosci Biotechnol Biochem* 67, 2477-2479

134. Kim, S.H., *et al.* (2004) Inhibition of the bacterial surface protein anchoring transpeptidase sortase by isoquinoline alkaloids. *Biosci Biotechnol Biochem* 68, 421-424

135. Park, B.S., *et al.* (2005) *Curcuma longa L.* constituents inhibit sortase A and Staphylococcus aureus cell adhesion to fibronectin. *J Agric Food Chem* 53, 9005-9009

136. Oh, K.B., *et al.* (2005) Bis(indole) alkaloids as sortase A inhibitors from the sponge *Spongosorites sp. Bioorg Med Chem Lett* 15, 4927-4931

137. Jang, K.H., *et al.* (2007) Aaptamines as sortase A inhibitors from the tropical sponge *Aaptos aaptos*. *Bioorg Med Chem Lett* 17, 5366-5369

138. Chang, Y.H., *et al.* (2008) Dihydroxystyrene metabolites from an association of the sponges *Poecillastra wondoensis* and *Jaspis sp. J Nat Prod* 71, 779-783

139. Kang, S.S., *et al.* (2006) Flavonols inhibit sortases and sortase-mediated *Staphylococcus aureus* clumping to fibrinogen. *Biol Pharm Bull* 29, 1751-1755

140. Lee, H.S., *et al.* (2005) Cyclic peptides of the nocardamine class from a marine-derived bacterium of the genus Streptomyces. *J Nat Prod* 68, 623-625

141. Scott, C.J., *et al.* (2002) Irreversible inhibition of the bacterial cysteine protease-transpeptidase sortase (SrtA) by substrate-derived affinity labels. *Biochem J* 366, 953-958

142. Frankel, B.A., *et al.* (2004) Vinyl sulfones: inhibitors of SrtA, a transpeptidase required for cell wall protein anchoring and virulence in *Staphylococcus aureus*. *J Am Chem Soc* 126, 3404-3405

143. Kruger, R.G., *et al.* (2004) Inhibition of the *Staphylococcus aureus* sortase transpeptidase SrtA by phosphinic peptidomimetics. *Bioorg Med Chem* 12, 3723-3729

144. Jung, M.E., *et al.* (2005) Synthesis of (2R,3S) 3-amino-4-mercapto-2-butanol, a threonine analogue for covalent inhibition of sortases. *Bioorg Med Chem Lett* 15, 5076-5079

145. Chenna, B.C., *et al.* (2008) Identification of novel inhibitors of bacterial surface enzyme *Staphylococcus aureus* Sortase A. *Bioorg Med Chem Lett* 18, 380-385

146. Oh, K.B., *et al.* (2004) Discovery of diarylacrylonitriles as a novel series of small molecule sortase A inhibitors. *J Med Chem* 47, 2418-2421

147. Maresso, A.W., *et al.* (2007) Activation of inhibitors by sortase triggers irreversible modification of the active site. *J Biol Chem* 282, 23129-23139

148. Mao, H., *et al.* (2004) Sortase-mediated protein ligation: a new method for protein engineering. *J Am Chem Soc* 126, 2670-2671

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

150. Nguyen, H.D., and Schumann, W. (2006) Establishment of an experimental system allowing immobilization of proteins on the surface of *Bacillus subtilis* cells. *J Biotechnol* 122, 473-482

151. Parthasarathy, R., *et al.* (2007) Sortase A as a novel molecular "stapler" for sequence-specific protein conjugation. *Bioconjug Chem* 18, 469-476

152. Chan, L., *et al.* (2007) Covalent attachment of proteins to solid supports and surfaces via Sortase-mediated ligation. *PLoS ONE* 2, e1164

153. Clow, F., *et al.* (2008) Immobilization of proteins to biacore sensor chips using *Staphylococcus aureus* sortase A. *Biotechnol Lett*

154. Popp, M.W., *et al.* (2007) Sortagging: a versatile method for protein labeling. *Nat Chem Biol* 3, 707-708

155. Tanaka, T., *et al.* (2008) Site-specific protein modification on living cells catalyzed by Sortase. *Chembiochem* 9, 802-807

156. Ton-That, H., *et al.* (1999) Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc Natl Acad Sci U S A* 96, 12424-12429

