Published in final edited form as:
Structure. 2004 July ; 12(7): 1147–1156. doi:10.1016/j.str.2004.06.001.

Structures of Sortase B from *Staphylococcus aureus* and *Bacillus anthracis* Reveal Catalytic Amino Acid Triad in the Active Site

Rongguang Zhang[1,5], **Ruiying Wu**[1,5], **Grazyna Joachimiak**[1,5], **Sarkis K. Mazmanian**[2,3], **Dominique M. Missiakas**[2,4], **Piotr Gornicki**[3], **Olaf Schneewind**[2,3], and **Andrzej Joachimiak**[1,*]

[1]Structural Biology Center and Midwest Center for Structural Genomics, Argonne National Laboratory, 9700 South Cass Avenue, Building 202, Argonne, Illinois 60439

[2]Committee on Microbiology, University of Chicago, 920 East 58th Street, Chicago, Illinois 60637

[3]Department of Molecular Genetics and Cell Biology, University of Chicago, 920 East 58th Street, Chicago, Illinois 60637

[4]Department of Biochemistry and Molecular Biology, University of Chicago, 920 East 58th Street, Chicago, Illinois 60637

Summary

Surface proteins attached by sortases to the cell wall envelope of bacterial pathogens play important roles during infection. Sorting and attachment of these proteins is directed by C-terminal signals. Sortase B of *S. aureus* recognizes a motif NPQTN, cleaves the polypeptide after the Thr residue, and attaches the protein to pentaglycine cross-bridges. Sortase B of *B. anthracis* is thought to recognize the NPKTG motif, and attaches surface proteins to m-diaminopimelic acid cross-bridges. We have determined crystal structure of sortase B from *B. anthracis* and *S. aureus* at 1.6 and 2.0 Å resolutions, respectively. These structures show a β-barrel fold with α-helical elements on its outside, a structure thus far exclusive to the sortase family. A putative active site located on the edge of the β-barrel is comprised of a Cys-His-Asp catalytic triad and presumably faces the bacterial cell surface. A putative binding site for the sorting signal is located nearby.

Introduction

Surface proteins of Gram-positive bacteria perform critical biological functions that are required for the colonization of host tissues, the evasion of immune defenses, and the acquisition of nutrients; they promote bacterial adhesion to specific tissues and invasion of host cells; and provide resistance to phagocytic killing (Foster and Höök, 1998; Navarre et al., 1999). Many surface proteins are anchored to the cell wall by sortases via a transpeptidation reaction requiring a C-terminal peptide sorting signal (Schneewind et al., 1992, 1993; Cossart and Jonquieres, 2000; Mazmanian et al., 1999; Mazmanian and Schneewind, 2004). Two classes of sortases can be distinguished based on the localization of the catalytic domain (Ilangovan et al., 2001). Class I enzymes harbor an N-terminal segment of hydrophobic amino acids which functions as a signal peptide for secretion and a stop-transfer signal for membrane anchoring. Class II enzymes contain a C-terminal segment of hydrophobic amino acids which

*Corresponence: andrzejj@anl.gov.
[5]These authors contributed equally to this work.

Accession Numbers
The atomic coordinates have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 1NG5 and 1RZ2).

serves as a membrane anchor, in addition to an N-terminal signal peptide. Class I sortases adopt a type 2 transmembrane topology (N terminus inside, C terminus outside the cytoplasm); class II enzymes are thought to represent type 1 transmembrane proteins (N terminus outside, C terminus inside). *S. aureus* and *B. anthracis* sortases are anchored to the cell membrane via their N-terminal hydrophobic leader sequence, and they appear to belong to class I transmembrane proteins, and therefore their C terminus containing the catalytic domain is located outside the cell.

During sorting process, the C-terminal peptide signal is removed and the mature protein is covalently attached to a specific component of the cell wall (Schneewind et al., 1995; Navarre and Schneewind, 1994; Ton-That et al., 1999; Navarre et al., 1998; Dhar et al., 2000). In *S. aureus*, the cleavage occurs after a Thr residue within a conserved 5 amino acid motif, either LPXTG (X, any amino acid) or NPQTN, and the Thr carboxyl group is amide linked to the amino group of pentaglycine cross-bridges of the staphylococcal cell wall peptidoglycan (Ton-That et al., 1997). It has been proposed that lipid II [undecaprenyl pyrophosphate-MurNAc(-L-Ala-D-iGln-L-Lys(NH_2-Gly_5)-D-Ala-D-Ala)-(β)1-4-GlcNAc], a membrane-anchored precursor of cell wall synthesis (Higashi et al., 1967, 1970), serves as the substrate for the protein anchoring (Perry et al., 2002; Ruzin et al., 2002). Surface proteins tethered to lipid II are subsequently incorporated into the peptidoglycan by means of the transglycosylation and transpeptidation reactions of bacterial cell wall synthesis (Strominger et al., 1967; Ton-That et al., 1998, 1999). Genes encoding sortases have been identified in *B. anthracis*, and these transpeptidases are thought to cleave surface protein sorting signals with LPXTG or NPKTG motifs (Read et al., 2003; Pallen et al., 2001; Mazmanian and Schneewind, 2004). Because the cell wall peptidoglycan of *B. anthracis* is synthesized from precursors with different lipid II structure [undecaprenyl pyrophosphate-MurNAc(-L-Ala-D-iGln-(NH_2)-*m*-Dpm-D-Ala-D-Ala)-(β)1-4-GlcNAc], surface proteins are thought to be attached to the m-diaminopimelic acid cross-bridges within the cell wall (Dhar et al., 2000; Schleifer and Kandler, 1972).

Several sortase genes have been characterized in Gram-positive bacteria (Mazmanian et al., 2000; Bolken et al. 2001; Lee and Boran, 2003; Barnett and Scott, 2002; Ton-That and Schneewind, 2003; Kharat and Tomasz, 2003; Hava et al., 2003; Bierne et al., 2002; Garandeau et al., 2002). *S. aureus* sortases, A (Sa-SrtA) and B (Sa-SrtB), show only low sequence similarity to each other (Mazmanian et al., 2002). Sa-SrtA, a calcium-dependent cysteine transpeptidase specific for LPXTG motif-containing proteins, is the most extensively studied sortase so far. The structure of its C-terminal fragment was determined by NMR (Ilangovan et al., 2001). The proposed catalytic mechanism of Sa-SrtA involves highly conserved Cys and His residues and places this enzymeinthe papain/cathepsin familyofcysteine proteases (Ton-That and Schneewind, 1999; Ton-That et al., 2000, 2002). A catalytic mechanism has been proposed, whereby a thioester acyl enzyme intermediate is formed between the thiol group of Cys in the active site of Sa-SrtA and the carboxyl group of the Thr residue at the C terminus of the mature protein (Ton-That et al., 1999). This acyl-enzyme intermediate is then resolved by nucleophilic attack of the amino group of the pentaglycine cross-bridge in the cell wall envelope (Ton-That et al., 2000). Kinetic data are consistent with a ping-pong mechanism with two distinct rate-limiting steps: formation of the acyl-enzyme intermediate in a transpeptidation reaction, followed by its hydrolysis (Huang et al., 2003). Questions whether a thiol ion pair intermediate plays a key role in the sortase-catalyzed reaction and which residues constitute the active site of Sa-SrtA remain unsolved (Ton-That et al., 2002; Conolly et al., 2003).

Sa-SrtB is required for anchoring surface proteins containing the NPQTN motif, whereas LPXTG motif-containing polypeptides are not a substrate for this trans-peptidase (Mazmanian et al., 2002). Sa-SrtB mutants are defective in anchoring the IsdC surface protein and accumulate precursor molecules with uncleaved C-terminal sorting signals (Mazmanian et al., 2002). Persistence of *S. aureus* animal infections is affected by such mutations (Jonsson et al.,

2002, 2003). In contrast to Sa-SrtA, Sa-SrtB does not require metal binding for activity (Mazmanian et al., 2002). Purified Sa-SrtB cleaves NPQTN peptides in vitro, suggesting that this enzyme functions in a manner similar to Sa-SrtA (Mazmanian et al., 2002). Other studies suggest that some property of the cell envelope component used in SrtB-mediated protein attachment is different than the property of lipid II used by SrtA, but the identity of the cell wall anchor of IsdC remains unknown (Mazmanian et al., 2003). The *srtB* gene is located in the *isd* locus of staphylococci (Mazmanian et al., 2003). Genes in this locus allow *S. aureus* to scavenge heme-iron from hemoglobin, which in turn allows bacteria to grow under otherwise iron-restrictive conditions of mammalian tissues (Mazmanian et al., 2003). *S. aureus* IsdC, a heme binding protein, is also involved in heme-iron transport (Mazmanian et al., 2003). The genome of *B. anthracis* also encompasses an *isd* operon, encoding *B. anthracis* sortase B Ba-SrtB, a protein with 41% sequence identity to Sa-SrtB, and IsdC (Mazmanian et al., 2002, 2003).

