The Structure of the \textit{Staphylococcus aureus} Sortase-Substrate Complex Reveals How the Universally Conserved LPXTG Sorting Signal Is Recognized

Received for publication, May 18, 2009, and in revised form, June 16, 2009. Published, JBC Papers in Press, July 10, 2009, DOI 10.1074/jbc.M109.022624

Nuttee Suree\textsuperscript{1,5}\textsuperscript{,}, Chu Kong Liew\textsuperscript{2}, Valerie A. Villareal\textsuperscript{1,5,}, William Thieu\textsuperscript{6}, Evgeny A. Fadeev\textsuperscript{1,5,4}, Jeremy J. Clemens\textsuperscript{1,5}, Michael E. Jung\textsuperscript{1}, and Robert T. Clubb\textsuperscript{1,5,6}

From the \textsuperscript{1}Department of Chemistry and Biochemistry and the \textsuperscript{2}UCLA-Department of Energy Institute for Genomics and Proteomics, UCLA, Los Angeles, California 90095-1570

In Gram-positive bacteria, sortase enzymes assemble surface proteins and pili in the cell wall envelope. Sortases catalyze a transpeptidation reaction that joins a highly conserved LPXTG sorting signal within their polypeptide substrate to the cell wall or to other pilin subunits. The molecular basis of transpeptidation and sorting signal recognition are not well understood, because the intermediates of catalysis are short lived. We have overcome this problem by synthesizing an analog of the LPXTG signal whose stable covalent complex with the enzyme mimics a key thioacyl catalytic intermediate. Here we report the solution structure and dynamics of its covalent complex with the \textit{Staphylococcus aureus} SrtA sortase. In marked contrast to a previously reported crystal structure, we show that SrtA adaptively recognizes the LPXTG sorting signal by closing and immobilizing an active site loop. We have also used chemical shift mapping experiments to localize the binding site for the triglycine portion of lipid II, the second substrate to which surface proteins are attached. We propose a unified model of the transpeptidation reaction that explains the functions of key active site residues. Since the sortase-catalyzed anchoring reaction is required for the virulence of a number of bacterial pathogens, the results presented here may facilitate the development of new anti-infective agents.

In vitro studies of SrtA have begun to define the mechanism of transpeptidation. SrtA consists of two parts: an unstructured amino-terminal tail that contains a stretch of nonpolar residues and the catalytic domain that competently performs the transpeptidation reaction of an amide bond between the carboxyl group of the threonine and the cell wall precursor molecule lipid II (undecaprenyl-pyrophosphate-MurNAc(-L-Ala-D-iGln-L-Lys(NH\textsubscript{2}-Gly\textsubscript{3})-D-Ala-D-Ala)-\textbeta-1-4-GlcNAc)), creating a protein-lipid II-linked product that is incorporated into the peptidylglycan via the transglycosylation and transpeptidation reactions of bacterial cell wall synthesis (18–20). Over 900 sortase-attached proteins in 72 different strains of bacteria have thus far been identified (21, 22). The vast majority of these proteins contain a COOH-terminal sorting signal harboring an LPXTG motif and are anchored to the cell wall by enzymes closely related to SrtA.

Bacterial surface proteins function as virulence factors that enable pathogens to adhere to sites of infection, evade the immune response, acquire essential nutrients, and enter host cells (1). Gram-positive bacteria use a common mechanism to covalently attach proteins to the cell wall. This process is catalyzed by sortase transpeptidase enzymes, which join proteins bearing a highly conserved Leu-Pro-X-Thr-Gly (LPXTG, where X is any amino acid) sorting signal to the cross-bridge peptide of the peptidylglycan (2–4). Sortases also polymerize proteins containing sorting signals into pili, filamentous surface exposed structures that promote bacterial adhesion (5, 6). The search for small molecule sortase inhibitors is an active area of research, since these enzymes contribute to the virulence of a number of important pathogens, including among others \textit{Staphylococcus aureus}, \textit{Listeria monocytogenes}, \textit{Streptococcus pyogenes}, and \textit{Streptococcus pneumoniae} (reviewed in Refs. 7 and 8). Sortase enzymes are also promising molecular biology reagents that can be used to site-specifically attach proteins to a variety of biomolecules (9–14, 72).

