Surface proteins in Gram-positive bacteria are incorporated into the cell wall through a peptide ligation reaction catalyzed by transpeptidase sortase. Six main classes (A–F) of sortase have been identified of which class A sortase is meant for housekeeping functions. The prototypic housekeeping sortase A (SaSrtA) from *Staphylococcus aureus* cleaves LPXTG-containing proteins at the scissile T–G peptide bond and ligates protein-LPXTG to the terminal Gly residue of the nascent cross-bridge of peptidoglycan lipid II precursor. Sortase-mediated ligation (“sortagging”) of LPXTG-containing substrates and Gly-terminated nucleophiles occurs *in vitro* as well as *in cellulo* in the presence of Ca\(^{2+}\) and has been applied extensively for protein conjugations. Although the majority of applications emanate from SaSrtA, low catalytic efficiency, LPXTG specificity restriction, and Ca\(^{2+}\) requirement (particularly for *in cellulo* applications) remain a drawback. Given that Gram-positive bacteria genomes encode a variety of sortases, natural sortase mining can be a viable complementary approach akin to engineering of wild-type SaSrtA. Here, we describe the structure and specificity of a new class E sortase (SavSrtE) annotated to perform housekeeping roles in *Streptomyces avermitilis*. Biochemical experiments define the attributes of an optimum peptide substrate, demonstrate Ca\(^{2+}\)-independent activity, and provide insights about contrasting functional characteristics of SavSrtE and SaSrtA. Crystal structure, substrate docking, and mutagenesis experiments have identified a critical residue that dictates the preference for a non-canonical LA\(\text{TG}\) recognition motif over LPXTG. These results have implications for rational tailoring of substrate tolerance in sortases. Besides, Ca\(^{2+}\)-independent orthogonal specificity of SavSrtE is likely to expand the sortagging toolkit.

Several proteins of Gram-positive bacteria after their translocation across the membrane are covalently attached to the cell wall by action of a unique family of cysteine transpeptidases referred to as sortases (1–3). Surface proteins meant for covalent anchoring harbor a cell wall sorting signal in the C-terminal region that contains a sortase-recognition LPXTG type of pentapeptide sequence motifs (4). The general mechanism of sortase-mediated anchoring involves two steps in which the enzyme first cleaves the T–G peptide bond and captures the protein by forming an enzyme-thioacyl intermediate. This thioacyl intermediate is resolved by nucleophilic attack of an amine moiety from the peptidoglycan cross-bridge (1). Importantly, many surface proteins that are sortase substrates play crucial roles in immune evasion, biofilm formation, and other critical processes associated with bacterial pathogenesis (5, 6). Sortase knock-outs display reduced virulence with retention of cellular viability (7–9). Hence, sortases are viewed as attractive targets for the development of novel therapeutics, especially against antibiotic-resistant strains (10, 11).

Gram-positive bacteria encode multiple sortases and a variety of substrates (12). Accordingly, sortases are grouped into six classes (A–F) based on their sequence similarity, membrane topology, function, or substrate specificity (13–15). Class A enzymes anchor a large number of proteins and act as a housekeeping sortase. Sortases belonging to classes B–D are meant to perform specialized functions. The roles of sortases E and F remain somewhat unclear, although class E sortases are believed to function like housekeeping enzymes. Sortase A (SrtA), first discovered in *Staphylococcus aureus*, is considered a prototype sortase (3). SaSrtA recognizes the LPXTG motif in *S. aureus* surface proteins and anchors them to the cell wall by forming a peptide bond to the pentaglycine arm of the peptidoglycan (16). Sortase-mediated peptide ligation between short synthetic peptides or large engineered proteins embedded with the LPXTG motif and an aminoglycine-derivatized moiety occurs well *in vitro* (17) and has enabled unprecedented applications of SaSrtA in protein engineering (18). SaSrtA or its variants remain the workhorse of most sortagging applications (19). The housekeeping sortases of *Streptococcus pyogenes* and
Lactobacillus plantarum are the only other enzymes that have been employed, although with limited utility (20, 21). Furthermore, newer challenging applications of sortases would require the availability of enzymes endowed with enhanced catalytic efficiency as well as orthogonal specificities in both donor and acceptor polypeptides. Moreover, Ca\(^{2+}\)-independent sortase activity is desirable to facilitate intracellular sortagging applications (19, 22, 23). The mining of bacterial genomes for improved sortase enzymes capable of recognizing diverse sorting motifs appear very attractive in this regard.

The structure of SaSrtA obtained by both NMR and X-ray highlights the presence of a typical “sortase” fold comprising an eight-stranded \(\beta\)-barrel architecture (24, 25). The active site is composed of the catalytic residues His-120, Cys-184, and Arg-197 situated at the end of the long groove along one side of the \(\beta\)-barrel connected by random coil loops and two short helices. The loops connecting \(\beta 2/\beta 3, \beta 3/\beta 4, \beta 6/\beta 7, \text{and } \beta 7/\beta 8\) strands form the walls of the groove. Residues located in the \(\beta 6/\beta 7\) loop play crucial roles in substrate sorting and facilitate catalytic turnover of those peptide substrates that are endowed with a “kinked” conformation due to the Pro residue in the LPXTG substrate (24). Accordingly, the LPXTG substrate containing Pro analogs and homologs capable of generating Pro-like conformational features is well tolerated against LAXTG or LGXTG peptide substrates (26). Interestingly, class E sortases present in some bacteria are annotated to process non-canonical LAXTG or other previously unknown sorting signals (13). The presence of multiple class E sortases is noted in the GC-rich genome of actinomycetes. Duong et al. (27) detected two class E sortases (SrtE1 and E2) in Streptomyces coelicolor and used knock-out experiments to demonstrate the requirement of sortase activity in aerial development. The aerial hyphae formation requires “chaplin” proteins, some of which contain LAXTG as a potential sorting signal (28). The genome of Streptomyces avermitilis has revealed the presence of four putative SrtE enzymes of which SrtE3 (SAV4333) is distinct from SrtE1 and SrtE2 of S. coelicolor (27, 29).

This study was initiated with a view to define the substrate specificity of SrtE3 from S. avermitilis and to delineate the structural features responsible for recognition of a non-canonical LAXTG peptide substrate. Our results of transeptidation assays carried out using a battery of peptides reveal the attributes of a good sortase E substrate not previously recognized in classical SaSrtA or other sortases. The crystal structure of the enzyme together with bioinformatics, modeling, and mutagenesis provide insights into the altered substrate specificity of this new class E housekeeping sortase.

**Results**

**Design and expression of a catalytic domain of SrtE3 (SavSrtE)**

*S. avermitilis* genome encodes at least four putative class E sortases. SrtE3 (SAV4333) is a 230-residue protein composed of an N-terminal transmembrane region (residues 12–32) and a catalytic domain (residues 83–214). We chose to express the N-terminal truncated version of the enzyme to facilitate solubility. Accordingly, the sequence representing residues 51–230 was expressed with a hexa-His tag at the N terminus. The

**Structure and specificity of a class E sortase**

![Figure 1. Characterization of recombinant SavSrtE. A, SDS-PAGE of purified recombinant SavSrtE comprising residues 51–230 expressed with an N-terminal hexa-His tag: lane 1, molecular weight marker; lane 2, SavSrtE. B, ES-MS of the purified protein. The deconvoluted spectrum produced two peaks (22,157 and 22,335 Da, respectively). A mass of 22,157 Da corresponds to the calculated mass of expressed sequence (22,289 Da) without a Met residue. The second peak, which is higher by 178 Da, presumably represents a glucosylated protein.](image-url)

N-terminal 50-residue deleted construct of SrtE3 referred to as SavSrtE was used for the study.