157. Carmona, A.K., *et al.* (2006) A continuous fluorescence resonance energy transfer angiotensin I-converting enzyme assay. *Nat Protoc* 1, 1971-1976

158. Ton-That, H., *et al.* (2000) Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. Sortase catalyzed in vitro transpeptidation reaction using LPXTG peptide and NH(2)-Gly(3) substrates. *J Biol Chem* 275, 9876-9881

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

160. Jeffery, C.J. (1999) Moonlighting proteins. *Trends Biochem Sci* 24, 8-11

161. O'Brien, P.J., and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities. *Chem Biol* 6, R91-R105

162. Schoemaker, H.E., *et al.* (2003) Dispelling the myths--biocatalysis in industrial synthesis. *Science* 299, 1694-1697

163. Straathof, A.J., *et al.* (2002) The production of fine chemicals by biotransformations. *Curr Opin Biotechnol* 13, 548-556

164. Mingeot-Leclercq, M.P., *et al.* (1999) Aminoglycosides: activity and resistance. *Antimicrob Agents Chemother* 43, 727-737

165. Vakulenko, S.B., and Mobashery, S. (2003) Versatility of aminoglycosides and prospects for their future. *Clin Microbiol Rev* 16, 430-450

166. Pratt, M.R., and Bertozzi, C.R. (2005) Synthetic glycopeptides and glycoproteins as tools for biology. *Chem Soc Rev* 34, 58-68

167. Griffith, B.R., *et al.* (2005) 'Sweetening' natural products via glycorandomization. *Curr Opin Biotechnol* 16, 622-630

168. Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* 3, 97-130

169. Doores, K.J., *et al.* (2006) Exploring and exploiting the therapeutic potential of glycoconjugates. *Chemistry* 12, 656-665

170. Mills, N.L., *et al.* (2006) An alpha-helical peptidomimetic inhibitor of the HIV-1 Rev-RRE interaction. *J Am Chem Soc* 128, 3496-3497

171. Rosen, C.A., and Pavlakis, G.N. (1990) Tat and Rev: positive regulators of HIV gene expression. *AIDS* 4, 499-509

172. Cullen, B.R. (1992) Mechanism of action of regulatory proteins encoded by complex retroviruses. *Microbiol Rev* 56, 375-394

173. Pollard, V.W., and Malim, M.H. (1998) The HIV-1 Rev protein. *Annu Rev Microbiol* 52, 491-532

174. Malim, M.H., *et al.* (1990) HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. *Cell* 60, 675-683

175. Holland, S.M., *et al.* (1990) Human immunodeficiency virus rev protein recognizes a target sequence in rev-responsive element RNA within the context of RNA secondary structure. *J Virol* 64, 5966-5975

176. Heaphy, S., *et al.* (1990) HIV-1 regulator of virion expression (Rev) protein binds to an RNA stem-loop structure located within the Rev response element region. *Cell* 60, 685-693

177. Dayton, E.T., *et al.* (1992) Extensive sequence-specific information throughout the CAR/RRE, the target sequence of the human immunodeficiency virus type 1 Rev protein. *J Virol* 66, 1139-1151

178. Hammarskjold, M.L. (1997) Regulation of retroviral RNA export. *Semin Cell Dev Biol* 8, 83-90

179. Malim, M.H., and Cullen, B.R. (1993) Rev and the fate of pre-mRNA in the nucleus: implications for the regulation of RNA processing in eukaryotes. *Mol Cell Biol* 13, 6180-6189

180. Hadzopoulou-Cladaras, M., *et al.* (1989) The rev (trs/art) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a cis-acting sequence in the env region. *J Virol* 63, 1265-1274

181. Frankel, A.D. (1992) Activation of HIV transcription by Tat. *Curr Opin Genet Dev* 2, 293-298

182. Churcher, M.J., *et al.* (1993) High affinity binding of TAR RNA by the human immunodeficiency virus type-1 tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. *J Mol Biol* 230, 90-110

183. Delling, U., *et al.* (1992) Conserved nucleotides in the TAR RNA stem of human immunodeficiency virus type 1 are critical for Tat binding and trans activation: model for TAR RNA tertiary structure. *J Virol* 66, 3018-3025