The sortase (SrtA)\textsuperscript{3} enzyme from \textit{S. aureus} is the prototypical member of the sortase enzyme family (15, 16). It anchors proteins to the murein sacculus that possess a COOH-terminal cell wall sorting signal that consists of a LPXTG motif, followed by a hydrophobic segment of amino acids and a tail composed of mostly positively charged residues (17). SrtA is located on the extracellular side of the membrane. After partial secretion of its protein substrate across the cell membrane, SrtA cleaves the LPXTG motif between the threonine and glycin residues, forming a thioacyl-linked protein-sortase intermediate (16). It then catalyzes the formation of an amide bond between the carboxyl group of the threonine and the cell wall precursor molecule lipid II (undecaprenyl-pyrophosphate-MurNAc(-L-Ala-D-iGln-L-Lys(NH\textsubscript{2}-Gly\textsubscript{3})-D-Ala-D-Ala)-\textbeta-1-4-GlcNAc)), creating a protein-lipid II-linked product that is incorporated into the peptidylglycan via the transglycosylation and transpeptidation reactions of bacterial cell wall synthesis (18–20). Over 900 sortase-attached proteins in 72 different strains of bacteria have thus far been identified (21, 22). The vast majority of these proteins contain a COOH-terminal sorting signal harboring an LPXTG motif and are anchored to the cell wall by enzymes closely related to SrtA.

In vitro studies of SrtA have begun to define the mechanism of transpeptidation. SrtA consists of two parts: an unstructured amino-terminal tail that contains a stretch of nonpolar residues that embed it in the membrane and an autonomously folded catalytic domain that competently performs the transpeptidation reaction \textit{in vitro} (SrtA\textsubscript{A19999}, residues 60–206) (16, 23–25).

\textsuperscript{§}The abbreviations used are: SrtA, sortase A; Cbz, benzoyloxycarbonyl; T*, (2R,3S)-3-amino-4-mercapto-2-butanol analog of threonine that replaces the carbonyl group with -CH\textsubscript{2}-SH; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; r.m.s., root mean square.

\textsuperscript{1}Recipient of DPST and JSTP scholarships from the Royal Thai Government.

\textsuperscript{2}Present address: Dept. of Molecular Cardiology and Biophysics, The Victor Chang Cardiac Research Institute, Level 6, 384 Victoria St., Darlinghurst, New South Wales 2010, Australia.

\textsuperscript{3}Present address: Department of Chemistry and Biochemistry, 611 Charles E. Young Dr., Los Angeles, CA 90095-1570. Tel.: 310-206-2334; Fax: 310-206-4749; E-mail: rclubb@mbi.ucla.edu.

\textsuperscript{4}Present address: University of California, Irvine, CA 92697.

\textsuperscript{5}Present address: Wyeth Research, 401 N. Middletown Rd., Pearl River, NY 10965.

\textsuperscript{6}To whom correspondence should be addressed: Dept. of Chemistry and Biochemistry, 611 Charles E. Young Dr., Los Angeles, CA 90095-1570. Tel.: 310-206-2334; Fax: 310-206-4749; E-mail: rclubb@mbi.ucla.edu.

\textsuperscript{§}The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. S1 and Table S1.

\textsuperscript{§}Supported by National Institutes of Health, NIGMS, Grant F31GM075564.

\textsuperscript{§}The abbreviations used are: SrtA, sortase A; Cbz, benzoyloxycarbonyl; T*, (2R,3S)-3-amino-4-mercapto-2-butanol analog of threonine that replaces the carbonyl group with -CH\textsubscript{2}-SH; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; r.m.s., root mean square.