The Ni-NTA affinity-purified SavSrtE migrated as a neat single band around 20–25 kDa on SDS-PAGE indicating high purity of the protein preparation (Fig. 1A). However, ES-MS analysis yielded two components (22,157 and 22,335 Da, respectively) whose mass differed by 178 Da (Fig. 1B). The mass of the first component fits to the calculated mass of the expressed protein sequence (22,289) minus the mass of a Met residue indicating processing of N-terminal Met by Escherichia coli aminopeptidases. The addition of 178 Da presumably arose due to N-glucosylation of the des-Met protein. Such modifications have been noted earlier in many N-terminal His-tagged proteins (30).

**Expressed SavSrtE is an active sortase and prefers Ala over Pro at the second position of the pentapeptide-sorting motif**

We carried out transeptidation assays with a variety of peptide substrates to define the substrate preference and tolerance of SavSrtE. We generated several substrate variants (Fig. 2A) by incorporating proline (Pro), 4-hydroxyproline, 3,4-dehydroproline, azetidine-2-carboxylic acid, glycine (Gly), alanine (Ala), \(\alpha\)-aminoisobutyric acid, and \(\beta\)-alanine (\(\beta\)-Ala) at the second position of the pentapeptide motif in a YALKX(T/G) model peptide template based on our earlier studies with SaSrtA (26).

We first evaluated SavSrtE-catalyzed transeptidation reaction of Ala-containing peptide (YALKX(T/G), where \(X = \text{Ala}\)) using GGKKY as an acceptor. The reaction was carried out by incubating the peptide (0.5 mM) with GGGKY (1 mM) and SavSrtE (50 \(\mu\)M) at 20 °C. An aliquot from the reaction mixture was analyzed by RP-HPLC (inset, Fig. 2B) as described under “Experimental procedures,” and the product was characterized by mass spectrometry (data not shown). The time course of the reaction depicted in Fig. 2B shows that the reaction attained equilibrium in about 5–8 h and produced a 20% yield. Subsequently, we carried out the transeptidation assay of each YALKX(T/G) peptide (\(X = 1–8\)) under identical conditions for
6 h and quantified the product (Fig. 2C). Interestingly, Gly- or β-Ala-containing peptides did not yield any product indicating a role of peptide conformation in substrate recognition. The α-aminoisobutyric acid peptide produced about 15% transpeptidation yield as compared with about 20% by Ala peptide. However, product formation was reduced to almost half that of Ala in Pro peptide (9%) and significantly diminished (5% or less) in other Pro analogs.

The above results show that Pro analogs with Pro-like conformational attributes are poor substrates of SavSrtE as compared with SaSrtA. In contrast, Ala or its analogs at this position are more effective substrates. However, the overall equilibrium yield of about 20% obtained with SavSrtE using YALANTGK substrate was less than half as compared with that produced by SaSrtA (45%) with the counterpart YALPNTGK peptide (26).

What could be the reason for this lower product conversion? Can residues flanking the LAXTG motif play a role? A closer inspection of the putative sortase protein substrates encoded in the *S. avermitilis* genome revealed the presence of neutral residues, such as Asn, Ala, or Gly flanking the LAXTG pentapeptide motif. This prompted us to investigate whether the presence of a Lys residue in YALANTGK exerted any deleterious effect on the transpeptidation reaction. Accordingly, we synthesized the Ala- (neutral) or Glu (acidic)-terminated peptides in addition to Lys-terminated YALANTGK to generate the LANTG substrate flanked by Ala/Ala, Ala/Glu, and Ala/Lys residues. Interestingly, peptide containing the Ala/Ala-flanking residue produced (40%) about 5-fold higher equilibrium transpeptidation yield as compared with Ala/Glu (8%) and 2-fold or more as compared with Ala/Lys (18%) peptide (Fig. 3A).

We also replaced the Ala residue preceding Leu of the motif and created a Asn/Ala-flanking pair and generated two peptides corresponding to the LAXTG and LPXTG motifs, respectively. The YNLAETGA peptide yielded about 35% product, which was comparable with that obtained with YALANTGK. Interestingly, equilibrium yield in the case of YNLPETGA peptide was found to be about 12–14% as compared with about 9% obtained with YALPNTGK suggesting that flanking residues are important in the context of both LAXTG and LPXTG substrates (Fig. 3B).
Structure and specificity of a class E sortase

SavSrtE crystals (dimensions ~0.5 × 0.1 × 0.05 mm) diffracted to 1.65 Å. The protein crystallized in the P3_{21}_{21} space group with unit cell parameters a = b = 85.84 Å, c = 48.20 Å, \( \alpha = \beta = 90^\circ, \gamma = 120^\circ \) and a monomer in the asymmetric unit (Table 1). The modeled structure lacked interpretable electron density for residues 51–73 (KGTVRPTAAPGASARTSP-KE) and residues 201–204 (EWGH). SavSrtE displayed typical eight-stranded \( \beta \)-barrel sortase fold despite low sequence identity with other sortases (Fig. 4A). The overall structure of SavSrtE resembles SrtA more closely than sortases of class B or C (Fig. 4B). However, the \( \beta_6/\beta_7 \) loop (21 residues) in SavSrtE contains two \( \psi_0 \) helices and is slightly longer than the equivalent loop in class A sortases (16–18 residues) but is shorter than in class B sortases (~38 residues). Although the first \( \psi_0 \) helix is similar to that of SaSrtA and is expected to contact the bound peptide, the second helix partially overlaps with the signature \( \alpha \)-helix present in sortase B. Additionally, the loop joining \( \beta_1 \) and \( \beta_2 \) strands is somewhat longer relative to other sortases. A \( \psi_0 \)-peptide (\( 165^G^P^166 \)) is present toward the C terminus in the \( \beta_6 \) strand resulting in an additional anti-parallel wide \( \beta \)-bulge following the anti-parallel classic \( \beta \)-bulge present in the \( \beta_6 \) strand of other sortase structures.

Description of active site

Catalysis in sortases is performed by a triad of highly conserved His, Cys, and Arg residues. The Cys residue is responsible for the formation of the thioacyl-enzyme intermediate; His acts as a proton donor for the leaving group, and Arg is believed to stabilize the intermediate and also facilitate the positioning of the substrate in the active-site cleft (31, 32). SavSrtE active-site residues, His-129 (C-terminal to B4), Cys-198 (C-terminal to \( \beta_7 \)), and Arg-207 (N-terminal to \( \beta_8 \)) are located at one edge of the \( \beta \)-barrel (Fig. 5, A and B). The walls of the active-site groove in SavSrtE are formed by residues from the B2/H1 loop, \( \beta_3/\beta_4 \) loop, \( \beta_4 \) strand, \( \beta_6/\beta_7 \) loop, \( \beta_7 \) strand, \( \beta_7/\beta_8 \) loop, and \( \beta_8 \) strand. Residue Cys-198, however, was associated with an additional blob of electron density that resembles a partial modification of Cys to (2-hydroxyethyl)thiocysteine (Cme,\(^6\)) out of the SG atom in the crystal structure) by \( \beta \)-mercaptoethanol present in the protein solution. The distance between the SG atom of the modified Cys and the ND1

---

**Figure 3. Influence of residues adjacent to the pentapeptide recognition motif.** The SavSrtE-catalyzed transpeptidation reactions were carried with respective LANTG or LPXG motif peptide (0.5 mM) and GGGKY (1 mM) with 50 \( \mu \)m enzyme at 20 °C for the indicated time. Each reaction mixture was subjected to RP-HPLC analysis as described under “Experimental procedures.” A, time course of transpeptidation reaction of LANTG motif: YALANTG (filled circle); YALANTGK (filled square); and YALANTE (filled triangles). B, comparative reaction time course of LAETG and LPETG peptide substrates; YNLAETGA (filled circles) and YNLPETGA (filled squares).