184. Dingwall, C., *et al.* (1989) Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro. *Proc Natl Acad Sci U S A* 86, 6925-6929

185. Aboul-ela, F., *et al.* (1995) The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J Mol Biol* 253, 313-332

186. Weeks, K.M., and Crothers, D.M. (1991) RNA recognition by Tat-derived peptides: interaction in the major groove? *Cell* 66, 577-588

187. Jones, K.A., and Peterlin, B.M. (1994) Control of RNA initiation and elongation at the HIV-1 promoter. *Annu Rev Biochem* 63, 717-743

188. Zapp, M.L., *et al.* (1993) Small molecules that selectively block RNA binding of HIV-1 Rev protein inhibit Rev function and viral production. *Cell* 74, 969-978

189. Lacourciere, K.A., *et al.* (2000) Mechanism of neomycin and Rev peptide binding to the Rev responsive element of HIV-1 as determined by fluorescence and NMR spectroscopy. *Biochemistry* 39, 5630-5641

190. Werstuck, G., *et al.* (1996) A non-canonical base pair within the human immunodeficiency virus rev-responsive element is involved in both rev and small molecule recognition. *Chem Biol* 3, 129-137

191. Wang, S., *et al.* (1998) Binding of neomycin to the TAR element of HIV-1 RNA induces dissociation of Tat protein by an allosteric mechanism. *Biochemistry* 37, 5549-5557

192. Mei, H.Y., *et al.* (1997) Discovery of selective, small-molecule inhibitors of RNA complexes--I. The Tat protein/TAR RNA complexes required for HIV-1 transcription. *Bioorg Med Chem* 5, 1173-1184

193. Moazed, D., and Noller, H.F. (1987) Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature* 327, 389-394

194. von Ahsen, U., *et al.* (1991) Antibiotic inhibition of group I ribozyme function. *Nature* 353, 368-370

195. Stage, T.K., *et al.* (1995) Inhibition of the hammerhead ribozyme by neomycin. *RNA* 1, 95-101

196. Rogers, J., *et al.* (1996) Inhibition of the self-cleavage reaction of the human hepatitis delta virus ribozyme by antibiotics. *J Mol Biol* 259, 916-925

197. Mikkelsen, N.E., *et al.* (1999) Inhibition of RNase P RNA cleavage by aminoglycosides. *Proc Natl Acad Sci U S A* 96, 6155-6160

198. Kirk, S.R., and Tor, Y. (1999) tRNA(Phe) binds aminoglycoside antibiotics. *Bioorg Med Chem* 7, 1979-1991

199. Luedtke, N.W., *et al.* (2003) RNA-ligand interactions: affinity and specificity of aminoglycoside dimers and acridine conjugates to the HIV-1 Rev response element. *Biochemistry* 42, 11391-11403

200. Park, W. (1996) Rapid Combinatorial Synthesis of Aminoglycoside Antibiotic Mimetics: Use of a Polyethylene Glycol-Linked Amine and a Neamine-Derived Aldehyde in Multiple Component Condensation as a Strategy for the Discovery of New Inhibitors of the HIV RNA Rev Responsive Element. *J Am Chem Soc* 118, 10150-10155

201. Litovchick, A., *et al.* (1999) Arginine-aminoglycoside conjugates that bind to HIV transactivation responsive element RNA in vitro. *FEBS Lett* 445, 73-79

202. Litovchick, A., *et al.* (2000) Aminoglycoside-arginine conjugates that bind TAR RNA: synthesis, characterization, and antiviral activity. *Biochemistry* 39, 2838-2852

203. Cabrera, C., *et al.* (2000) Anti-human immunodeficiency virus activity of novel aminoglycoside-arginine conjugates at early stages of infection. *AIDS Res Hum Retroviruses* 16, 627-634

204. Litovchick, A., *et al.* (2001) Neomycin B-arginine conjugate, a novel HIV-1 Tat antagonist: synthesis and anti-HIV activities. *Biochemistry* 40, 15612-15623

205. Borkow, G., *et al.* (2003) Structure-activity relationship of neomycin, paromomycin, and neamine-arginine conjugates, targeting HIV-1 gp120-CXCR4 binding step. *Antiviral Res* 60, 181-192