\textsuperscript{§}This work was supported, in whole or in part, by National Institutes of Health Grant AI52217 (to R. T. C. and M. E. J.).
NMR Structure of the Covalent Sortase-sorting Signal Complex

Catalysis occurs through a ping-pong mechanism that is initiated when the thiol group of amino acid Cys\textsuperscript{184} nucleophilically attacks the carbonyl carbon of the threonine residue within the LPXTG sorting signal (16, 23–25). This forms a transient tetrahedral intermediate that, upon breakage of the threonine-glycine peptide bond, rearranges into a more stable thioacyl enzyme-substrate linkage. SrtA then joins the terminal amine group within the pentaglycine branch of lipid II to the carbonyl carbon of the threonine, creating a second tetrahedral intermediate that is resolved into the lipid II-linked protein product (23).

Sortase enzymes contain three conserved residues within their active sites: His\textsuperscript{120}, Cys\textsuperscript{184}, and Arg\textsuperscript{197} (SrtA numbering). These residues play a critical role in catalysis, since their mutation in SrtA causes severe reductions in enzyme activity (16, 26–30). Although it is well established that Cys\textsuperscript{184} forms a covalent linkage to the sorting signal, the functions of His\textsuperscript{120} and Arg\textsuperscript{197} are controversial. A variety of disparate functions have been ascribed to Arg\textsuperscript{197}. These include deprotonating Cys\textsuperscript{184} (28), deprotonating lipid II (31), or stabilizing the binding of the LPXTG sorting signal (28, 32) or oxyanion intermediates (31, 32). Different functions have also been proposed for His\textsuperscript{120}. Originally, it was suggested that it activated Cys\textsuperscript{184} by forming an imidazolium-thiolate ion pair (26). However, subsequent ρ\textsubscript{Ka} measurements revealed that both His\textsuperscript{120} and Cys\textsuperscript{184} are predominantly uncharged at physiological pH values, leading to the suggestion that His\textsuperscript{120} functions as a general base during catalysis (33). Most recently, it has been proposed that the most active form of the enzyme contains His\textsuperscript{120} and Cys\textsuperscript{184} in their charged states but that only a small fraction of SrtA exists in this form (~0.06%) prior to binding to the sorting signal (25).

NMR and crystal structures of SrtA\textsubscript{AN59} have revealed that it adopts an eight-stranded β-barrel fold (31, 34). Other sortase enzymes have also been shown to possess a similar overall structure, including SrtB from \textit{S. aureus} (27, 35), SrtB from \textit{Bacillus anthracis} (27, 36), SrtA from \textit{S. pyogenes} (37), and the SrtC-1 and SrtC-3 enzymes from \textit{S. pneumoniae} (38). However, the molecular basis of substrate recognition remains poorly understood, because all of the structures reported to date have not contained a sorting signal bound to the enzyme. The lone exception is the crystal structure of SrtA\textsubscript{AN59} bound to an LPETG peptide (31). However, in this structure the peptide substrate is bound nonspecifically (see below) (32, 39).

In this paper, we report the structure and dynamics of SrtA covalently bound to an analog of the LPXTG sorting signal. The structure of the complex resembles the thioacyl intermediate of catalysis, providing insights into the molecular basis of binding of the LPXTG sorting signal and the functions of key active site residues. Notably, the mechanism of substrate binding visualized in the NMR structure differs substantially from a previously reported crystal structure of SrtA\textsubscript{AN59} non-covalently bound to a LPETG peptide (31). We have also used NMR chemical shift mapping experiments to localize the binding site for a triglycerine cell wall substrate analog. A mechanism of transpeptidation compatible with these new data is proposed.