---

**Table 1**

| Description of active site | Value |
|---------------------------|-------|
| Resolution range (Å)      | 42.92–1.65 (1.68–1.65) |
| Space group               | P 3_{21}_{21} |
| Unit cell parameters      | a (Å) 85.84, c (Å) 48.20, \( \alpha = \beta = 90^\circ, \gamma = 120^\circ \) |
| Multiplicity              | 9.3 (9.4) |
| Total no. of observed reflections | 231,825 (11,672) |
| Unique reflections        | 24,903 (1245) |
| Rmerge (%)                | 8.5 (7.5) |
| No. of molecules in the asymmetric unit | 1 |
| Solvent content (%)       | 46.6 |
| Overall R-factor from Wilson plot (Å\(^2\)) | 26.5 |
| Rmerge/Rmerge (%)         | 16.12 / 19.08 |
| Root mean square bond lengths (Å) | 0.014 |
| Root mean square bond angles (°) | 1.58 |
| Average R-factors (Å\(^2\)) | 19.6 |
| Protein                   | 32.9 |
| Gly                       | 38.4 |
| Ramachandran map statistics (%) | 86.8 |
| Most favored region       | 13.2 |
| Additional allowed region | 0.0 |
| Generously allowed region | 0.0 |
| Outliers                  | 0.0 |

\(^6\) The abbreviations used are: Cme, S,S-(2-hydroxyethyl)thiocysteine; PDB, Protein Data Bank; Fmoc, N-(9-fluorenyl)methoxycarbonyl; RIN, residue interaction network; RP-HPLC, reverse-phase HPLC.

---

**Table 2**

| Summary of the data collection and processing statistics for SavSrtE (PDB code 5GO5, values are for the highest resolution shell) | Value |
|-------------------------------------------------------------|-------|
| Resolution range (Å)                                        | 42.92–1.65 (1.68–1.65) |
| Space group                                                 | P 3_{21}_{21} |
| Unit cell parameters                                        | a (Å) 85.84, c (Å) 48.20, \( \alpha = \beta = 90^\circ, \gamma = 120^\circ \) |
| Multiplicity                                                | 9.3 (9.4) |
| Total no. of observed reflections                           | 231,825 (11,672) |
| Unique reflections                                          | 24,903 (1245) |
| Rmerge (%)                                                  | 8.5 (7.5) |
| No. of molecules in the asymmetric unit                     | 1 |
| Solvent content (%)                                         | 46.6 |
| Overall R-factor from Wilson plot (Å\(^2\))                 | 26.5 |
| Rmerge/Rmerge (%)                                           | 16.12 / 19.08 |
| Root mean square bond lengths (Å)                           | 0.014 |
| Root mean square bond angles (°)                            | 1.58 |
| Average R-factors (Å\(^2\))                                 | 19.6 |
| Protein                                                     | 32.9 |
| Gly                                                         | 38.4 |
| Ramachandran map statistics (%)                             | 86.8 |
| Most favored region                                         | 13.2 |
| Additional allowed region                                   | 0.0 |
| Generously allowed region                                   | 0.0 |
| Outliers                                                    | 0.0 |

---

**Figure 3. Influence of residues adjacent to the pentapeptide recognition motif.** The SavSrtE-catalyzed transpeptidation reactions were carried with respective LANTG or LPXG motif peptide (0.5 mM) and GGGKY (1 mM) with 50 \( \mu \)m enzyme at 20 °C for the indicated time. Each reaction mixture was subjected to RP-HPLC analysis as described under “Experimental procedures.” A, time course of transpeptidation reaction of LANTG motif: YALANTG (filled circle); YALANTGK (filled square); and YALANTE (filled triangles). B, comparative reaction time course of LAETG and LPETG peptide substrates; YNLAETGA (filled circles) and YNL PETGA (filled squares).
Structure and specificity of a class E sortase

A

SauA.2kid
SpyA.3fn5
SagA.3rcoc
SagA.3rcb
SpyB.3paq
Bass.1r2
Bass.1r12
SauA.1ng5
SauA.1ng6
CofB.4u7
CofD.2w1j
CofD.2w1k
CofD.4d70
RaoD.21n7
SavE

B

SrtA
SrtB
SrtC
SrtD
SrtE

β1
α1
β3
β4
β1
β5
β6
β7
β8

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT

TT
atom of His-129 is 4.8 Å, compared with >5 Å in otherapo-
sortase A structures except for a value of 3.6 Å observed in
Bacillus anthracis sortase A (PDB code 2kw8).

**SavSrtE lacks a Ca$^{2+}$-binding site and contains a “preformed” pocket for substrate binding**

The β6/β7 loop in SavSrtA and other sortases has been
demonstrated to play an incisive role in substrate recognition
(32–34). In SavSrtA, this highly mobile loop undergoes confor-
mational transition from a “disordered open state” (PDB codes
1ja and 1t2p) to an “ordered closed conformation” (PDB code
2kid) upon substrate binding with the formation of a 3₁₀ helix
that is stabilized by a Ca$^{2+}$ ion (Fig. 5C). Analysis of the holo-
structure bound with Ca$^{2+}$ (PDB code 2kid) showed that Ca$^{2+}$
balances the electrostatic repulsion among residues Glu-105,
Glu-108, Asp-112 on the 3₁₀ loop, and Glu-171 following the 3₁₀

---

**Figure 5. Analysis of SavSrtE structure.** A, crystal structure of SavSrtE, with the catalytic residues shown as blue sticks (the dots represent part of the β7/β8 loop whose electron density is not observed). The insertion in the β6/β7 loop, containing a 3₁₀ helix, is highlighted by a dotted box. B, surface representation of SavSrtE, colored by B-factors from low (blue) to high (red). Active-site loops and catalytic residues are labeled. The first 3₁₀ helix is expected to interact with the bound peptide, has low B-factors, whereas the portion of the β7/β8 loop with observed electron density has high B-factors (the dots represent part of the β7/β8 loop for which interpretable electron density is not observed). C, β6/β7 loop in the apo-structure of S. aureus SrtA (SaSrtA, PDB code 1ja, pink) is in open conformation, whereas a 3₁₀ helix is induced upon Ca$^{2+}$ binding (green sphere), which closes over the bound substrate (not shown) in the holo-structure (PDB code 2kid, cyan). Comparison of apo-SavSrtE (blue) with holo-SaSrtA and apo-structure from B. anthracis SrtA (BaSrtA, PDB code 2kw8, orange) shows the 3₁₀ helix in β6/β7 loop and its closed conformation, resulting in a preformed binding pocket. D, SaSrtA contains a cluster of negatively charged residues which are stabilized by Ca$^{2+}$ binding. The equivalent residues in SavSrtE do not form a charged cluster that requires neutralization by Ca$^{2+}$ ion. Interactions of equivalent residues in BaSrtA involve a Lys in charge neutralization. E, electrostatic surface potential in SaSrtA showing the negatively charged residues, whereas SavSrtE has no such charged cluster. Surface electrostatics was calculated using Accelrys Discovery studio software. The color scheme ranges from blue (for electropositive regions) to red (for electronegative regions). F, effect of Ca$^{2+}$ on transpeptidation reaction. The reactions were carried out at 20 °C using 50 μM SavSrtE with YNLAEI GA (0.5 mM) and GGGKY (1 mM) in the absence and presence of Ca$^{2+}$ (5 mM). The reaction mixture was processed by RP-HPLC, and the product was quantified as described under “Experimental procedures.”