In this paper, we report high-resolution crystal structures of SrtB from *S. aureus* and *B. anthracis*. The structures, similar despite rather divergent protein sequence, reveal a sortase-unique protein fold and identify the active site which contains a Cys-His-Asp catalytic triad. SrtB is a novel member of the cysteine protease family, whose members, wide spread in pathogenic bacteria, play many critical functions now extended to anchoring surface proteins to the cell wall.

Results and Discussion

Structure of Sortase B

The *S. aureus srtB* gene encodes 244 amino acid protein and the *B. anthracis srtB* gene encodes 254 amino acid protein, including 29 and 34 amino acid N-terminal signal/membrane anchor peptides, respectively (Figure 1) (Mazmanian et al., 2000). Sa-SrtB and Ba-SrtB expressed as fusions with N-terminal His tags to aid affinity purification (not removed prior to crystallization) did not contain the signal/membrane anchor peptide portions of the ORFs (Mazmanian et al., 2002). Purified proteins were monomers (determined by size exclusion chromatography; data not shown).

Crystal structures of Ba-SrtB and Sa-SrtB determined at high resolution are very similar (rmsd for Ca atoms 1.13 Å with 41.3% sequence identity) (Figure 2).We used the structure of Sa-SrtB as a reference in this discussion, as the biochemical data for this protein are rather abundant. SrtB monomer has α/β fold that is composed of a central β-barrel decorated on the outside with several α helices (Figure 2). The core of the protein is formed by an eight-stranded β-barrel containing mainly antiparallel β strands (the numbering scheme is for Sa-SrtB [Figure 3A]): β1 ↓ (42–47), β2↑ (54–58), β3↓ (81–83), β4↑ (95–101), β7↑ (187–194), β8↓ (203–213), β6↑ (134–147), and β5↓ (126–131); and five α helices: α1 (6–23), α2 (29– 38), α3 (108–116), α4 (118–123), and α5 (160–174); and several loops and a hairpin (Figure 3A). In Ba-SrtB the loop region between strands β7 and β8 (235–240) is disordered. The only other differences between Ba-SrtB and Sa-SrtB are observed in the N-terminal region: additional seven N-terminal residues are visible in the electron density in Ba-SrtB, and there is an insertion of three residues between α1 and α2 resulting in formation of a one-turn helix. The structures of SrtA and SrtB represent a fold that is unique to the sortase family in which the β-barrel assumes a novel—"double banner" design. The fold is comprised of two structural motifs each containing one long and three short β strands (Figure 3), although each motif has different β strand order (Figure 3A). One motif is composed of strands β1↓-β2↑-β3↓-β4↑, and a second motif is composed of strands β7↑-β8↓-β6↓-β5↑. The two motifs are connected through interaction of two long parallel β strands β4 and β7 to form a pseudo 2-fold arrangement (Figure 3B). The cylinder is completed by sealing the edges between parallel β2 and β3 strands as shown in Figure 3C. The β-barrel is distorted because of the difference in length and relative position of

strands in each motif. Putative active site residues are located on the edge of the barrel (Figure 2 and Figure 3D).

The β-barrel has some interesting features. One end of the barrel is covered by two short α helices (residues 108–123), whereas the other site has several extended loops (Figure 2). Three additional α helices shield two sides of the β-barrel from solvent. Two N-terminal helices α1 and α2 shield β strands 1–3. They weakly interact with the main body of the protein and may anchor SrtB to the cell wall and position the face of the protein containing the putative active site toward the cell surface. Helix α5 shields edges of β strands 5–8.

Comparison of the Structure of Sortase A and B

SrtA and SrtB perform a very similar function in *S. aureus* and other Gram-positive bacteria (Mazmanian et al., 2002), despite their rather low sequence similarity and numerous structural differences. Sa-SrtB and Sa-SrtA sequences are only 23% identical (over 184 equivalent residues) and Ba-SrtB shows negligible identity with Sa-SrtA. Interestingly, Sa-SrtB shows some sequence similarity with a domain of cyclodextrin glucanotransferase (24% identity in a 138 amino acid overlap, PDB entry 1D7F), but no structural homology. Sortases A and B show a similar overall β-barrel core design (Figure 4). The SrtA structure obtained by NMR is composed almost exclusively of β strands and loops with only one short α helix (not present in SrtB). In contrast, SrtB is a α/β fold with two long and three short α helices and nearly equal percentage of β and α secondary structure elements. However, both proteins show very similar structure and the same connectivity of the β strands in the barrel. Several conserved residues in the SrtB family are confined to this core unit including putative catalytic resides located on the edges of the barrel (Figure 2–Figure 4, and discussion below). However, major structural differences between SrtA and SrtB exist. The N terminus (residues 5–41) is unstructured in Sa-SrtA and forms two α helices and a turn in SrtB. Sequence Thr62-Ser80 forms a turn and a loop in SrtB and the only α-helix in Sa-SrtA. This region includes highly conserved Gly79 and Ser80. Sequence Val108-Lys123 forms two short α helices in SrtB and is a loop in SrtA. Region Lys145-Asp186 comprises a long α helix (α5) and two loops in SrtB and is unstructured in SrtA. This region is in direct vicinity to the proposed active site and adopts a very different conformation in both proteins. It contains several conserved residues including Thr146, Leu167, Lys173, and Ser174. In addition, several loops and turns show significantly different conformations: Gly88-His94, Asn131-Gly134, Cys194-Arg204, and the C terminus.

Overall, more than a half of the structure of Sa-SrtB shows significant structural differences with Sa-SrtA. Some of these differences may be related to the recognition of the specific sorting signals and different peptidoglycan substrates, as well as different catalytic properties. Interestingly, in Sa-SrtA the N- and C termini are on the same face of the protein, but in SrtB (both Ba-SrtB and Sa-SrtB) they are on opposite faces. This difference may have a significant implication for the orientation of these proteins with respect to the cell wall surface. In SrtA, the putative active site may point away from the cell wall, but in SrtB it points toward the cell wall (Figure 2 and Figure 4).

The Sortase Family

Analysis of protein sequences in nonredundant Gen-Bank CDS and NCBI database using multiple iterations with PSI=BLAST revealed that the sortase family (COG3764/4509) includes over 150 proteins (E values cut off 3e-05). They are found mainly in Gram-positive bacteria but their distant sequence homologs are also found in eukaryota including humans. Sortase homologs were also found as domains in larger proteins (for example the predicted unknown protein from *S. pyogenes*, gi|4033715|). Figure 1 compares sequences of SrtB from nine gram-positive bacteria (*B. anthracis, B. cereus, B. halodurans, L. innocua, L. monocytogenes, E. faecalis, S. pyogenes, C. perfringensis*, and *S. aureus*). In general, the

sortase B family shows considerable sequence diversity with only 17 residues strictly conserved among these sequence homologs.

Comparison with other protein structures using DALI server (Holm and Sander, 1999) shows only one protein—SrtA from *S. aureus* (PDB deposit 1IJA_A)—with structural similarity to SrtB (Z score 9.6, rmsd 2.9 Å, with 120 amino acids) (less than half of SrtB residues) occupying equivalent positions. Comparisons with all other protein structures show Z scores lower than 2. We conclude that the sortases A and B structures represent a new protein fold thus far exclusive to sortase family and may represent a new superfamily of proteins with transpeptidase activity.