EXPERIMENTAL PROCEDURES

\textbf{Preparation of the Covalent Complex for NMR Studies—}\textit{Wild-type SrtA from \textit{S. aureus} containing amino acid residues 60–206 (SrtA\textsubscript{AN59}) was produced as described previously (29). Uniformly \textsuperscript{15}N- and \textsuperscript{13}C- or \textsuperscript{15}N-labeled SrtA\textsubscript{AN59} protein was covalently attached to an analog of the LPXTG sorting signal, Cbz-LPAT* (where * T is (2R,3S)-3-amino-4-mercapto-2-butanol, and Cbz is a carbobenzyloxy protecting group). The methods used to synthesize the analog and to prepare its covalent complex with SrtA\textsubscript{AN59} have been described previously (40). Three ~1 mM samples of the complex were studied. Each was dissolved in 50 mM Tris-HCl (pH 6.0), 100 mM NaCl, 20 mM CaCl\textsubscript{2}, and 0.01% NaN\textsubscript{2}. The complexes contained either 1) \textsuperscript{15}N-labeled SrtA\textsubscript{AN59} bound to the unlabeled peptide dissolved in H\textsubscript{2}O (7% \textsuperscript{2}H\textsubscript{2}O), 2) \textsuperscript{13}C,\textsuperscript{15}N-labeled protein bound to the unlabeled peptide dissolved in H\textsubscript{2}O (7% \textsuperscript{2}H\textsubscript{2}O), or 3) \textsuperscript{13}C,\textsuperscript{15}N-labeled protein bound to the unlabeled peptide dissolved in \textsuperscript{2}H\textsubscript{2}O.}

\textbf{NMR Spectroscopy and Structure Determination—}\textit{NMR spectra were acquired at 302 K on Bruker Avance 500-, 600-, and 800-MHz spectrometers equipped with triple resonance cryogenic probes. \textsuperscript{1}H, \textsuperscript{13}C, and \textsuperscript{15}N protein chemical shift assignments were obtained using standard methods (41, 42). Chemical shift assignments for the sorting signal were obtained by analyzing two-dimensional (F1,F2) \textsuperscript{13}C-filtered NOESY (43) and (F1) \textsuperscript{13}C-filtered TOCSY (44) spectra. Distance restraints to define the structure of the protein were obtained from three-dimensional \textsuperscript{15}N- and \textsuperscript{13}C-edited NOESY spectra (mixing time 100 ms), and intramolecular restraints for the sorting signal analog were obtained by analyzing two-dimensional (F1,F2) \textsuperscript{13}C-filtered NOESY spectra. Intermolecular distance restraints between SrtA\textsubscript{AN59} and the bound peptide were identified in three-dimensional (F1) \textsuperscript{13}C,\textsuperscript{15}N-filtered (F2) \textsuperscript{13}C-edited NOE\textsuperscript{HSQC} and (F1) \textsuperscript{13}C,\textsuperscript{15}N-filtered (F2) \textsuperscript{15}N-edited NOESY-\textsuperscript{HSQC} spectra (45) and in a two-dimensional (F1) \textsuperscript{13}C-edited NOE\textsuperscript{HSQC} spectrum (43). \textsuperscript{3}J\textsubscript{\textit{HN-HA}} values were measured from a water flip-back three-dimensional HNHA spectrum (46), and backbone \(\psi\) and \(\phi\) dihedral angles were obtained using the program TALOS (47). The NMR data were processed using NMRPipe (48) and analyzed using the PIPP (49) and CARA (version 1.4.1) (50) software packages.}

Structure calculations were performed using the ATNOS/\textit{CANDID} and NIH-XPLOR programs (51–53). Three three-dimensional NOE\textsuperscript{HSQC} data sets were used as input for ATNOS/\textit{CANDID}: three-dimensional \textsuperscript{13}C-edited NOE\textsuperscript{HSQC} and \textsuperscript{15}N-edited NOE\textsuperscript{HSQC} spectra of the complex dissolved in H\textsubscript{2}O and a \textsuperscript{13}C-edited NOE\textsuperscript{HSQC} spectrum of the complex dissolved in \textsuperscript{2}H\textsubscript{2}O. These data were supplemented with restraints for the backbone and side chain dihedral angles. Seven cycles of ATNOS/\textit{CANDID} calculations yielded a converged ensemble of the protein in the complex. The structure was then refined in an iterative manner by manually checking all of the NOEs assigned by \textit{CANDID} and by including \textsuperscript{3}J\textsubscript{\textit{HN-HA}} couplings and carbon chemical shifts in the calculations. New manually identified intra- and intermolecular distance restraints were also added, and at the final stages of refinement, hydrogen bonds were identified and included as distance
obtained from Sigma. A 500 mM concentrated stock solution of pET15b expression vector. The presence of the histidine tag were performed in an identical manner.