---

**Figure 4. Comparison of SavSrtE to sortases of classes A–D.** A, structure-based sequence alignment of SavSrtE (SavE in the figure) to other sortases of known structure, namely S. aureus SrtA (SaSrtA), S. pyogenes SrtA (SpyA), S. agalactiae SrtA (Saga), B. anthracis SrtA (BanA), S. pyogenes SrtB (SpyB), B. anthracis SrtB (BanB), S. aureus SrtB (SaSrtB), C. difficile SrtB (CdSrtB), S. pneumoniae SrtC1 (SpnC1), S. pneumoniae SrtC2 (SpnC2), S. pneumoniae SrtC3 (SpnC3), C. perfringens SrtD (CprD), and B. anthracis SrtD (BanD). The corresponding PDB codes are mentioned. The numerals 1–5 refer to the five different classes. Conserved residues are marked in white on a red background, and similar residues are in red. The secondary structure of S. aureus SrtA is shown. The catalytic residues are marked by red triangles; Ca$^{2+}$-binding residues in S. aureus SrtA are marked by blue triangles; and a Tyr residue in SavSrtE (Tyr-112), found to be conserved in all sortase E enzymes listed in the Sortase database, has been marked by an inverted brown triangle. B, structure of sortases (class A–E) depicting the eight-stranded β-barrel fold (bottom) and the corresponding topological diagram (top). Strands are in yellow; helices are in red, and loops are in green or black. The catalytic residues are shown as blue sticks. The structures depicted are S. aureus SrtA (PDB code 2kid), S. aureus SrtB (PDB code 1ng5), S. pneumoniae SrtC1 (PDB: 2w1j) with the Lid region that closes the active site in the absence of substrate colored in blue, B. anthracis SrtD (PDB code 2in7), and S. avermitilis SrtE (PDB code 5GOS).
helix in the $\beta_6/\beta_7$ loop (Fig. 5, D and E), thereby stabilizing the closed conformation of the loop (24). The equivalent residues in SavSrtE, namely Ala-113, Ala-116, Gln-120, and Pro-179, are neutral and interact via non-bonded interactions with their neighboring residues obviating the need for extraneous Ca$^{2+}$ ion. The first $3_1^\alpha$ helix in the $\beta_6/\beta_7$ loop of SavSrtE shows relatively lower $B$-factors compared with the second half of the $\beta_6/\beta_7$ loop (Fig. 5B), and it assumes an ordered closed conformation, indicating the presence of a preformed binding pocket as observed in SpySrtA and BaSrtA, which lack the calcium-binding site (35, 36). Consistent with the crystal structure data, the time course of the SavSrtE-catalyzed transpeptidation reaction carried out in the absence or presence of Ca$^{2+}$ ion (5 mM) under identical conditions generated almost super-imposable product yield curves (Fig. 5F).

**Binding of the second substrate**

Comparison of the apo- and substrate-bound structures of SaSrtA (PDB codes 1ija and 2kid) and BaSrtA (PDB codes 2kw8 and 2rui) showed that the $\beta_7/\beta_8$ loop undergoes a large conformational change upon substrate binding. The active site Cys residue is shifted away from the peptide displacing the $\beta_7/\beta_8$ loop and resulting in a second groove leading to the active site (Fig. 6, A and B). This second groove, formed by the $\beta_4/\beta_2$ loop, $\beta_2$ helix, and $\beta_7/\beta_8$ loop, has been suggested to be the binding site for the amine nucleophile (24). A similar groove is also observed in the SavSrtE structure. Interestingly, this groove contained additional electron density that could be modeled as a glycine residue (Fig. 6, C–E).

**SavSrtE shows strict specificity for Gly-based amine nucleophiles**

We further explored the propensity of AAKY, Tri-DAP (L-alanyl-γ-d-glutamyl-meso-diaminopimelic acid), and a pilin motif VHVPYKN subsequent sequence to residues 185–191 of FimA pilus protein (37) to undergo SavSrtE-catalyzed transpeptidation reaction with YNLAETGA peptide substrate (Fig. 6F). However, none of these peptide nucleophiles were able to generate any transpeptidation product. The presence of a small amount of YNLAET peptide was observed due to slow hydrolysis of the substrate in the absence of a productive amine nucleophile (data not shown). Thus, SavSrtE appears to display strict specificity for GGGKY as the second substrate.

**Modeling of protein-substrate interactions gives insight into preference for LAXTG peptide and implicates a role for Tyr-112**

In the absence of a substrate-bound structure of SavSrtE, we modeled the enzyme-substrate complex to understand the structural basis of substrate preference. The ALPNT or ALANT peptides were docked into the preformed binding pocket in the SavSrtE structure. The best docking pose for each ligand, as ascertained from the docking scores, contained the recognition motif placed in the active-site pocket of SavSrtE in an orientation similar to that observed in the peptide-bound SaSrtA structure (Fig. 7A). The side chain of Leu in the peptide is positioned in a hydrophobic cavity formed by the N-terminal stretch of the $\beta_6/\beta_7$ loop, whereas the succeeding Ala/Pro interacts with residues from the $\beta_2/\beta_1$ loop and $\beta_3/\beta_4$ loop.

The binding interaction was analyzed in terms of docking energy score and residue interaction network (Table 2). Lower docking energy scores for ALPNT binding to SaSrtA as compared with ALANT binding suggest the former as the preferred substrate for SaSrtA. Consistent with this, ALANT forms a stronger interaction network to SaSrtA than ALANT. Inspection of the docked structures shows that replacing bulky Pro by an Ala residue reduces the interaction of the peptide with Leu-169 in the $\beta_6/\beta_7$ loop.

In contrast, ALANT establishes a stronger interaction network than ALPNT in SavSrtE in accordance with the preference of Ala.
at the second position of the pentapeptide motif (ALANT) relative to Pro (ALPNIT). This may be partly ascribed to the presence of a bulky Tyr-112 as against an equivalent Ala residue (Ala-104) in SaSrtA, which is implicated in substrate recognition (38). Tyr-112 in SavSrtE presumably hinders the Pro-containing peptide from extending deeper into the active-site groove (Fig. 7, B and C), resulting in limited protein-peptide interactions. Besides, the –OH group of Tyr-112 is involved in a hydrogen bond with the backbone nitrogen of the Ala residue of the recognition motif in the ALANT, which is abolished in ALPNIT peptide. 

**Figure 7. Modeling protein-substrate interactions shows the importance of a Tyr residue at the active site.** The molecular surface of the protein has been rendered in gray, and the peptides are shown as sticks. Interactions of the peptides to the active site Arg are shown by black dots. Important active site residues have been labeled. A, ALPNT (cyan) peptide in the active site pocket of SaSrtA, modeled using the LPAT*-bound SaSrtA structure (PDB code 2kid). B, ALPNIT (cyan), and C, ALANT (green) peptides docked into the active-site pocket of SavSrtE. The presence of Tyr-112 hinders better binding of the Pro-containing peptide and favors the Ala-containing peptide. D, percent identity matrix showing the variation in sequence identity (in shades of gray; black indicates identity of 100%, and white indicates low identity) among SrtE sequences in the Sortase database. Sequences sharing higher identity have been grouped together. The sequences are as follows: 1) NP_789137 (Streptomyces coelicolor); 2) NP_787692 (Tropheryma whipplei); 3) cory_diphtheriae.6302c (Corynebacterium diphtheriae); 4) NP_739396 (Corynebacterium glutamicum); 5) NP_602126 (Corynebacterium efficiens); 6) thermo_fusca.3646 (Thermobifida fusca); 7) bifido_longum_DJO10A.7548; 8) NP_695779 (Bifidobacterium longum); 9) NP_825510 (SAV4333); 10) NP_826383 (SAV206); 11) NP_825514 (SAV4373); 12) NP_825513 (SAV4336) (Streptomyces avermitilis); 13) NP_628037; 14) NP_628038 (Streptomyces coelicolor). E, multiple sequence alignment of class E sortases depicting the presence of a Tyr (inverted brown triangle) equivalent to Tyr-112 at the active site of SavSrtE. The catalytic His residue is marked by a red triangle.