Active Site of SrtB Contains a Cys-His-Asp Catalytic Triad

In SrtB, the putative active site residues are located on the edge of the β-barrel (Figure 2–Figure 4). The proposed active site includes a Cys residue (Cys194 in Sa-SrtB and Cys233 in Ba-SrtB) which was shown to be critical for the Sa-SrtB activity (Mazmanian et al., 2002). The putative active site also includes His and Asp residues (His101 and Asp196 in Sa-SrtB, and His140 and Asp234 in Ba-SrtB). The arrangementofthesethree amino acids, Cys-His-Asp, is reminiscent of the catalytic triad of Ser/Cys proteases which in turn is similar to human deubiquitinating enzyme UCH-L3 (Johnston et al., 1997) (PDB entry 1UCH). In Sa-SrtB, the Asp residue forms a strong hydrogen bond with nitrogen NE2 of the His residue (2.79 and 3.06 Å in two subunits of the AU, respectively, Figure 5A). The corresponding residues of Ba-SrtB form a strong hydrogen bond (2.67 Å) as well. The interaction of the His residue with the Cys residue is weaker in both proteins. The distance between these two residues is 4.0 Å in Sa-SrtB and 4.2 Å in Ba-SrtB. A similar catalytic site witha Cys-His-Asp triad ispresent in cysteine peptidases: YopT from *Yersinia pestis* and AvrPphB from *Pseudomonas* (Shao et al., 2002). YopT cleaves posttranslationally modified Rho GTPases near their carboxyl termini releasing these proteins from the membrane. AvrPphB is a virulence protein. YopT and AvrPphB, whose cysteine proteases activity dependents on the invariant Cys-His-Asp triad, define a family of 19 roteins involved in bacterial pathogenesis (Shao et al., 2002).

The catalytic triad in Ba-SrtB and Sa-SrtB is not formed entirely by structurally equivalent residues. The Cys and His residues, both strictly conserved among members of the sortase family, assume structurally equivalent positions, although different rotamers are suggested for His residues in Ba-SrtB and Sa-SrtB. The Asp residues, on the other hand, occupy different positions in the structure (and in sequence, Cys+1 in Ba-SrtB and Cys+2 in Sa-SrtB) but they both interact with the His residue in a manner consistent with their potential role in catalysis. In the SrtB family the Cys+1 residue is always Asp or Glu and Cys+2 residue is either Asp or an amino acid with hydrogen bonding potential (Tyr, Thr, or Asn, with exception of SrtB from *C. perfringens*) which could participate in the formation of the active site (Figure 1). A Cys-His-Asn catalytic triad is found in other hydrolases, for example, in human cathepsin X (Guncar et al., 2002) (PDB entry 1EF7). If such alternative configurations of the active site are considered, the SrtB-like active site with the catalytic triad can be predicted for all its homologs. These alternative configurations of amino acid residues in the active site of the family members may reflect adjustments required to accept different sorting signals and cell wall peptidoglycan attachment sites. All three putative catalytic residues (Cys-His-Asp) are present in Sa-SrtA, but only Cys126 (rmsd 1.10 Å) and His128 (rmsd 0.69 Å) residues are in positions structurally similar with the equivalent Sa-SrtB residues. It is plausible that other residues near the active site may play an important role in catalysis and can alter catalytic and kinetic properties of sortases. The question still remains open whether a thiol intermediate is formed during catalysis. Answering these questions will require more detailed mutational, kinetic, and additional structural analysis.

Putative Substrate Binding Sites

Two grooves arranged in a V shape are present on the surface of Sa-SrtB (Figure 5B). A larger groove is adjacent to the active site Cys194/233 and is formed by residues on strands β3, β4, β7, and β8 (Figure 4C). This groove includes a number of highly conserved residues (Asn63, Asp84, Asn87, Tyr99, Thr193, and Arg204). It is lined with residues with strong hydrogen binding potential suitable for binding of the NPQTN/NPKTG motifs. The sulfhydryl group of Sa-SrtB Cys194, located at the end of this groove, was shown to participate in the cleavage reaction (after the Thr residue of the motif). Figure 5C shows the approximate position of the substrate peptide NPKTG modeled onto the structure of Ba-SrtB. A second groove is formed by strands β1 and β2, the C terminus of α2, and the loop between β2 and β3 (Figures 5B and 5C) which contain several highly conserved residues (Asn39, Trp45, Gln59, Tyr66, and Ser80) located on the protein surface. We propose that the Gly or m-diaminopimelic acid residues of lipid II are presented from this groove for the transpeptidation reaction. Cocrystallization experiments of SrtB with different peptides aiming at the characterization of the proposed interactions thus far did not yield any crystals (data not shown).

In summary, *B. anthracis* and *S. aureus* are both human pathogens. *S. aureus* is a major cause of infections in hospitals worldwide and *B. anthracis* poses a significant bioterrorism threat (Brumfitt and Hamilton-Miller, 1989; Dixon et al., 1999). In recent years, new strains have emerged with resistance to all currently available antibiotics (Tenover et al., 2001). Therefore, there is an urgent need to develop new gene targets and drugs to control these persistent pathogens. Cell surface proteins that play many important functions in pathogenesis are potential and attractive targets for medical intervention (Foster and Höök, 1998). Their functional domains are located on the outside of the plasma membrane should therefore be easily accessible to drugs. Among these proteins are sortases, which anchor other proteins to the cell surface via a transpeptidation reaction requiring a C-terminal sorting signal. Therefore sortases can thus be considered as potential targets for the development of inhibitory molecules (Cossart and Jonquieres, 2000). SrtB appears to be an excellent example of such a target. Its gene is located in the *isd* locus of staphylococci and other Gram-positive bacteria. Genes in this locus allow bacteria to scavenge heme-iron from hemoglobin, which permits bacteria to grow under otherwise iron-restrictive conditions of mammalian tissues. A successful inhibition of SrtB could lead to a deficient decoration of the cell with surface proteins which in turn could inhibit bacterial growth and help combat infections. The determination of high-resolution crystal structures of SrtB and detailed view of its putative active sites in Ba-SrtB and Sa-SrtB, and contribution of various side chains to the active site presented here will aid full exploration of the biology of this important protein and potential drug target.

Experimental Procedures

Gene Cloning, Protein Expression, and Purification—Sa-SrtB ORF was amplified by PCR from chromosomal DNA of *S. aureus* (strain Newman), using two primers with the NdeI and BamHI sites engineered at the translation start codon and immediately downstream of the translation stop codon, respectively, and cloned between the NdeI and BamHI sites of the pET15b vector (Novagen) in frame with His10 tag. The construct did not include 29 N-terminal amino acids (signal peptide) and encoded 216 amino acid long Sa-SrtB protein with a 20 residue His tag at its N terminus. Ba-SrtB ORF was amplified by PCR from chromosomal DNA of *B. anthracis* (strain Sterne), using two primers compatible with the ligation independent cloning vector pMCSG7 and fusing the ORF in frame with His6 tag. The construct did not include 36 N-terminal amino acids (signal peptide) and encoded 242 amino acid Ba-SrtB protein with a 24 residue His tag at its N terminus.

Cells were grown at 37°C in LB media in the presence of 100 μg/ml ampicillin and 30 μg/ml kanamycin at final concentration, respectively. Expression of the His-tagged fusion protein in

E. coli strain BL21(DE3) was induced with 1 mM isopropyl-α-D-thiogalactoside when the optical density at 600 nm reached 0.6. Cells were harvested after 4 hr of culture at 37°C; suspended in 50 mM phosphate buffer (pH 8.0), 300 mM NaCl, 10 mM imidazol, 10 mM β-mercaptoethanol, and 10% glycerol; and lysed by sonication. The fusion protein was purified by affinity chromatography using Ni-NTA Superflow resin (Qiagen). The protein was concentrated with simultaneous buffer exchange using Centriplus-3 (Amicon) (3 kDa cutoff). A 2 mM protein stock solution in 10 mM Tris-HCl (pH 7.4), 20 mM NaCl, and 1 mM DTT was used for crystallization. Selenomethionine-labeled Ba-SrtB and Sa-SrtB were prepared using methionine biosynthesis inhibition method as described earlier (Walsh et al., 1999). The best crystals of Sa-SrtB were obtained using vapor diffusion and hanging droplets from 2.4 M sodium sulfate and 100 mM sodium citrate (pH 5.6) at room temperature, and the best crystals of Ba-SrtB were obtained using vapor diffusion and hanging droplets from 20% PEG 3350, 0.2 M sodium chloride, and 100 mM HEPES buffer (pH 7.0), at room temperature.