206. Riguet, E., *et al.* (2004) A peptide nucleic acid-neamine conjugate that targets and cleaves HIV-1 TAR RNA inhibits viral replication. *J Med Chem* 47, 4806-4809

207. Lee, J., *et al.* (2004) An approach to enhance specificity against RNA targets using heteroconjugates of aminoglycosides and chloramphenicol (or linezolid). *J Am Chem Soc* 126, 1956-1957

208. Hamasaki, K., and Ueno, A. (2001) Aminoglycoside antibiotics, neamine and its derivatives as potent inhibitors for the RNA-protein interactions derived from HIV-1 activators. *Bioorg Med Chem Lett* 11, 591-594

209. Kumagai, I., *et al.* (2000) Construction of HIV Rev peptides containing peptide nucleic acid that bind HIV RRE IIB RNA. *Bioorg Med Chem Lett* 10, 377-379

210. Takahashi, T., *et al.* (1999) Design of artificial pepetides that recognize the HIV RRE IIB RNA. *Nucleic Acids Symp Ser*, 271-272

211. Kjems, J., *et al.* (1992) Specific binding of a basic peptide from HIV-1 Rev. *EMBO J* 11, 1119-1129

212. Zhang, Q., *et al.* (2001) Structural characterization of the complex of the Rev response element RNA with a selected peptide. *Chem Biol* 8, 511-520

213. Chaloin, L., *et al.* (2007) Potent inhibition of HIV-1 replication by backbone cyclic peptides bearing the Rev arginine rich motif. *J Biomed Sci* 14, 565-584

214. Choudhury, I., *et al.* (1998) Inhibition of HIV-1 replication by a Tat RNA-binding domain peptide analog. *J Acquir Immune Defic Syndr Hum Retrovirol* 17, 104-111

215. Friedler, A., *et al.* (2000) Development of a functional backbone cyclic mimetic of the HIV-1 Tat arginine-rich motif. *J Biol Chem* 275, 23783-23789

216. Rigl, C.T., *et al.* (1997) Structural RNA mimetics: N3'-->P5' phosphoramidate DNA analogs of HIV-1 RRE and TAR RNA form A-type helices that bind specifically to Rev and Tat-related peptides. *Biochemistry* 36, 650-659

217. Moehle, K., *et al.* (2007) Design of beta-hairpin peptidomimetics that inhibit binding of alpha-helical HIV-1 Rev protein to the rev response element RNA. *Angew Chem Int Ed Engl* 46, 9101-9104

Publications

J|A|C|S
COMMUNICATIONS

Published on Web 01/30/2008

Peptide–Sugar Ligation Catalyzed by Transpeptidase Sortase: A Facile Approach to Neoglycoconjugate Synthesis

Sharmishtha Samantaray, Uttara Marathe, Sayani Dasgupta, Vinay K. Nandicoori, and Rajendra P. Roy*

National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India

Received September 23, 2007; E-mail: rproy@nii.res.in

Development of new methods for linking sugars to peptides or proteins is an active area of research because natural glycopeptides or neoglycoconjugates play important roles in biology and medicine and are indispensable tools for probing several biological processes.[1–4] However, despite dramatic progress in synthetic carbohydrate and protein chemistry in recent years, glycoconjugate synthesis involving sugar and polypeptide remains a formidable task. This is principally because synthetic protocols are quite demanding and involve multiple reaction steps with requirements of rather extensive protection of reactive functionalities. The problem may be in part or completely obviated through the intermediary of enzymes. Indeed, glycosidases and glycosyl transferases, in appropriate situations, have made the synthesis of oligosaccharides much simpler.[5–7] Given the current ease with which peptides are assembled by solid phase methodology and proteins obtained from expression systems, the availability of enzymes capable of covalently linking a presynthesized sugar and a polypeptide would greatly facilitate the convergent semisynthesis of glycoconjugates with exquisite biological properties. We report here a novel enzymatic approach, using an unprecedented sortase-catalyzed transamidation reaction, for the facile one-pot synthesis of glycoconjugates comprising amino sugars and native polypeptides.