Mutational analyses reveal a critical role for Tyr-112 residue in substrate recognition and specificity discrimination

To probe the possible role of Tyr-112 in substrate recognition, we generated four mutants of SavSrtE by replacing Tyr by Ala, Gly, Trp, and Phe (Table 3). The choice of Ala and Phe was based on equivalent sites in SaSrtA and sortases of class D, respectively. Gly and Trp were chosen to gauge the effect of local flexibility or steric constraint on substrate recognition. The mutants were expressed and purified as described under "Experimental procedures," and their ability to catalyze transpeptidation reaction was explored with ALPNT* (Ala substrate) and YNLPETGA (Pro substrate) using GGGKY as the nucleophile. RP-HPLC profile of the reaction carried out with either Ala or Pro substrates with Tyr-112 mutants in which Tyr was replaced by Ala (Y112A), Gly (Y112G), or Trp (Y112W) did not show the presence of any transpeptidation or hydrolytic product peak. In contrast, Y112F mutant was active on both substrates, although to different extents relative to the wild-type enzyme (Fig. 8A). At equilibrium, the Y112F mutant yielded about ~18% product with the Pro, which was higher or comparable with the wild-type enzyme (~14%). However, product yield associated with the Ala substrate for Y112F was considerably diminished (~2% for Y112F versus 38% for wild type) indicating that Y112F mutation was particularly detrimental to the Ala substrate specificity of the enzyme (Fig. 8B).

**Kinetics of Y112F mutant with Ala and Pro substrates**

We carried out steady-state kinetic experiments to further characterize the role of the Tyr-112 residue in substrate recognition and specificity discrimination. Accordingly, kinetic parameters of wild-type SavSrtE and Y112F mutant were determined on the Ala substrate (acetyl-YNLAETGA) and Pro substrate (acetyl-YNLPEFGGA), respectively (Table 4). The individual values of $K_{cat}$ and $K_m$ with the Ala substrate for Y112F or wild-type SavSrtE was easily obtained by fitting the data to standard Michaelis-Menten kinetics or to the substrate inhibition model. However, both enzymes were subjected to inhibition by the Pro substrate above 3 mM and yielded an inhibition profile that did not fit well to the classical substrate inhibition model. Frankel et al. (39) have also observed such a complex inhibition of SaSrtA mutants with the Pro substrate and estimated $K_{cat}/K_m$ values of the Pro substrate at low concentrations from the slope of the linear portion of Lineweaver-Burk plot. We followed similar analyses to compare the relative specificity of the Pro substrate for wild-type and mutant sortase.

The Y112F mutation in SavSrtE resulted in a decrease of about 12-fold in $K_{cat}$ and an increase of 4-fold in $K_m$ for the Ala substrate yielding a $K_{cat}/K_m$ (1.03 ± 0.08 μM⁻¹ s⁻¹) that was about 50-fold less than the wild-type enzyme (50.94 ± 7.09 μM⁻¹ s⁻¹). In contrast, $K_{cat}/K_m$ of the mutant for the Pro substrate was found to be quite similar to the wild-type enzyme (8.60 versus 10.71 μM⁻¹ s⁻¹). The data suggest that Y112F mutation preferentially affected the Ala substrate specificity of SavSrtE.

We also generated kinetic parameters for SaSrtA with the above substrates to contrast the divergent substrate specificity of classical housekeeping class A sortase (SaSrtA) with class E sortase (SavSrtE) in view of the fact that SaSrtA is known to
Structure and specificity of a class E sortase

Table 2
Comparison of the residue preference of SavSrtE and SaSrtA at the second position of the recognition motif ascertained using docked models

| Parameters                      | SavSrtE | SaSrtA (PDB code 2kld) |
|---------------------------------|---------|------------------------|
| Docking energy score            | −34.7   | −2.53                  |
| Interaction network parameters  | 138     | 106                    |
| No. of edges (E)                | 22      | 18                     |
| No. of nodes (n)                | 6.3     | 5.9                    |
| E/n                             |         |                        |

* The docking score (provided by FiberDock) is not the absolute binding free energy but is proportional to it.

Table 3
Sequence of primers used for engineering Tyr-112 mutants of SavSrtE

| Mutants | Sequences (5′–3′) |
|---------|------------------|
| Y112A   | Forward, AAGAAGGGGTCCACCTGCCGCTCGGCCG
          | Reverse, TCTGGCCGAGCCCCTTCTT |
| Y112F   | Forward, AAGAAGGGGTCCACCTGCCGCTCGGCCG
          | Reverse, TCTGGCCGAGCCCCTTCTT |
| Y112G   | Forward, AAGAAGGGGTCCACCTGCCGCTCGGCCG
          | Reverse, TCTGGCCGAGCCCCTTCTT |
| Y112W   | Forward, AAGAAGGGGTCCACCTGCCGCTCGGCCG
          | Reverse, TCTGGCCGAGCCCCTTCTT |

Discussion

Development of novel sortases endowed with enhanced catalytic efficiency and newer specificity is a topical area of intense interest (19). The major motivation for this emanates from the immediate need to expand sortagging applications beyond the specificity restrictions of SaSrtA. In this endeavor, current efforts are centered on generating newer variants of SaSrtA employing rational protein engineering (34, 40) and directed evolution strategies (33, 38, 41). Although these approaches appear promising, contemporaneous mining of natural sortases offer rich possibilities of excavating useful enzymes with altered specificity, improved turnover, and stability. Here, we report the substrate specificity and crystal structure of a hitherto unexplored class E enzyme from *S. avermitilis*.

Our results demonstrating the preference of “LAXTG” over “LPXTG” sorting sequence (600-fold shift in Ala/Pro substrate specificity ratio from SaSrtA), the influence of the residue following the pentapeptide motif in the substrate recognition, and the calcium ion-independent catalytic activity of SavSrtE distinguish this enzyme from the prototype housekeeping sortase of *S. aureus*. Importantly, equilibrium product conversions obtained using corresponding substrates (SavSrtE, YALANTGGA, versus SaSrtA, YALPNTGK) under optimized conditions for SavSrtE (~40%) compares well with SaSrtA (~45%) corroborating the utility of SavSrtE in peptide ligation and protein engineering endeavors. This is important considering that sortases with robust activity are hard to find, and perhaps for this reason the activity of many new sortases is demonstrated through *in situ* detection of the product by mass spectrometric analysis of the reaction mixture. It is pertinent to mention here that Duong *et al.* (27) recently probed the substrate specificity of SrtE1 and SrtE2 of *S. coelicolor* by direct ES-MS analyses of the total reaction mixture and found multiple cleavages in the LAXTG sorting signal. Curiously, they reported predominant cleavage between the second (Ala) and the third residue (Xaa) rather than at the T-G peptide bond of the motif that is recognized by all sortases. We, however, did not see any trace of such aberrant proteolysis or transpeptidation indicating high specificity of SavSrtE for the scissile T-G peptide bond of the pentapeptide recognition motif.

The finding that the polar residue adjacent to the T-G scissile peptide bond of the sorting motif exerts detrimental influence on substrate recognition by SavSrtE is an important result that differs from the prototype housekeeping SaSrtA, which tolerates both Ala and Lys residues at this position equally well (26). Interestingly, residues juxtaposed to the LPXTG motif in pilin sortases are known to play a critical role in substrate discrimination by pilin and housekeeping sortases during pilus biogenesis (42). The neutral residues flanking pentapeptide sorting sequence may be relevant *in vivo* for performing some regulatory function.

The housekeeping sortases prefer either oligoglycine, oligoalanine, or both as the second substrate in the transpeptidation reaction. For example, SaSrtA prefers oligoglycine, and SrtA of *S. pyogenes* prefers oligoalanine (24, 35). Some housekeeping class A or other sortases might use diaminopimelic acid as has been experimentally demonstrated in the transpeptidation reaction catalyzed by SrtB of *Clostridium difficile* (43). In contrast, pilin sortases use an ε-amine of a lysine residue present in the “pilin motif” of pilus constituent proteins (44). The stringent specificity of aminoglycine for SavSrtE is consistent with the serendipitous presence of a bound Gly residue in the putative site for the second substrate (Fig. 6D). Interestingly, the docking program ClusPro (45) positioned the N-terminal Gly of a GGG peptide into the same pocket in a model of LANT-bound SavSrtE reinforcing this observation.