Data Collection—Diffraction data were collected at 100 K at the 19ID beamline of the Structural Biology Center at the Advanced Photon Source, Argonne National Laboratory. For Sa-SrtB the three-wavelength inverse-beam MAD data set (peak: 12.6603 keV [0.9794 Å]; inflection point: 12.6620 keV [0.9793 Å]; high-energy remote: 13.0000 keV [0.95385Å]) was collected from a Se-Met-labeled protein crystal at 100 K. One crystal (0.1 × 0.2 × 0.2 mm) was used to collect all MAD data to 2.48Å resolution, with 5 s exposure/1°/frame using 200 mm crystal to the detector distance. The crystal was annealed at room temperature for 15 s prior to data collection to lower mosaicity. The total oscillation range was 150° as predicted using strategy module within HKL2000 suite (Otwinowski and Minor, 1997). The space group was $P2_12_12$ with cell dimension of a = 71.208, b = 104.367, c = 58.087, α = β = γ = 90°. All data was processed and scaled with HKL2000 (Table 1) to an R_{merge} of 9.1%, 9.2%, and 8.3% for inflection point, peak, and remote respectively. The native data were collected at 1.0332 Å wavelength to 1.97 Å resolution (R_{merge} of 11.0%) from a single crystal at the SBC 19ID beamline and were used for phase extension and model refinement.

For the Ba-SrtB the three-wavelength inverse-beam MAD data set (peak: 12.6603 keV [0.9794 Å]; inflection point: 12.6620 keV [0.9793 Å]; high-energy remote: 13.0000 keV [Otwinowski and Minor, 1997] [0.9538 Å]) was collected from a Se-Met-labeled protein crystal at 100 K. One crystal (0.1 × 0.2 × 0.2 mm) was used to collect all MAD data to 1.60 Å resolution, with 5 s exposure/1°/frame using 150 mm crystal to the detector distance. The total oscillation range was 180° as predicted using strategy module within HKL2000 suite (Otwinowski and Minor, 1997). The space group was $P2_1$ with cell dimension of a = 40.47 Å, b = 64.60 Å, c = 42.96 Å, α = 90°, β = 105.77°, γ = 90°. All data was processed and scaled with HKL2000 (Table 2) to an R_{merge} of 9.0%, 10.2%, and 7.2% for inflection point, peak, and remote respectively.

Structure Determination and Refinement—The structures of Ba-SrtB and Sa-SrtB were determined independently by MAD phasing using CNS (Brunger, et al., 1998) and refined initially to 1.6 and 2.0 Å, respectively, using CNS against the averaged peak data. The initial models were built manually using QUANTA. The models were further refined against MAD peak data. For the Sa-SrtB the final R was 23.7% and the free R 24.8% with zero σ cutoff (Table 3), and for the Ba-SrtB, the final R was 22.0% and the free R 26.3% with zero σ cutoff (Table 4). The stereochemistry of the structures was checked with PROCHECK (Laskowski et al., 1993) and the Ramachandran plot. For the Sa-SrtB, the main chain torsion angles for all residues, except two, are in allowed regions, and the two residues are in additional allowed regions. Four residues from N terminus of the construct do not have well-defined electron density and appear disordered. For the Ba-SrtB, the main chain torsion angles for all residues are in allowed regions and the three residues are in additional allowed regions. 34 residues from N terminus and 4 residues (235–240) of the construct do not have well-defined electron density and appear disordered.

Acknowledgments

We thank all members of the SBC at ANL for their help conducting experiments and Lindy Keller for her help in preparation of this manuscript. This work was supported by grants from the National Institutes of Health, National Institute of General Medical Sciences to A.J. (GM62414), D.M.M. (GM58266), and O.S. (AI38897 and AI52474); by a University of Chicago-Argonne National Laboratory joint research grant under H.28 of the prime contract; by the U.S. Department of Energy, Office of Biological and Environmental Research, under contract W-31-109-Eng-38; and by the Regional Center of Excellence in Biodefense and Emerging Infectious Diseases (RCE) Program. The authors wish to acknowledge membership and support from the Region V (Great Lakes) RCE, award 1-U54-AI-057153 from the National Institutes of Allergy and Infectious Diseases. The submitted manuscript has been created by the University of Chicago as Operator of Argonne National Laboratory ("Argonne") under Contract No. W-31-109-ENG-38 with the U.S. Department of Energy. The U.S. Government retains for itself, and others acting on its behalf, a paid-up, nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.

References

Barnett TC, Scott JR. Differential recognition of surface proteins in *Streptococcus pyogenes* by two sortase gene homologs. J. Bacteriol 2002;184:2181–2191. [PubMed: 11914350]

Bierne H, Mazmanian SK, Trost M, Pucciarelli MG, Liu G, Dehoux P, Jansch L, Garcia-del Portillo F, Schneewind O, Cossart P, et al. Inactivation of the *srtA* gene in *Listeria monocytogenes* inhibits anchoring of surface proteins and affects virulence. Mol. Microbiol 2002;43:869–881. [PubMed: 11929538]

Bolken TC, Franke CA, Jones KF, Zeller GO, Jones CH, Dutton EK, Hruby DE. Inactivation of the *srtA* gene in *Streptococcus gordonii* inhibits cell wall anchoring of surface proteins and decreases *in vitro* and *in vivo* adhesion. Infect. Immun 2001;69:75–80. [PubMed: 11119491]

Brumfitt W, Hamilton-Miller J. Methicillin-resistant Staphylococcus aureus. N. Engl. J. Med 1989;320:1188–1199. [PubMed: 2651925]

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, et al. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr 1998;54:905–921. [PubMed: 9757107]

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 2003;31:3497–3500. [PubMed: 12824352]

Conolly KM, Smith BT, Pilpa R, Ilangovan U, Jung ME, Clubb RT. Sortase from *Staphylococcus aureus* does not contain a thiolate-imidazolium ion pair in its active site. J. Biol. Chem 2003;278:34061–34065. [PubMed: 12824164]

Cossart P, Jonquieres R. Sortase, a universal target for therapeutic agents against gram-positive bacteria? Proc. Natl. Acad. Sci. USA 2000;97:5013–5015. [PubMed: 10805759]

Dhar G, Faull KF, Schneewind O. Anchor structure of cell wall surface proteins in *Listeria monocytogenes*. Biochemistry 2000;39:3725–3733. [PubMed: 10736172]

Dixon TC, Meselson M, Guillemin J, Hanna PC. Anthrax. N. Engl. J.Med 1999;341:815–826.

Foster TJ, Höök M. Surface protein adhesins of *Staphylococcus aureus*. Trends Microbiol 1998;6:484–488. [PubMed: 10036727]

Garandeau C, Reglier-Poupet H, Dubail I, Beretti JL, Berche P, Charbit A. The sortase SrtA of *Listeria monocytogenes* is involved in processing of internalin and in virulence. Infect. Immun 2002;70:1382–1390. [PubMed: 11854224]

Guncar G, Klemencic I, Turk B, Turk V, Karaoglanovic-Carmona A, Juliano L, Turk D. Crystal structure of cathepsin X: a flip-flop of teh ring of His 23 allows carboxy-monopeptidase and carboxy-dipeptidase activity of the protease. Struct. Fold. Des 2002;8:305–313.