The transpeptidase sortase, present in the cell envelope of most gram-positive bacteria, catalyzes the covalent anchoring of several bacterial surface proteins to the peptidoglycan cross-bridges of the cell wall.[8,9] Sortase A of *Staphylococcus aureus* recognizes a LPXTG pentapeptide sequence motif located near the C-terminus of the target proteins, cleaves at Thr–Gly peptide bond, and catalyzes the formation of a new peptide bond between the threonyl carboxyl and amino group of the peptidoglycan pentaglycine cross-bridges.[10] The transpeptidation reaction proceeds in two steps without the aid of any extraneous molecule; the active site cysteine residue first attacks the target LPXTG substrate forming an acyl–enzyme intermediate which in the second step is resolved by the nucleophilic attack of the amino group of the terminal Gly residue of the peptidoglycan. In the absence of a suitable amino nucleophile, the LPXTG peptide substrate is slowly hydrolyzed. Sortase-mediated transpeptidation reaction involving LPXTG and aminoglycine containing polypeptides proceeds smoothly in vitro and has been applied to the synthesis of protein–peptide and peptide–nucleic acid conjugates that would have been difficult to obtain by purely chemical or genetic means.[11–13] Interestingly, ligation of LPXTG substrates can occur even with polypeptides containing a single Gly residue at the amino terminus. This observation of relaxed specificity for the amine nucleophile and the fact that the transpeptidation reaction does not require intermediary of high-energy phosphates prompted us to explore sortase-mediated ligation of polypeptides to amino sugars with a view to develop an enzymatic approach to glycoconjugate synthesis.

We considered 6-aminohexoses as potential sugar substrates with the idea that the $-CH_2-NH_2$ moiety present in these sugars might mimic some elements of the glycine structure. Accordingly, we tested the potential of sortase to ligate 6-deoxy-6-aminoglucose and 6-deoxy-6-aminomannose to a model YALPETGK peptide substrate. HPLC assays followed by MALDI-TOF or ESMS analyses revealed the formation of respective YALPET–sugar adducts, suggesting that the above amino sugars indeed acted as nucleophiles in the transamidation reaction (Supporting Information Figure 1). In contrast, the substrate peptide was hydrolyzed to YALPET without the formation of the YALPET–sugar adduct when glucosamine was used as a substrate. Consistent with the known tolerance of sortase for the LPXTG recognition motif,[14] the YALPMTGK peptide sequence also reacted with 6-deoxy-6-aminoglucose or 6-deoxy-6-aminomannose but not with glucosamine.

To further probe the specificity requirements as well as to see if 6-aminohexoses can serve as recognition tags for peptide–sugar ligations, we investigated the ability of sortase to ligate peptides to an aminoglycoside class of therapeutically important antibiotics. These antibiotics are built up by a variety of amino sugars of the 6-amino or the 2,6-diamino type besides containing several other amino functionalities[15] (Figure 1). The central scaffold of aminoglycoside antibiotics is the 2-deoxystreptamine ring to which amino sugars are substituted at positions 4 and 6 (as in tobramycin and kanamycins) or 4 and 5 (as in ribostamycin, neomycin, and paromomycin). Sortase-mediated ligation of model peptide substrates to aminoglycoside antibiotics proceeded smoothly. Analyses of the reaction products by reversed phase HPLC (Figure 2) followed by MALDI (Supporting Information Table 2) revealed the formation of specific conjugates between antibiotics and peptides in the yields varying from 35 to 70% for the kanamycin class, and about 18–30% for the ribostamycin class of antibiotics. Electrospray mass spectrometry (Supporting Information Figures 2–5) of the respective conjugates produced fragmentations that unambiguously showed occurrence of peptide ligation exclusively at a single 6-amino site in ring A of kanamycins, tobramycin, and ribostamycin or ring D of paromomycin and neomycin. Thus, conjugation of peptide substrates was limited to the 6-amino site in the antibiotics despite the presence of a plethora of amino groups, indicating rather strict specificity and selectivity for the sugar amino groups by sortase.