Modeling of the peptide substrate in the crystal structure of SavSrtE in the absence of an enzyme-substrate complex structure has been quite insightful and implicates a role for the Tyr-112 residue that is strategically located for interaction with Ala/Pro residue of the sorting sequence and may better complement the Ala side chain in the substrate as compared with...
Pro. This notion is further corroborated by the fact that Tyr-112 is conserved in all class E sortases (with LAXTG-containing substrates) in the Sortase database despite the lack of high sequence identity between them (Fig. 7, D and E), although it is not seen as conserved in the sequence alignment of sortases from all classes (Fig. 4A). The equivalent position in class D sortases, which are annotated to process an LP recognition motif, is generally occupied by a Phe residue (from structure-based sequence alignment of all sortases in the Sortase database). Analysis of B. anthracis SrtD (BaSrtD, PDB code 2ln7, NMR structure) shows that the side chain of an equivalent Phe-46 residue adopts a range of orientations of which only six are similar to the conformation of Tyr-112 in SavSrtE. There appears to be some flexibility at this Phe position in SrtD structures, whereas Tyr-112 in SavSrtE is found to be relatively rigid (all-atom B-factor, 13.2 Å²) and may cause partial steric occlusion of the Pro-containing substrate from the binding pocket. The experimental results showing the absence of activity in Y112A, Y112G, and Y112W mutants is consistent with this analysis. Furthermore, display of similar Pro substrate specificity as the wild-type enzyme concomitant with a 50-fold decrease in Ala substrate specificity in Y112F mutant indicates the importance of the hydrogen bond between the –OH group of Tyr-112 and the Ala residue of the LAETG motif. Phe-46 residue adopts a range of orientations of which only six are similar to the conformation of Tyr-112 in SavSrtE. There appears to be some flexibility at this Phe position in SrtD structures, whereas Tyr-112 in SavSrtE is found to be relatively rigid (all-atom B-factor, 13.2 Å²) and may cause partial steric occlusion of the Pro-containing substrate from the binding pocket. The experimental results showing the absence of activity in Y112A, Y112G, and Y112W mutants is consistent with this analysis. Furthermore, display of similar Pro substrate specificity as the wild-type enzyme concomitant with a 50-fold decrease in Ala substrate specificity in Y112F mutant indicates the importance of the hydrogen bond between the –OH group of Tyr-112 and the Ala residue of the LAXTG recognition motif.

The demonstration of Tyr-112 residue of SavSrtE as a critical determinant of specificity is instructive in view of the fact that no single residue of a sortase in isolation has been observed to exert considerable discriminatory influence on substrate preference. Previously reported SaSrtA variants evolved for improved activity toward the LAETG motif contained multiple mutations together with mutation of Ala-104 residue equivalent to Tyr-112 in SavSrtE (38). Interestingly, mutation of Ala-104 in isolation (single mutant) produced an inactive enzyme (29), but reversion of the 104 site to Ala in the multiple mutation setting resulted in severalfold gain of preference for the LPETG substrate (38). The retention of native-like Pro specificity concomitant with a severalfold loss of Ala specificity in a
Structure and specificity of a class E sortase

The single Y112F mutant suggests some degree of natural divergence of LPXTG to LAXTG specificity in class E sortase. SavSrtE is endowed with considerable propensity for LAXTG specificity and may be exploited as a suitable scaffold for directed evolution of sortases with enhanced catalytic efficiency.

In summary, the crystal structure and substrate specificity data of SavSrtE presented here represent the first detailed description of a class E sortase. SavSrtE produces useful transpeptidation yield and can fruitfully complement SaSrtA in protein engineering applications. The Ca\(^{2+}\)-independent activity together with preference for an altered substrate makes this enzyme an attractive tool for intracellular protein labeling. Besides, the SavSrtE crystal structure can form the basis for interrogation of substrate specificity and facilitate inhibitor design especially against the homologous class E sortases of pathogenic organisms.

Experimental procedures

Cloning, expression, and purification of SrtE3 (SavSrtE) from S. avermitilis

The gene corresponding to the full-length SrtE3 (residues 1–230) was custom-synthesized from GeneScript cloned in pUC57 vector. A gene encoding truncated version of the protein sequence corresponding to residues 51–230 was subcloned in pET28b(+) using appropriate primers (5'-CCCCCATATGAAGGGGACGGTGCCCGGA-3' and 5'-CCCCCTCTGAGCTAACGGCGTAGAGCCTC-3') nested with NdeI and XhoI restriction sites and amplified using pUC57 as template. The identity of the clone was established by DNA sequencing. The plasmid DNA isolated from the clone was used to transform BL21 (DE3) pLysS (Novagen) or BL21-CodonPlus-RP strain (Agilent) for protein expression. Subsequent to transformation, colonies of BL21 (DE3) pLysS (Novagen) or BL21-CodonPlus-RP cells were inoculated in 5 ml of LB medium containing 50 \(\mu\)g/ml kanamycin and allowed to grow overnight at 37 °C. This overnight grown culture was used as the seed for large scale culturing of cells.

The cells were allowed to grow at 37 °C until an optical density of 0.6 was attained at 600 nm. The protein expression was induced by the addition of 1 mM isopropyl \(\beta\)-D-thiogalactopyranoside, and the culture was grown for 10 h at 25 °C. Cells were harvested by centrifugation (12,000 rpm for 10 min), resuspended in 10 mM Tris buffer (pH 7) containing 40 mM NaCl and 10 mM imidazole, and lysed by sonication. The lysate was further clarified by centrifugation (10,000 rpm for 30 min), and the protein was purified using standard Ni-NTA affinity chromatography. The excess imidazole present in the eluted protein was removed using a PD-10 desalting column. The purity of the protein was checked by SDS-PAGE and mass spectrometry.

Site-directed mutagenesis

The above SavSrtE pET28b(+) clone was used as a template to introduce Y112A, Y112G, Y112W, and Y112F mutation by PCR with desired primers (Table 3) using QuickChange mutagenesis kit. Identity of the clone was confirmed by sequencing. The mutant proteins were expressed, purified, and characterized in much the same way as the wild-type SavSrtE.

Expression and purification of sortase A (SaSrtA) from S. aureus

The sequence comprising residues 60–204 of SaSrtA previously cloned in pET23b vector was expressed in BL-21 E. coli cells. The expressed protein was purified as described earlier (46).

Peptide synthesis

Peptides were synthesized using solid-phase methodology employing standard Fmoc/1-hydroxybenzotriazole/N,N-dicyclohexylcarbodiimide chemistry as described previously (26). The desired amino acids pre-loaded on Wang resin were used for elaboration of peptide sequences. After cleavage from the resin with 95% TFA, crude peptide was precipitated in cold ether and purified by RP-HPLC on a C18 column (Phenomenex Luna 10 μ, 250 × 30 mm) using acetonitrile/TFA/water solvent system. A linear gradient of 8–72% acetonitrile containing 0.1% TFA was used to effect the separation of the peptides. The mass of each peptide was confirmed using MALDI-TOF or ES-MS measurements.

Acetylation of the terminal amino group was carried out on the resin after completion of the final coupling and Fmoc deprotection step. Briefly, the dried resin was treated with a mixture of 10% acetic anhydride and 10% N,N-dimethylformamide in N,N-dimethylformamide for 1 h at room temperature. The resin was filtered, dried, processed for cleavage by TFA, subjected to purification, and mass spectrometric characterization as described above.