Hava DL, Hemsley CJ, Camilli A. Transcriptional regulation in the *Streptococcus pneumoniae rlrA* pathogenicity islet by RlrA. J. Bacteriol 2003;185:413–421. [PubMed: 12511486]

Higashi Y, Strominger JL, Sweeley CC. Structure of a lipid intermediate in cell wall peptidoglycan synthesis: a derivative of C55 isoprenoid alcohol. Proc. Natl. Acad. Sci. USA 1967;57:1878–1884. [PubMed: 5231417]

Higashi Y, Strominger JL, Sweeley CC. 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 1970;245:3697–3702. [PubMed: 4248530]

Holm L, Sander C. Protein folds and families: sequence and structural alignments. Nucleic Acid. Res 1999;27:244–247. [PubMed: 9847191]

Huang X, Aulabaugh A, Ding W, Kapoor B, Alksne L, Tabei K, Ellestad G. Kinetic mechanism of *Staphylococcus aureus* sortase SrtA. Biochemistry 2003;42:11307–11315. [PubMed: 14503881]

Ilangovan U, Ton-That H, Iwahara J, Schneewind O, Clubb RT. Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. Proc. Natl. Acad. Sci. USA 2001;98:6056–6061. [PubMed: 11371637]

Johnston SC, Larsen CN, Cook WJ, Wilkinson KD, Hill CP. Crystal structure of a deubiquitinating enzyme (human UCH-L3) at 1.8Å resolution. EMBO J 1997;16:3787–3796. [PubMed: 9233788]

Jonsson IM, Mazamanian SK, Schneewind O, Vendrengh M, Bremell T, Tarkowski A. On the role of *Staphylococcus aureus* sortase and sortase-catalyzed surface protein anchoring in murine septic arthritis. J. Infect. Dis 2002;185:1417–1424. [PubMed: 11992276]

Jonsson IM, Mazmanian SK, Schneewind O, Bremell T, Tarkowski A. The role of *Staphylococcus aureus* sortase A and sortase B in murine arthritis. Microbes Infect 2003;5:775–780. [PubMed: 12850203]

Kharat AS, Tomasz A. Inactivation of the *srtA* gene affects localization of surface proteins and decreases adhesion of *Streptococcus pneumoniae* to human pharyngeal cells *in vitro*. Infect. Immun 2003;71:2758–2765. [PubMed: 12704150]

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr 1993;26:283–291.

Lee SF, Boran TL. Roles of sortase in surface expression of the major protein adhesin P1, saliva induced aggregation and adherence, and cariogenicity of *Streptococcus mutans*. Infect. Immun 2003;71:676–681. [PubMed: 12540545]

Mazmanian, SK.; Schneewind, O. Cell wall anchored surface proteins and lipoproteins of Gram-positive bacteria. In: Sonenshine, A.; Losick, R.; Hoch, J., editors. *Bacillus subtilis* and Other Gram-Positive Bacteria. Vol. Second Edition. New York: ASM Press; 2004. in press

Mazmanian SK, Liu G, Ton-That H, Schneewind O. *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. Science 1999;285:760–763. [PubMed: 10427003]

Mazmanian SK, Liu G, Jensen ER, Lenoy E, Schneewind O. *Staphylococcus aureus* mutants defective in the display of surface proteins and in the pathogenesis of animal infections. Proc. Natl. Acad. Sci. USA 2000;97:5510–5515. [PubMed: 10805806]

Mazmanian SK, Ton-That H, Su K, Schneewind O. An iron-regulated sortase enzyme anchors a class of surface protein during *Staphylococcus aureus* pathogenesis. Proc. Natl. Acad. Sci. USA 2002;99:2293–2298. [PubMed: 11830639]

Mazmanian SK, Skaar EP, Gasper AH, Humayun M, Gornicki P, Jelenska J, Joachimiak A, Missiakas DM, Schneewind O. Passage of heme-iron across the envelope of *Staphylococcus aureus*. Science 2003;299:906–909. [PubMed: 12574635]

Navarre WW, Schneewind O. Proteolytic cleavage and cell wall anchoring at the LPXTG motif of surface proteins in gram-positive bacteria. Mol. Microbiol 1994;14:115–121. [PubMed: 7830549]

Navarre WW, Ton-That H, Faull KF, Schneewind O. Anchor structure of staphylococcal surface proteins. II. COOH-terminal structure of muramidase and amidase-solubilized surface protein. J. Biol. Chem 1998;273:29135–29142. [PubMed: 9786922]

Navarre WW, Ton-That H, Faull KF, Schneewind O. Multiple enzymatic activities of the murein hydrolase from staphylococcal phage f11. Identification of a D-alanyl-glycine endopeptidase activity. J. Biol. Chem 1999;274:15847–15856. [PubMed: 10336488]

Otwinowski Z, Minor W. Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.

Pallen MJ, Lam AC, Antonio M, Dunbar K. An embarrassment of sortases: a richness of substrates. Trends Microbiol 2001;9:97–101. [PubMed: 11239768]

Perry AM, Ton-That H, Mazmanian SK, Schneewind O. Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. III. Lipid II is an *in vivo* peptidoglycan substrate for sortase-catalyzed surface protein anchoring. J. Biol. Chem 2002;277:16241–16248. [PubMed: 11856734]

Read TD, Peterson SN, Tourasse N, Baille LW, Paulsen IT, Nelson KE, Tettelin H, Fouts DE, Eisen JA, Gill SR, et al. The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. Nature 2003;423:81–86. [PubMed: 12721629]

Ruzin A, Severin A, Ritacco F, Tabei K, Singh G, Bradford PA, Siegel MM, Projan SJ, Shlaes DM. Further evidence that a cell wall precursor [C(55)-MurNAc-(peptide)-GlcNAc] serves as an acceptor in a sorting reaction. J. Bacteriol 2002;184:2141–2147. [PubMed: 11914345]

Schleifer KH, Kandler O. Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol. Rev 1972;36:407–477. [PubMed: 4568761]

Schneewind O, Model P, Fischetti VA. Sorting of protein A to the staphylococcal cell wall. Cell 1992;70:267–281. [PubMed: 1638631]

Schneewind O, Mihaylova-Petkov D, Model P. Cell wall sorting signals in surface protein of Gram-positive bacteria. EMBO J 1993;12:4803–4811. [PubMed: 8223489]

Schneewind O, Fowler A, Faull KF. Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. Science 1995;268:103–106. [PubMed: 7701329]

Shao F, Merritt PM, Bao Z, Innes RW, Dixon JE. A *Yersinia* effector and a *Pseudomonas* avirulence protein define a family of cysteine proteases functioning in bacterial pathogenesis. Cell 2002;109:575–588. [PubMed: 12062101]

Strominger JL, Izaki K, Matsuhashi M, Tipper DJ. Peptidoglycan transpeptidase and D-alanine carboxypeptidase: penicillin-sensitive enzymatic reactions. Fed. Proc 1967;26:9–18. [PubMed: 4959942]

Tenover FC, Biddle JW, Lancaster MV. Increasing resistance to vancomycin and other glycopeptides in *Staphylococcus aureus*. Emerg. Infect. Dis 2001;7:327–332. [PubMed: 11294734]

Ton-That H, Schneewind O. Anchor structure of staphylococcal surface proteins. IV. Inhibitors of the cell wall sorting reaction. J. Biol. Chem 1999;274:24316–24320. [PubMed: 10446208]

Ton-That H, Schneewind O. Assembly of pili on the surface of *C. diphtheriae*. Mol. Microbiol 2003;50:1429–1438. [PubMed: 14622427]

Ton-That H, Faull KF, Schneewind O. Anchor structure of staphylococcal surface proteins. I. A branched peptide that links the carboxyl terminus of proteins to the cell wall. J. Biol. Chem 1997;272:22285–22292. [PubMed: 9268378]

Ton-That H, Labischinski H, Berger-Bächi B, Schneewind O. Anchor structure of staphyococcal surface proteins. III. The role of the FemA, FemB, and FemX factors in anchoring surface proteins to the bacterial cell wall. J. Biol. Chem 1998;273:29143–29149. [PubMed: 9786923]

Ton-That H, Liu G, Mazmanian SK, Faull KF, Schneewind O. Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. Proc. Natl. Acad. Sci. USA 1999;96:12424–12429. [PubMed: 10535938]

Ton-That H, Mazmanian H, Faull KF, Schneewind O. Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. I. Sortase catalyzed *in vitro* transpeptidation reaction using LPXTG peptide and NH_2-Gly_3 substrates. J. Biol.Chem 2000;275:9876–9881. [PubMed: 10734144]