Site-directed Mutagenesis and Enzyme Assays—Five single amino acid mutants of SrtAΔN59 containing a COOH-terminal six-histidine tag were produced in *Escherichia coli* from a pET15b expression vector. The presence of the histidine tag

restraints. The latter were obtained by inspecting the structures of the complex and by identifying NOE patterns characteristic of distinct secondary structures. All of the NOEs within the active site and binding pocket were manually checked. Residual dipolar couplings were measured by taking the difference in J couplings between partially aligned and unaligned protein samples. The protein was aligned using 15% (w/v) charged bicelles (30:10:1 molar ratio of dimyristoylphosphatidylcholine/dihexanoyl-phosphatidylcholine/hexadecyl(cetyl)trimethylammonium bromide). The programs MOLMOL (54) and PyMOL (55) were used to generate figures.

**NMR Relaxation Studies**—NMR data were collected using the 15N-labeled sample of the complex at 600 MHz. The strategy used to collect and analyze the relaxation data has been described previously (56, 57). The well resolved 1H-15N HSQC spectrum enabled the reliable measurement of relaxation parameters for 86 of a total of 148 residues. The average relaxation time of 8.56 ns, and the axial symmetric model was preferred over the isotropic model. A total of 76 of 86 quantifiable residues could be fit satisfactorily. The data from Ala81, Val87, Tyr88, Arg99, Ser102, Asn132, Ala135, Lys137, Met155, and Asp165 could not be fit to any model, possibly because they undergo more complicated motions. Residues were classified as follows: model 1 (S-only) was selected for 52 residues, 2 residues fit to model 2 (S and τ), 15 residues fit to model 3 (S and Rτs), 1 residue fit to model 4 (Sτ, S, and Rτs), and 6 residues fit to model 5 (Sτ, Sτ, and τ). For residues located in regions of regular secondary structure, the average order parameter is 0.93 ± 0.01. The relaxation data were analyzed using the suite of analysis programs kindly provided by Prof. Arthur G. Palmer III (60–63).

**Localization of the Lipid II Binding Pocket**—These studies made use of a 1 mM sample of the complex containing 15N-labeled SrtAΔN59 covalently attached to the sorting signal analog (NMR buffer: 50 mM Tris-HCl, 100 mM NaCl, 20 mM CaCl2, 0.01% NaN3, and 7% 2H2O, pH 6.0). Triglycine (Gly3) was dissolved in NMR buffer was used. A series of two-dimensional 1H-15N HSQC spectra were recorded at 302 K after the addition of small aliquots of Gly3. A total of 12 spectra were acquired with the following molar ratios of Gly3 to the SrtAΔN59-LPAT* complex: 0:1, 0.5:1, 1:1, 2:1, 3:1, 4:1, 8:1, 16:1, 40:1, 50:1, 80:1, and 100:1. No significant chemical shift changes were observed after a 40:1 molar ratio was achieved. A normalized chemical shift change (Δδ) was calculated as Δδ = ((ΔδN)2 + (ΔδH/6.493))^1/2, where ΔδN and ΔδH are, respectively, the amide nitrogen and proton chemical shift difference for a given residue in the presence and absence of Gly3. The titration experiments using the apo-form of the enzyme were performed in an identical manner.

**Site-directed Mutagenesis and Enzyme Assays**—Five single amino acid mutants of SrtAΔN59 containing a COOH-terminal six-histidine tag were produced in *Escherichia coli* from a pET15b expression vector. The presence of the histidine tag
NMR Structure of the Covalent Sortase-sorting Signal Complex

![Diagram](http://www.jbc.org/content/suppl/2009/07/10/M109.022624.DC1.html)