Transpeptidation assay

The SavSrtE-catalyzed transpeptidation reaction was carried out using relevant donor (peptides with LAXTG or LPXTG sortase-recognition motif) and acceptor (GGGKY) peptides (26). Initially, test assays were performed to determine the optimum pH and temperature of the reaction. Accordingly, assays were carried out in 100 mM Tris-HCl buffer (pH 7) containing 150 mM NaCl and 2 mM \(\beta\)-mercaptoethanol at 20 °C. The reaction was set up with appropriate amounts of peptide substrates and initiated by addition of known amounts of SavSrtE. The reaction was quenched at the desired time by adding 10-fold excess of 0.1% TFA and analyzed by analytical RP-HPLC (C18, 5 μ, 4.6 × 250 mm, linear gradient of 4–72% acetonitrile in 0.1% TFA). The product yield was estimated from the peak area using a software provided by the manufacturer (LCSolution, Shimadzu Corp., Japan). The product in each case was characterized by mass spectrometry (data not shown).

The above protocol was also followed for assaying SaSrtA activity except that the reaction was carried out at 37 °C in Tris–HCl buffer (pH 7.5) containing 5 mM calcium chloride, 150 mM NaCl, and 2 mM \(\beta\)-mercaptoethanol.

Steady-state kinetics

For measurement of steady-state kinetics, transpeptidation assays were carried out with varying concentrations of acetyl-YNLAEHTGA (Ala substrate) or acetyl-YNLPETGA (Pro substrate) against GGGKY fixed at 1 mM. All kinetic assays were performed ensuring linearity of the product formation. Fur-
Structure and specificity of a class E sortase

thermore, the product yields at all tested substrate concentrations were in the range of 0–10%. The enzyme concentration and transpeptidation assay times were as follows: 50 μM wild-type SavSrtE or Y112F mutant, 30 min for both Ala and Pro substrate; 25 μM SaSrtA, 30 min for Pro substrate, and 150 μM SaSrtA, 120 min for Ala substrate. The standard Michaelis-Menten kinetics (Equation 1) or modified model incorporating substrate inhibition (Equation 2) was used to analyze the data (GraphPad Prism 5, GraphPad Software, Inc.).

\[ v_0 = \frac{V_{\text{max}} \cdot [S]}{K_M + [S]} \]  

(Eq. 1)

where \( v_0 \) is the observed velocity at a given concentration; \( V_{\text{max}} \) is the apparent maximal velocity; \( [S] \) is the substrate concentration; \( K_M \) is the Michaelis-Menten constant; and \( K_i \) is the apparent inhibitor dissociation constant for substrate binding.

The data obtained with the Ala substrate for the SaSrtA (substrate concentration range, 0.25–17.5 mM) and the Y112F (range, 0.25–14 mM) mutant of SavSrtE were analyzed by standard Michaelis-Menten kinetics (Equation 1), and the inhibition model (Equation 2) was used for the analysis of kinetics data obtained with the Ala substrate (range, 0.25–18 mM) for wild-type SavSrtE. In the case of Pro substrate, \( K_{cat}/K_m \) values for both wild-type SavSrtE and Y112F mutant were ascertained from the linear portion of the Lineweaver-Burk plot.

Crystallization, diffraction data collection, and processing

The screen was prepared for crystallization conditions with commercially available screens using the hanging drop vapor diffusion method. Initial crystals were obtained in 2–3 days in a condition containing 2.0 M ammonium sulfate, 0.1 M citric acid (pH 3.5). Bigger crystals were grown in a drop containing 1 μl of protein (4 mg/ml in 10 mM Tris buffer (pH 7.2), 100 mM NaCl, and 2 mM β-mercaptoethanol) and 1 μl of a solution containing 1.6 M ammonium sulfate, 0.1 M citric acid (pH 3.75). The crystals were cryo-protected in a 10% sucrose solution, prepared by adding a concentrated stock of sucrose to the crystallization drop. Diffraction data were collected using synchrotron radiation (λ = 0.95372 Å) at the European Synchrotron Radiation Facility (BM-14, ESRF), Grenoble, France, with a CCD MAR-mosaic 225 detector. The crystal was annealed by blocking the mosaic 225 detector. The crystal was annealed by blocking the edge-to-node ratio (E/n) was calculated using 5% of the reflections set aside at the beginning of refinement procedure.

Structure analyses

The geometry of the refined model was analyzed using PROCHECK (56). Pairwise root mean square deviations between sortase structures were calculated using SSM superposed in CCP4. Structure-based multiple sequence alignment was performed using MUSTANG and rendered using ESPript 3.0 (57, 58). Structural motifs were studied from the PDBSUM output (59). Surface electrostatics was calculated using Accelrys Discovery studio software.

Modeling protein-substrate interactions from peptide docking

Hetero-atoms and water molecules were removed from the SavSrtE crystal structure, and the missing residues in the protein were modeled using COOT (55). The peptides YALANTGA and YALPNTGA were generated using the coordinates of LPAT in the substrate-bound structure (PDB code 2kid) of S. aureus SrtA (SaSrtA as scaffold), and additional residues in the peptide were added using COOT (24, 55). The substrate-bound structures of SavSrtE were generated using the ClusPro server that generates multiple docked conformations using an FFT-based rigid body docking program PIPER (60). Residues from the putative binding site region of SavSrtE were provided to the program through the “attraction” option. The docked conformations were filtered by ClusPro using an empirical potential, clustered by root mean square deviation, ranked by cluster size, and subjected to a minimization procedure to remove side chain clashes. The energy score for the best docked conformation for each peptide, obtained from the top ranked cluster, was obtained using the FiberDock server (61). To compare the substrate preference of SavSrtE with SaSrtA, ALANT- and ALANT-bound models of SaSrtA were generated by appropriate modifications in the peptide-bound NMR structure (PDB code 2kid), and the structures were energy-minimized.

Interactions between the protein and the peptide were studied from residue interaction networks (RINs) generated by the RINerator (62). Hydrogen atoms were added to the peptide-bound protein model using “reduce,” and the non-covalent interactions (van der Waals interaction, hydrogen bond, and atomic clashes) were identified by Probe (63, 64). Interactions with a positive score (>0) were studied, and a higher score indicated stronger interaction between the connected residue pair. The RIN consisted of nodes (n) representing residues connected by edges (E), which represent the interaction between them. The edges were weighted by the interaction score. A sub-network of the protein residues directly interacting with the peptide was derived using Cytoscape and the edge-to-node ratio (E/n) was calculated (65). A higher E/n ratio is an indicator of a stronger interaction network (enhanced interaction).
**Structure and specificity of a class E sortase**

**Author contributions**—R. P. R. and S. R. conceptualized the study. S. D., V. S. P., A. K. S., and V. D. performed the experiments. S. D., V. S. P., R. P. R., and S. R. wrote the paper. All the authors reviewed and approved the manuscript.

**Acknowledgment**—We thank Dr. S. Kumaran for helpful discussions and Sushma Nagpal for technical help with peptide synthesis. The National Institute of Immunology was supported by core grants from the Department of Biotechnology (Government of India).