Ton-That H, Mazmanian SK, Alksne L, Schneewind O. Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. II. Cysteine 184 and histidine 120 of sortase A form a thiolate imidazolium ion pair for catalysis. J. Biol. Chem 2002;277:7447–7452. [PubMed: 11714722]

Walsh MA, Dementieva I, Evans G, Sanishvili R, Joachimiak A. Taking MAD to the extreme: ultra fast protein crystal structure determination. Acta Crystallogr. D Biol. Crystallogr 1999;55:1168–1173. [PubMed: 10329779]

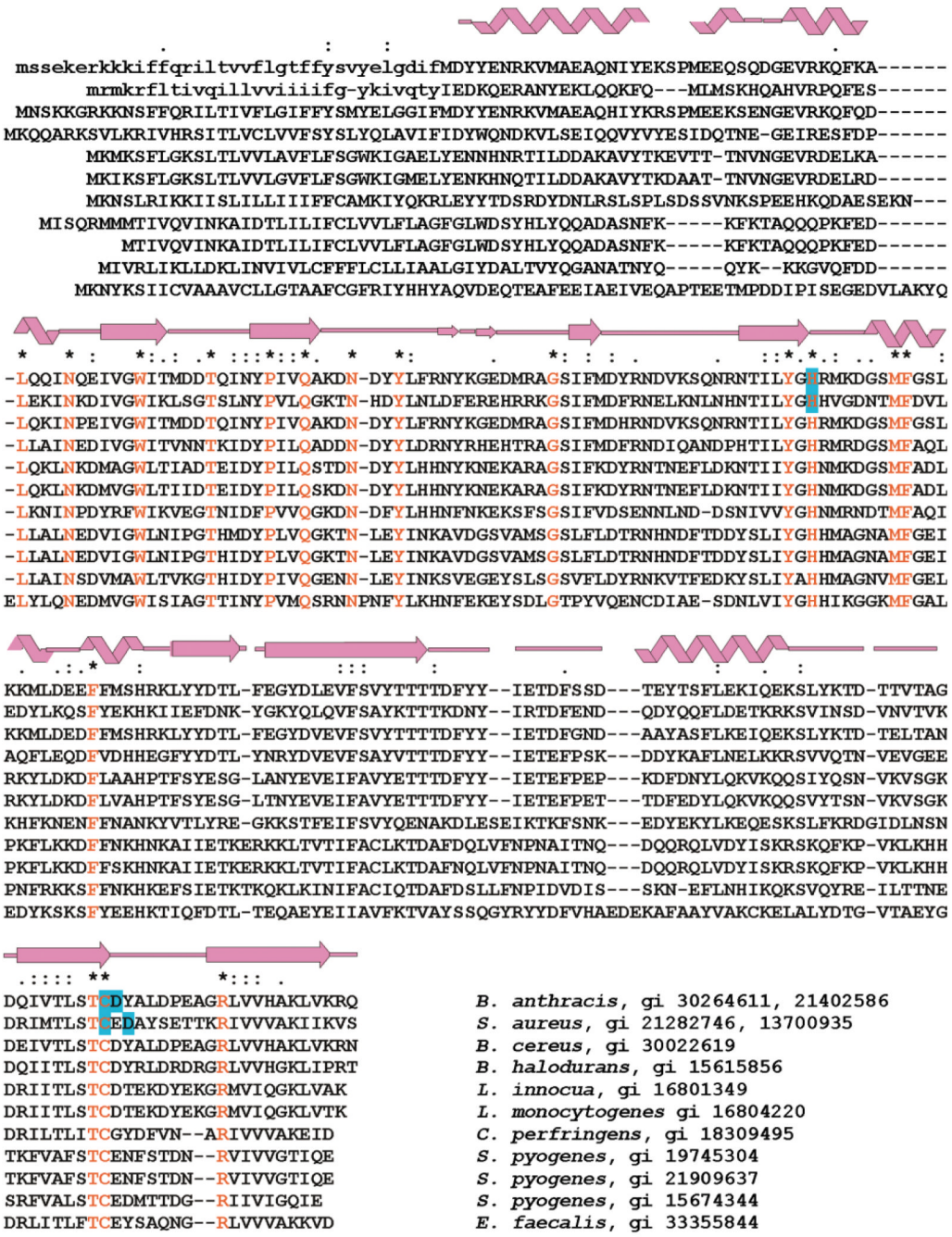

Figure 1. Multiple Sequence Alignment of SrtB Homologs Created Using ClustalX v. 1.82

"*," identical residues shown in red; "." and ":," conserved residues. Amino acids of the catalytic triad (this work) are highlighted in blue. N-terminal signal peptide sequences replaced with His tags in *B. anthracis* and *S. aureus* proteins used for crystallization are shown in lower case. Secondary structure elements are indicated above the sequence alignment. Chenna et al. (2003) are the authors of the software ClustalX v. 1.82 that was used to calculate alignment.

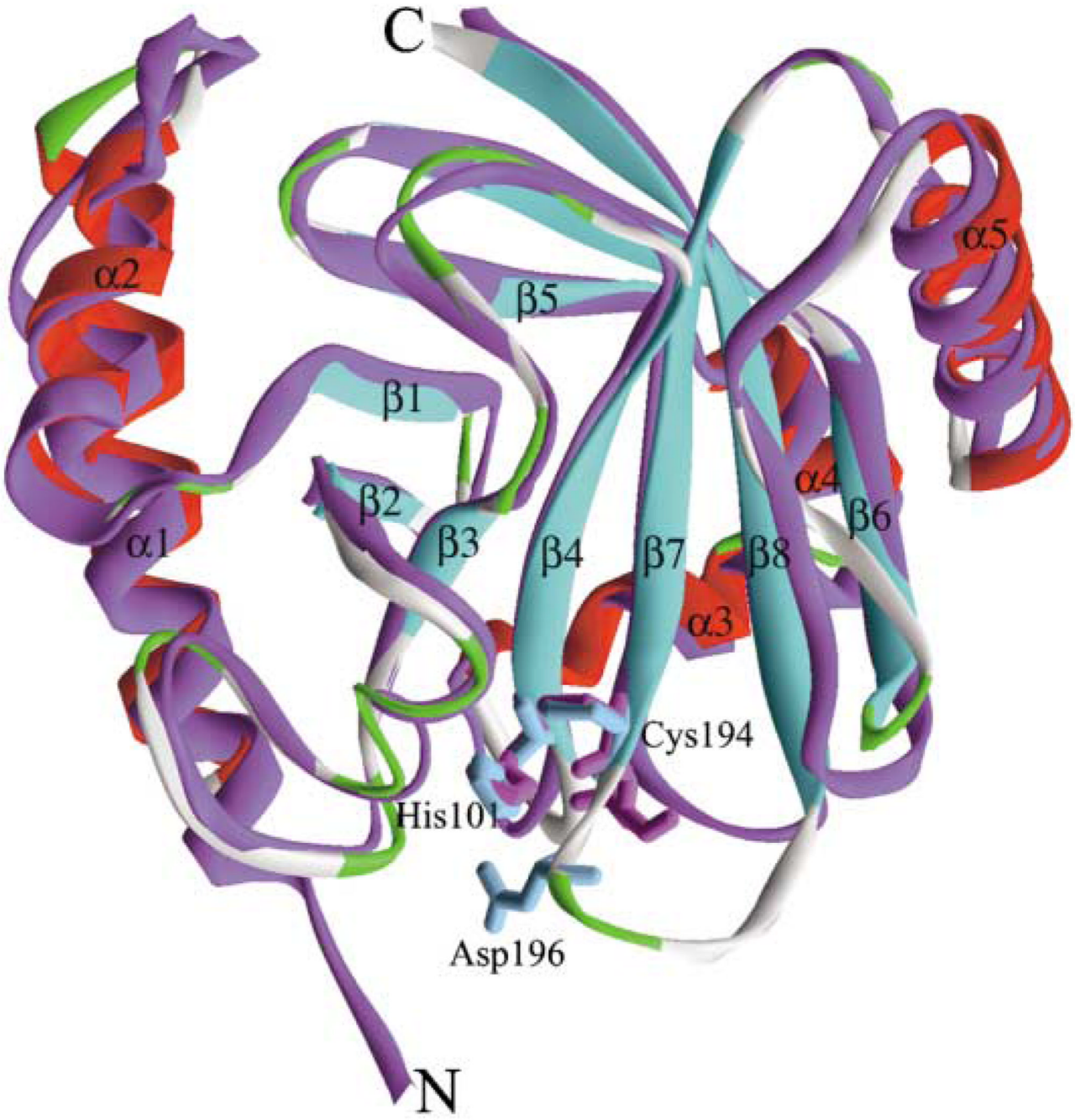

Figure 2. Comparison of the Structures of Ba-SrtB and Sa-SrtB

For Sa-SrtB helices are marked, red; strands, blue; loops, green; and turns, silver. Ba-SrtB is depicted in purple. N- and C termini, as well as presumed active site residues in Sa-SrtB are labeled. The figure was prepared with WebLabViewPro.