Next, we explored conjugation of biologically relevant peptides to aminoglycoside antibiotics. Toward this, we considered peptide sequences derived from or based on Tat and Rev proteins of HIV because these proteins play important roles in virus replication through their interactions with structured viral RNA target sites, TAR in the case of Tat and RRE in the case of Rev.[16] Interestingly, aminoglycoside antibiotics and analogues,[17] as well as short

10.1021/ja077358g CCC: $40.75 © 2008 American Chemical Society

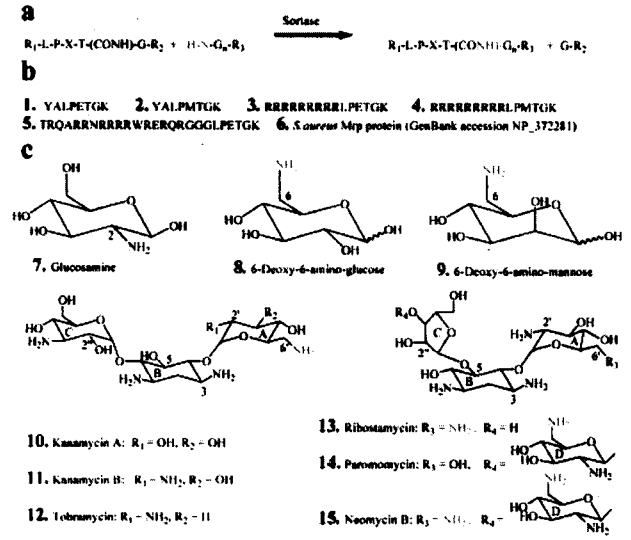


Figure 1. (a) General peptide ligation reaction catalyzed by sortase. (b) LPXTG peptide substrates used in the study. (c) Amino sugars used in the study. The 6-amino site is shown in blue.

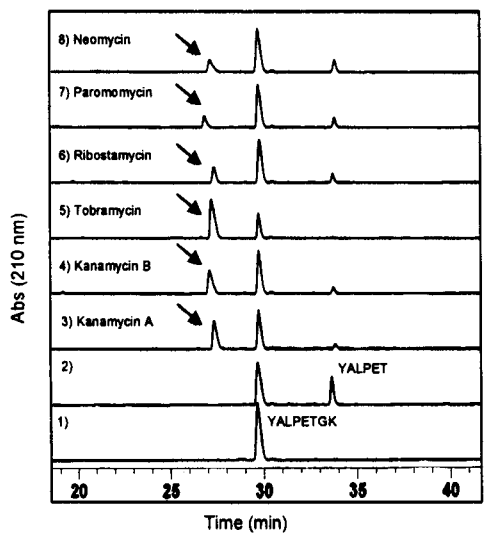


Figure 2. Sortase-catalyzed conjugation of YALPETGK model peptide to aminoglycosides. Reversed phase HPLC profiles 1 and 2 are peptide samples in the absence and presence of sortase, respectively. Chromatograms numbered 3 to 8 represent reactions of peptide with aminoglycoside antibiotics. The peptide–antibiotics conjugate is indicated by an arrow.

arginine-rich sequences derived from Tat and Rev[17,18] or even a nonaarginine peptide mimic, have been shown to interfere with Tat–TAR or Rev–RRE interactions, leading to inhibition of virus replication.[19] We prepared conjugates of nonaarginine (**4**) or a Rev sequence (**5**) with several antibiotics (Supporting Information Table 2 and Figure 6). We used the neomycin–Rev conjugate, as a test case, for evaluating the extent of RRE RNA binding. The gel retardation assays (Supporting Information Figure 7) yielded a RRE binding affinity (K_d) of 9.3 nM for the conjugate as compared to 114.3 nM for the Rev peptide, suggesting about 10-fold or more improvement of RNA binding in the conjugate. Together, these results demonstrate the utility of sortase for generating useful conjugates.

Finally, we tested the feasibility of using sortase for site-specific conjugation of sugars to proteins. For this, we expressed a protein (Mrp protein, NP_372281) from *Staphylococcus aureus* nested with a LPNTG sequence motif in its carboxy terminal region. We employed HPLC assays and MALDI-TOF to investigate the ligation of tobramycin to Mrp (Supporting Information Figures 8 and 9). Incubation of Mrp alone with sortase led to generation of two fragments expected from hydrolysis at the T–G peptide bond. However, incubation of the protein with sortase in the presence of tobramycin produced specific conjugate (Mrp–LPNT–tobramycin) with a yield of more than 45% in 6 h.