**References**

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

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

3. Mazmanian, S. K., Liu, G., Ton-That, H., and Schneewind, O. (1999) *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science* 285, 760–763

4. Novick, R. P. (2000) Sortase: the surface protein anchoring transpeptidase and the LPXTG motif. *Trends Microbiol.* 8, 148–151

5. Lévesque, C. M., Voronejskaia, E., Huang, Y. C., Mair, R. W., Ellen, R. P., and Cvitkovitch, D. G. (2005) Involvement of sortase anchoring of cell wall proteins in biofilm formation by *Streptococcus mutans*. *Infect. Immun.* 73, 3773–3777

6. Manetti, A. G., Zingaretti, C., Falugi, F., Capo, S., Bombaci, M., Bagnoli, F., Gambellini, G., Bensi, G., Mora, M., Edwards, A. M., Musser, J. M., Graviss, E. A., Telford, J. L., Grandi, G., and Margariti, I. (2007) *Streptococcus pyogenes* pili promote pharyngeal cell adhesion and biofilm formation. *Mol. Microbiol.* 64, 968–983

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

8. Khara, 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

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

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

11. Marraffini, L. A., Dedent, A. C., and Schneewind, O. (2008) Sortase as a target of anti-infective therapy. *Pharmacol. Rev.* 60, 128–141

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

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

15. Bradshaw, W. J., Davies, A. H., Chambers, C. J., Roberts, A. K., Shone, C. C., and Acharya, K. R. (2015) Molecular features of the sortase enzyme family. *FEBS J.* 282, 2097–2114

17. Kruger, R. G., Otvos, B., Frankel, B. A., Bentley, M., Dostal, P., and McCafferty, D. G. (2004) Analysis of the substrate specificity of the *Staphylococcus aureus* sortase transpeptidase SrtA. *Biochemistry* 43, 1541–1551

18. Popp, M. W., and Ploegh, H. L. (2011) Making and breaking peptide bonds: protein engineering using sortase. *Angew. Chem. Int. Ed. Engl.* 50, 5024–5032

19. Antos, J. M., Truttmann, M. C., and Ploegh, H. L. (2016) Recent advances in sortase-catalyzed ligation methodology. *Curr. Opin. Struct. Biol.* 38, 111–118

20. Antos, J. M., Chew, G. L., Guimaraes, C. P., Yoder, N. G., Grotenberg, G. M., Popp, M. W., and Ploegh, H. L. (2009) Site-specific N- and C-terminal labeling of a single polypeptide using sortases of different specificity. *J. Am. Chem. Soc.* 131, 10800–10801

21. Matsumoto, T., Takase, R., Tanaka, T., Fukuda, H., and Kondo, A. (2012) Site-specific protein labeling with amine-containing molecules using *Lactobacillus plantarum* sortase. *Biotechnol. J.* 7, 642–648

23. Strijbis, K., Spooner, E., and Ploegh, H. L. (2012) Protein ligation in living cells using sortase. *Traffic* 13, 780–789

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

25. Biswas, T., Pawale, V. S., Choudhury, D., and Roy, R. P. (2014) Sorting of LPXTG peptides by archetypical sortase A: role of invariant substrate residues in modulating the enzyme dynamics and conformational signature of a productive substrate. *Biochemistry* 53, 2515–2524

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

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

30. Piotukh, K., Geltinger, B., Heinrich, N., Gerth, F., Beyermann, M., Freund, C., and Schwarzer, D. (2011) Directed evolution of sortase A mutants with altered substrate selectivity profiles. *Traffic* 12, 1363–1376

31. Frankel, B. A., Kruger, R. G., Robinson, D. E., Kelleher, N. L., and McCafferty, D. G. (2005) *Staphylococcus aureus* sortase transpeptidase SrtA: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* 367, 169–184

32. Frankel, B. A., Kruger, R. G., Robinson, D. E., Kelleher, N. L., and McCafferty, D. G. (2005) *Staphylococcus aureus* sortase transpeptidase SrtA: the β6/β7 loop from SrtB confers NPQTN recognition to SrtA. *J. Biol. Chem.* 280, 6924–6933

33. Ploegh, H. L. (2001) Enzyme mechanism of sortase. *Cell* 104, 883–895
Structure and specificity of a class E sortase

36. Weiner, E. M., Robson, S., Marohn, M., and Clubb, R. T. (2010) The sortase A enzyme that attaches proteins to the cell wall of Bacillus anthracis contains an unusual active site architecture. J. Biol. Chem. 285, 23433–23443

37. Mishra, A., Das, A., Cisar, I. O., and Ton-That, H. (2007) Sortase-catalyzed assembly of distinct heteromeric fibriniae in Actinomyces naeslundii. J. Bacteriol. 189, 3156–3165

38. Dorr, B. M., Ham, H. O., An, C., Chaiok, E. L., and Liu, D. R. (2014) Reprogramming the specificity of sortase enzymes. Proc. Natl. Acad. Sci. U.S.A. 111, 13343–13348

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

40. Hirakawa, H., Ishikawa, S., and Nagamune, T. (2012) Design of Ca²⁺-independent Staphylococcus aureus sortase A mutants. Biotechnol. Bioeng. 109, 2952–2961

41. Chen, L., Dorr, B. M., and Liu, D. R. (2011) A general strategy for the evolution of bond-forming enzymes using yeast display. Proc. Natl. Acad. Sci. U.S.A. 108, 11399–11404

42. Douillard, F. P., Rasinkangas, P., von Ossowski, I., Reunanen, J., Palva, A., and de Vos, W. M. (2014) Functional identification of conserved residues involved in Lactobacillus rhamnosus strain GG sortase specificity and pilus biogenesis. J. Biol. Chem. 289, 15764–15775

43. Chambers, C. J., Roberts, A. K., Shone, C. C., and Acharya, K. R. (2015) Structure and function of a Clostridium difficile sortase enzyme. Sci. Rep. 5, 9449

44. Dasgupta, S., Samantaray, S., Sahai, D., and Roy, R. P. (2011) Isopeptide ligation catalyzed by quintessential sortase A: mechanistic cues from cyclic and branched oligomers of indolocidin. J. Biol. Chem. 286, 23996–24006

45. Comeau, S. R., Gatchell, D. W., Vajda, S., and Camacho, C. J. (2004) ClusPro: a fully automated algorithm for protein-protein docking. Nucleic Acids Res. 32, W96–W99

46. Samantaray, S., Marathe, U., Dasgupta, S., Nandicoori, V. K., and Roy, R. P. (2008) Peptide-sugar ligation catalyzed by transpeptidase sortase: a facile approach to neoglycoconjugate synthesis. J. Am. Chem. Soc. 130, 2132–2133

47. Leslie, A. G., and Powell, H. R. (2007) Processing diffraction data with Mosflm. Evolving Methods for Macromol. Crystallogr. 245, 41–51

48. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. 69, 1204–1214

49. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242

50. Kantardjeff, K. A., and Rupp, B. (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. Protein Sci. 12, 1865–1871

51. Matthews, B. W. (1968) Solvent content of protein crystals. J. Mol. Biol. 33, 491–497

52. Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S., and Tucker, P. A. (2005) Auto-rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. Acta Crystallogr. D Biol. Crystallogr. 61, 449–457

53. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674

54. Murshudov, G. N., Skubák, P., Lebedev, A. A., Panuw, N. S., Steiner, R. A., Nichols, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. 67, 355–367

55. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501

56. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26, 283–291

57. Konagurthu, A. S., Whisstock, J. C., Stucy, P. J., and Lesk, A. M. (2006) MUSTANG: a multiple structural alignment algorithm. Proteins 64, 559–574

58. Gouet, P., Robert, X., and Courcelle, E. (2003) ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. 31, 3320–3323

59. Laskowski, R. A. (2001) PDBsum: summaries and analyses of PDB structures. Nucleic Acids Res. 29, 221–222

60. Kozakov, D., Brenke, R., Comeau, S. R., and Vajda, S. (2006) PIPER: an FFT-based protein docking program with pairwise potentials. Proteins 65, 392–406

61. Maschii, E., Nussinov, R., and Wolfson, H. J. (2010) FiberDock: a web server for flexible induced-fit backbone refinement in molecular docking. Nucleic Acids Res. 38, W457–W461

62. Doncheva, N. T., Klein, K., Domingues, F. S., and Albrecht, M. (2011) Analyzing and visualizing residue networks of protein structures. Trends Biochem. Sci. 36, 179–182

63. Word, J. M., Lovell, S. C., Richardson, J. S., and Richardson, D. C. (1999) Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J. Mol. Biol. 285, 1735–1747

64. Word, J. M., Lovell, S. C., LaBean, T. H., Taylor, H. C., Zalis, M. E., Presley, B. K., Richardson, J. S., and Richardson, D. C. (1999) Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. J. Mol. Biol. 285, 1711–1733

65. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504