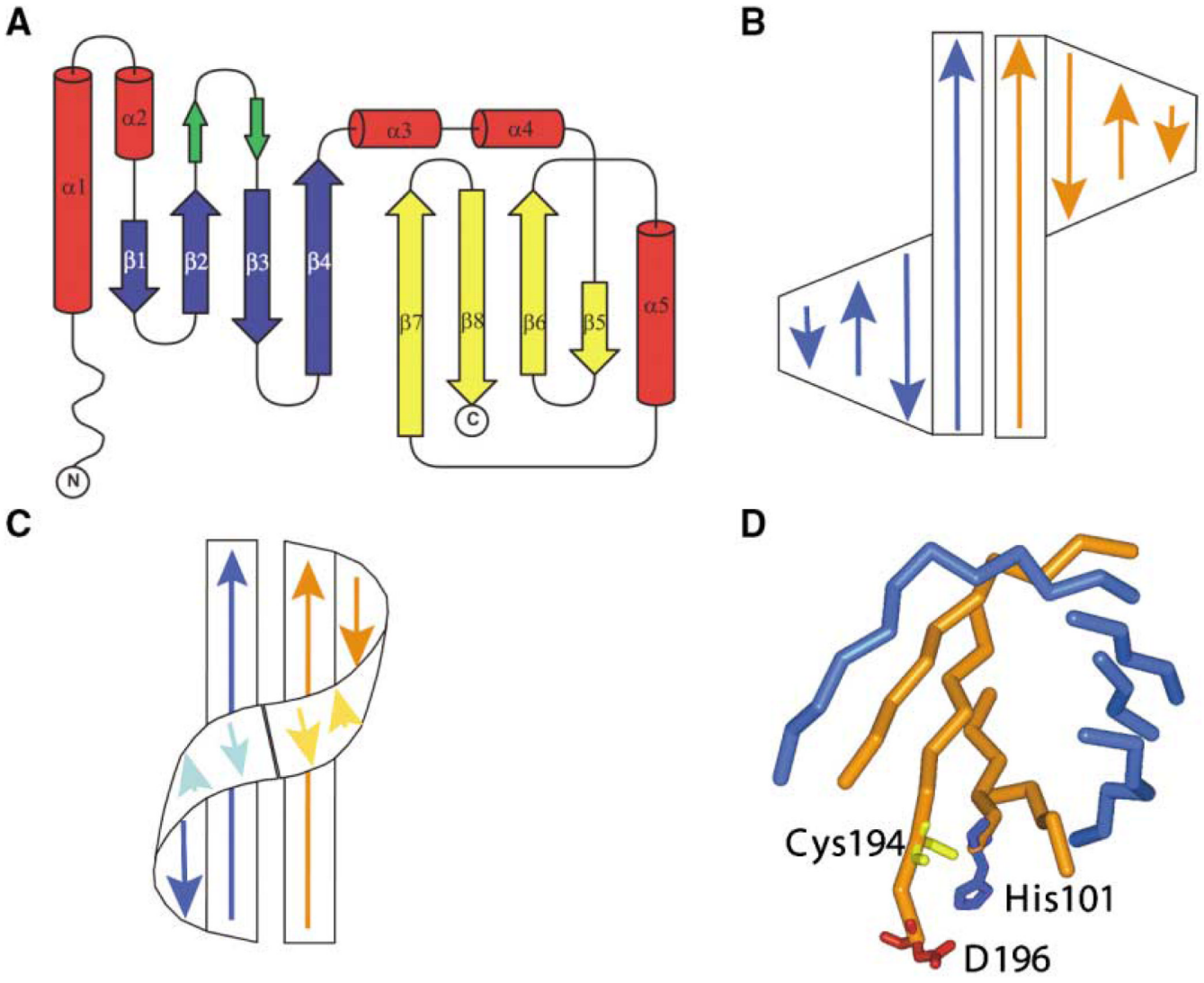

Figure 3. The Sortase Fold

(A) The secondary structure diagram of SrtB, strands forming the barrel are in blue and yellow, helices are in red, and the hairpin is in green.

(B) The core β-barrel is shown in a schematic "open" representation. One motif (in blue) is composed of strands β1↑-β2↓-β3↓-β4↑, and second motif (orange) is composed of strands β7↓-β8↑-β6↓-β5↑. The two motifs are connected through interaction of two long parallel β strands (poles) β4 and β7 to form a pseudo 2-fold arrangement.

(C) The cylinder is completed by sealing the edges between the parallel β1 and β5 strands.

(D) The putative catalytic triad of the active site located on the edge of the barrel. The β strands of the barrel color coded as in (B) and (C).

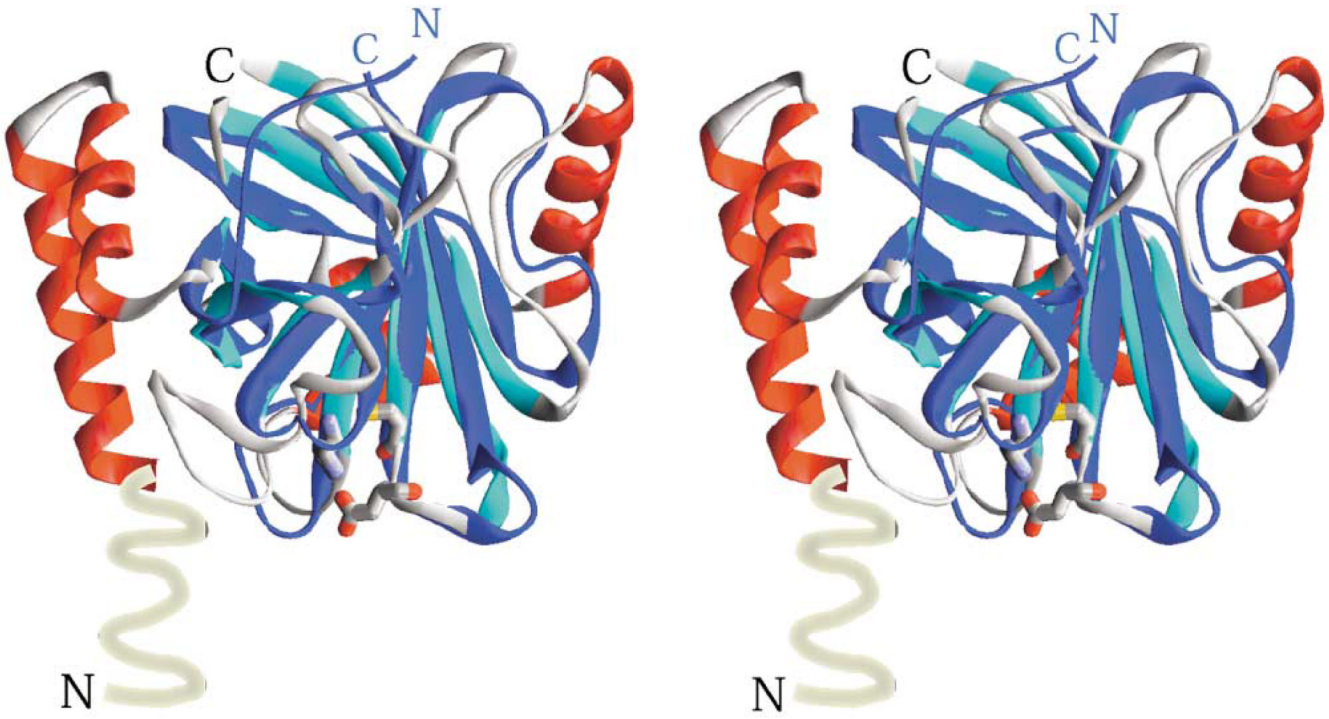

Figure 4. A Stereoview Comparison of Sortases A and B Structures

The structure of Sa-SrtA (dark blue, 1IJA_A) was superimposed on the structure of Sa-SrtB (blue β strands, red α helices, and sliver loops). N- and C-termini of SrtA and B are labeled in blue and black respectively, N-terminal signal (membrane anchor) peptide is shown as unstructured region. Residues of the putative active site of Sa-SrtB (His101, Cys194, and Asp196) are shown in a stick representation.