**FIGURE 2.** NMR solution structure of the SrtA<sub>N59</sub>-LPETG<sup>*</sup> complex. **a**, stereo image showing the ensemble of 40 lowest energy structures of the SrtA<sub>N59</sub>-LPETG<sup>*</sup> complex. The protein backbone heavy atoms (blue) and the covalently linked peptide (red) are shown. A yellow sphere represents the calcium ion. **b**, ribbon drawing of the structure of the SrtA<sub>N59</sub>-LPETG<sup>*</sup> complex. The covalently bound peptide is shown in a red ball-and-stick representation with its amino acids labeled. **c**, expanded view of the SrtA<sub>N59</sub>-substrate interface showing how the leucine and proline residues of the sorting signal are recognized. A ribbon drawing of the SrtA protein (white) with select side chains (yellow) is shown. **d**, expanded view of the positioning of the threonine and alanine residues of the sorting signal. The side chains of His<sup>120</sup> and Arg<sup>197</sup> are also shown, demonstrating their roles in catalysis. The panel is colored as in **c**. The -CH<sub>2</sub>-SH group of the peptide is covalently bonded with the side chain thiol group of Cys<sup>184</sup> through a disulfide bridge (pale yellow spheres).

(Submitted by Biopeptide Co., Inc.) Each peptide in the library contains the amino acid sequence from the sorting signal of protein A (LPETG) but randomizes the leucine position. NMR spectroscopy. To overcome this problem, we synthesized a peptide analog of the sorting signal that covalently modifies the enzyme. The peptide contains the amino acid sequence Cbz-LPETG<sup>*</sup>, where Cbz is a carbobenzyloxy protecting group and T<sup>*</sup> is a threonine derivative that replaces the carbonyl group with -CH<sub>2</sub>-SH (Fig. 1a) (40). Via its T<sup>*</sup> moiety, the peptide forms a disulfide bond with the active site Cys<sup>184</sup> thiol generating a covalent SrtA<sub>N59</sub>-LPETG<sup>*</sup> complex that structurally mimics the thioacyl intermediate of catalysis (compared in Fig. 1b).

The structure of the SrtA<sub>N59</sub>-LPETG<sup>*</sup> complex was determined using heteronuclear NMR methods (41). Previously, we assigned the backbone chemical shifts of the SrtA<sub>N59</sub> protein in the complex (29). In this report, to solve the structure of the complex, we assigned nearly all of the <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N chemical shifts of the protein and the <sup>1</sup>H chemical shifts of the bound peptide. As shown in Fig. 1c, the covalent SrtA<sub>N59</sub>-LPETG<sup>*</sup> complex exhibits good quality NMR spectra, enabling 36 intermolecular NOE distance restraints between the protein and peptide to be identified. The structure of the complex was calculated using 3,186 experimental restraints: 2,454 intraprotein

**RESULTS AND DISCUSSION**

Structure of the Covalent Complex between Sortase and an Analog of the Sorting Signal—The molecular basis of sorting signal recognition and transpeptidation is not well understood, because the reaction intermediates of catalysis are short lived and thus difficult to visualize by crystallography or analysis. A 40-μl reaction containing SrtA<sub>N59</sub> and the library was incubated at 37 °C for 16 h (final reaction concentrations: 0.1 mg/ml peptide library, 15 μM SrtA<sub>N59</sub> dissolved in 20 mM HEPES, 5 mM CaCl<sub>2</sub>, and 2 mM Gly<sub>3</sub>, pH 7.5). A 4-μl aliquot from the reaction was then quenched by the addition of 4 μl of 0.2% trifluoroacetic acid. After mixing with an equal amount of α-cyano-4-hydroxycinnamic acid, the products and reactants were analyzed by MALDI-TOF using a Voyager-DE STR Biospectrometry Work station (Applied Biosystems) under the positive ion mode. Both substrates and transpeptidation products can be simultaneously observed in the mass spectrum.
NMR Structure of the Covalent Sortase-sorting Signal Complex

The notation of the NMR structures is as follows: (SA) are the final 20 simulated annealing structures; (SA), is the average energy-minimized structure. The number of terms for each restraint is given in parentheses.