In summary, our work demonstrates that sortase can transfer peptide substrates to oligosaccharides appended with a 6-deoxy-6-aminohexose moiety in a selective manner as that of an oligoglycine sequence. Such an enzymatic activity of peptide–sugar ligation, presumably promiscuous in origin for sortase, is hitherto unknown. This robust reaction provides a simple and straightforward method for covalent ligation of a prefabricated sugar containing a 6-aminohexose tag to synthetic peptides and expressed proteins encoded with a C-terminal LPXTG sortase recognition sequence. We envision several biotechnological applications of this methodology, including generation of novel glycopeptide-based immunovaccines comprising oligosaccharides substituted with multiple peptides and glycolabeling of proteins. Besides, the facile assembly of aminoglycoside antibiotic conjugates offers tremendous possibilities of generating new RNA ligands and therapeutics. The work on anti-HIV activity of antibiotics–Tat/Rev conjugates is in progress.

Acknowledgment. We thank Gokhale Laboratory and Dr. P. Sahai for help with mass measurements. This work was supported by NII core, National Bioscience Award, and DBT-ICMR grants to R.P.R. from the Department of Biotechnology, India.

Supporting Information Available: Peptide–sugar ligations, product characterization, and RNA binding data. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Pratt, M. R.; Bertozzi, C. R. *Chem. Soc. Rev.* **2005**, *34*, 58–68.
(2) Griffith, B. R.; Langenhan, J. M.; Thorson, J. S. *Curr. Opin. Biotechnol.* **2005**, *16*, 622–630.
(3) Varki, A. *Glycobiology* **1993**, *3*, 97–130.
(4) Doores, K. J.; Gamblin, D. P.; Davis, B. G. *Chemistry* **2006**, *12*, 656–665.
(5) Bennett, C. S.; Wong, C.-H. *Chem. Soc. Rev.* **2007**, *36*, 1227–1238.
(6) Blanchard, S.; Thorson, J. S. *Curr. Opin. Chem. Biol.* **2006**, *10*, 263–271.
(7) Li, B.; Zeng, Y.; Hauser, S.; Song, H.; Wang, L. X. *J. Am. Chem. Soc.* **2005**, *127*, 9692–9693.
(8) Mazmanian, S. K.; Liu, G.; Ton-That, H.; Schneewind, O. *Science* **1999**, *285*, 760–763.
(9) Perry, A. M.; Ton-That, H.; Mazmanian, S. K.; Schneewind, O. *J. Biol. Chem.* **2002**, *277*, 16241–16248.
(10) Marraffini, L. A.; Dedent, A. C.; Schneewind, O. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 192–221.
(11) Mao, H.; Hart, S. A.; Schink, A.; Pollok, B. A. *J. Am. Chem. Soc.* **2004**, *126*, 2670–2671.
(12) Pritz, S.; Wolf, Y.; Kraetke, O.; Klose, J.; Bienert, M.; Beyermann, M. *J. Org. Chem.* **2007**, *72*, 3909–3912.
(13) Parthasarathy, R.; Subramanian, S.; Boder, E. T. *Bioconjugate Chem.* **2007**, *18*, 469–476.
(14) Huang, X.; Aulabaugh, A.; Ding, W.; Kapoor, B.; Alksne, L.; Tabei, K.; Ellestad, G. *Biochemistry* **2003**, *42*, 11307–11315.
(15) Busscher, G. F.; Rutjes, F. P.; van Delft, F. L. *Chem. Rev.* **2005**, *105*, 775–791.
(16) Zapp, M. L.; Stern, S.; Green, M. R. *Cell* **1993**, *74*, 969–978.
(17) Litovchick, A.; Lapidot, A.; Eisenstein, M.; Kalinkovich, A.; Borkow, G. *Biochemistry* **2001**, *40*, 15612–15623.
(18) Calnan, B. J.; Tidor, B.; Biancalana, S.; Hudson, D.; Frankel, A. D. *Science* **1991**, *252*, 1167–1171.
(19) Luedtke, N. W.; Baker, T. J.; Goodman, M.; Tor, Y. *J. Am. Chem. Soc.* **2000**, *122*, 12035–12036.

JA077358G