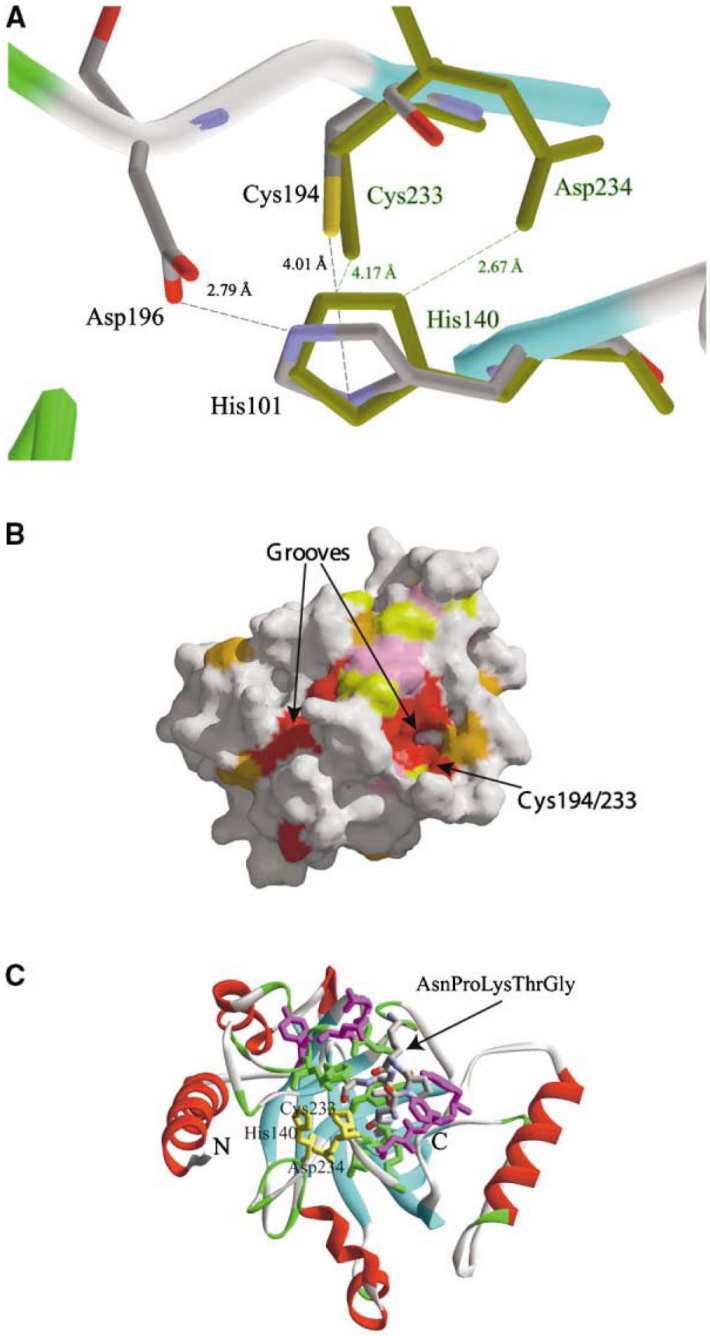

Figure 5. Comparison of the Active Sites of Ba-SrtB and Sa-SrtB

(A) Interaction between key residues in the catalytic triad of Sa-SrtB and Ba-SrtB (in green).

(B) Surface representation of the area near the putative Sa-SrtB active site shows two adjacent grooves arranged in a V shape on the protein surface. One groove includes a cluster of highly conserved residues (red) around the active site Cys194. The second groove presents a potential substrate binding site.

(C) The presumed peptide substrate NPKTG was modeled onto the 1.5 Å structure of Ba-SrtB using QUANTA. The peptide was regularized, but the protein side chains were not altered. N- and C-termini are labeled. Active site residues are in yellow. Side chains of residues in the groove that are conserved among sortase B family are in green. Side chains of residues in the

groove that are different between Sa-SrtB and Ba-SrtB that may explain difference in peptide specificity are in magenta. The protein is viewed from the surface of the cell wall.

Table 1

Sa-SrtB: Summary of Crystal and MAD Data

Unit cell	a = 71.208 Å, b = 104.367 Å, c = 58.087 Å, α = β = γ = 90°			
Space group	$P2_12_12$			
MW Da (residues)	23,760 (216)			
Mol (AU)	2			
SeMet (AU)	10			
	MAD Data Collection			
	Edge	Peak	Remote	High Resolution
Wavelength (Å)	0.9794	0.9793	0.9538	1.0332
Resolution range (Å)	2.48	2.48	2.48	1.97
No. of unique reflections	15,446	15,867	15,651	29,674
Completeness (%)	96.5	99.2	97.8	94.4
R_{merge} (%)	9.1	9.2	8.3	11.0

Table 2

Ba-SrtB: Summary of Crystal and MAD Data

Unit cell	a = 40.474 Å, b = 64.599 Å, c = 42.964 Å, β = 105.770°		
Space group	$P2_1$		
MW Da (residues)	29,210 (254)		
Mol (AU)	1		
SeMet (AU)	11		
	MAD Data Collection		
	Edge	Peak	Remote
Wavelength (Å)	0.9794	0.9793	0.9538
Resolution range (Å)	1.6	1.6	1.6
No. of unique reflections	27,014	27,074	26,950
Completeness (%)	98.6	98.9	98.4
R_{merge} (%)	9.0	10.2	7.2

Table 3

Sa-SrtB: Crystallographic Statistics

Phasing								
	Centric		**Acentric**			**All**		
Resolution range (Å)	**FOM**	**Phasing Power**	**FOM**	**Phasing Power**	**No.**	**FOM**	**Phasing Power**
50.0–2.5	0.59	1.49	0.39	1.34	28,607	0.403	1.354
Density modification						0.875	
Refinement							
Resolution range (Å)	50–2.0						
No. of reflections	50,697						
σ cutoff	0						
R value (%)	23.7						
Free R value (%)	24.8 (2,466)						
Rms deviations from ideal geometry							
Bond length (1–2) (Å)	0.009						
Angle (°)	1.3						
Dihedral (°)	23.7						
Improper (°)	0.75						
No. of atoms							
Protein	3,492						
Water	319						
Mean B factor (Å^2)							
All atoms	31.7						
Ramachandran plot statistics (%)							
Residues in most favored regions	83.7						
Residues in additional allowed regions	16.3						
Residues in disallowed region	0.0						

Table 4

Ba-SrtB: Crystallographic Statistics

Phasing								
	Centric		**Acentric**			**All**		
Resolution range (Å)	**FOM**	**Phasing Power**	**FOM**	**Phasing Power**	**No.**	**FOM**	**Phasing Power**
50.0–1.6	0.72	3.42	0.647	3.11	53,007	0.648	3.12
Density Modification						0.820	
Refinement							
Resolution range (Å)	50–1.60						
No. of reflections	53,007						
σ cutoff	0						
R value (%)	22.0						
Free R value (%)	26.3 (2,643)						
Rms deviations from ideal geometry							
Bond length (1–2) (Å)	0.008						
Angle (°)	1.3						
Dihedral (°)	24.1						
Improper (°)	0.72						
No. of atoms							
Protein	1,781						
Water	187						
Mean B factor (Å^2)							
All atoms	26.1						
Ramachandran plot statistics (%)							
Residues in most favored regions	98.6						
Residues in additional allowed regions	1.4						
Residues in disallowed region	0.0						