**TABLE 1**

Structural statistics for the NMR structure of the SrtA_{AN59}-LPAT^+ complex

| r.m.s. deviations from NOE interproton distance restraints (Å) |  |
|---|---|---|---|
| All (2456) | 0.027 ± 0.011 | 0.039 |
| Intermolecular (36) | 0.030 ± 0.017 | 0.054 |
| r.m.s. deviations from hydrogen-bonding restraints (Å) (94) | 0.023 ± 0.012 | 0.024 |
| r.m.s. deviations from dihedral angle restraints (degrees) (260) | 0.304 ± 0.141 | 0.712 |
| r.m.s. deviations from \( ^3 \)J_{HN} coupling constants (Hz) (88) | 0.579 ± 0.241 | 0.726 |

| r.m.s. deviations from secondary \( ^1 \)C shifts (ppm) |  |
|---|---|---|
| (127) | 0.91 ± 0.04 | 1.38 |
| (61) | 1.05 ± 0.15 | 1.54 |

<table>
<thead>
<tr>
<th>Residual dipolar coupling R-factors (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DNH (66)</td>
<td>4.6 ± 0.1</td>
</tr>
<tr>
<td>DNC (61)</td>
<td>6.4 ± 0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Deviations from idealized covalent geometry</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonds (Å)</td>
<td>0.002 ± 0.001</td>
</tr>
<tr>
<td>Angles (degrees)</td>
<td>0.055 ± 0.044</td>
</tr>
<tr>
<td>Impropers (degrees)</td>
<td>0.181 ± 0.174</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PROCHECK-NMR</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Most favored regions (%)</td>
<td>71 ± 1</td>
</tr>
<tr>
<td>Additionally allowed regions (%)</td>
<td>27 ± 1</td>
</tr>
<tr>
<td>Generously allowed regions (%)</td>
<td>2.5 ± 0.8</td>
</tr>
<tr>
<td>Disallowed regions (%)</td>
<td>0.0 ± 0.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coordinate precision (Å)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein and peptide backbone</td>
<td>0.28 ± 0.06</td>
</tr>
<tr>
<td>All protein and peptide heavy atoms</td>
<td>0.87 ± 0.08</td>
</tr>
</tbody>
</table>

---

The signal rests against the β6/β7 loop, where residues Val\(^{166}\)–Leu\(^{169}\) adopt a 3_10 helix that only forms when the substrate is bound (Fig. 2c). Helix formation enables the leucine methyl groups of the analog to be partially encircled by hydrophobic contacts. From above, the leucine side chain is in close proximity to the α-protons of Thr\(^{164}\) and Val\(^{166}\), whereas from below, it is contacted by the side chains of Val\(^{168}\) and Arg\(^{197}\). The proline ring of the sorting signal is buried in the binding groove by contacts from the side chains of Ile\(^{182}\) (β7) and Ala\(^{118}\) (β4) that project from the underlying sheet and by contacts from residues within both walls of the groove (Leu\(^{169}\) (3_10 helix), Ala\(^{92}\) (β2/H1 loop), and Ala\(^{103}\) (β3/β4 loop)). This latter interaction is supported by the observation of strong intermolecular NOEs between the methyl groups of the alanine residues and protons within the proline ring.

The LPAT^+ peptide adopts an “L-shaped” structure as a result of a ~90° kink at the alanine-proline peptide bond, which is in a trans conformation. The kink redirects the trajectory of the signal, enabling it to approach the active site in parallel with the underlying β-strands of the binding groove. In the LPAT^+ peptide, an alanine residue mimics the X position of the LPXTG motif. In the structure, the alanine is packed against the side chain of Leu\(^{97}\) located in helix H1 as a result of several strong NOEs to the Leu\(^{97}\) Hδ protons (Fig. 2d). This explains the demonstrated promiscuity of SrtA for this site within the sorting signal, since model building studies suggest that larger side chains could project away from the enzyme via a cleft located between helix H1 and His\(^{120}\) (39). Recognition is completed by packing of the threonine γ-methyl group beneath the indole ring of Trp\(^{194}\), as substantiated by several NOEs between the methyl and the Hε1 proton of the tryptophan. These contacts partially shield the active site from the solvent and help to project the -CH2-SH portion of the threonine analog toward Cys\(^{184}\) for disulfide bond formation.

**Sorting Signal Binding Closes and Immobilizes the β6/β7 Loop**—Based on biochemical studies, the active site β6/β7 loop plays an important role in catalysis, since amino acid mutations in this structural element significantly impair enzyme activity and alter the substrate specificity of the enzyme (32, 66). This is compatible with the structure of the complex, since residues from the NH2-terminal portion of the loop directly contact the leucine and proline residues of the sorting signal (Fig. 2, b and c). NMR and x-ray studies of apo-SrtA\(^{AN59}\) indicate that the β6/β7 loop is unstructured and mobile in the absence of the sorting signal (31, 34). Interestingly, a comparison of the NMR structures of SrtA\(^{AN59}\) solved in the presence and absence of the sorting signal reveals that substrate binding causes the loop to transition from a structurally disordered and open conformation to an ordered “closed” conformation containing a 3_10 helix (Fig. 3, a and b). The substrate-induced structural changes are extensive and involve a ~10-Å displacement of the loop toward the active site (Fig. 3c). The loop is also shortened, as residues Thr\(^{156}\) to Val\(^{161}\) become incorporated into strand β6 by forming a network of new hydrogen bonds to residues within strand β8. Notably, this large structural change was not observed in a previously reported crystal structure of the SrtA\(^{AN59}\)-LPETG complex (see below) (31).

Distances, 36 intermolecular distances, 94 hydrogen bonds, 88 \(^3\)H\(_{N\alpha}\) couplings, 254 carbon chemical shifts, and 260 dihedral angles. This produced an ensemble of 20 conformers that possess good covalent geometry with no NOE, dihedral angle, or scalar coupling violations greater than 0.5 Å, 5°, or 2 Hz, respectively (Fig. 2a). The amino acids of the peptide analog and residues Pro\(^{61}\)–Ile\(^{87}\) of SrtA\(^{AN59}\) are well defined by the NMR data and have backbone and heavy atom coordinate root mean square deviations to the mean structure of 0.28 ± 0.06 and 0.87 ± 0.08 Å, respectively. Complete structure and restraint statistics are presented in Table 1.

**Structural Basis of LPXTG Binding**—SrtA recognizes the LPXTG sorting signal through a large groove that leads into the active site (Fig. 2b). Residues in strands β4 and β7 form the floor of the groove, whereas the walls are formed by surface loops that connect strand β6 to strand β7 (β6/β7 loop), strand β7 to strand β8 (β7/β8 loop), strand β3 to strand β4 (β3/β4 loop), and strand β2 to helix H1 (β2/H1 loop). The leucine residue of distances, 36 intermolecular distances, 94 hydrogen bonds, 88 \(^3\)H\(_{N\alpha}\) couplings, 254 carbon chemical shifts, and 260 dihedral angles. This produced an ensemble of 20 conformers that possess good covalent geometry with no NOE, dihedral angle, or scalar coupling violations greater than 0.5 Å, 5°, or 2 Hz, respectively (Fig. 2a). The amino acids of the peptide analog and residues Pro\(^{61}\)–Ile\(^{87}\) of SrtA\(^{AN59}\) are well defined by the NMR data and have backbone and heavy atom coordinate root mean square deviations to the mean structure of 0.28 ± 0.06 and 0.87 ± 0.08 Å, respectively. Complete structure and restraint statistics are presented in Table 1.

**Structural Basis of LPXTG Binding**—SrtA recognizes the LPXTG sorting signal through a large groove that leads into the active site (Fig. 2b). Residues in strands β4 and β7 form the floor of the groove, whereas the walls are formed by surface loops that connect strand β6 to strand β7 (β6/β7 loop), strand β7 to strand β8 (β7/β8 loop), strand β3 to strand β4 (β3/β4 loop), and strand β2 to helix H1 (β2/H1 loop). The leucine residue of

---

Nonetheless, this large structural change was not observed in a previously reported crystal structure of the SrtA\(^{AN59}\)-LPETG complex (see below) (31).

Distances, 36 intermolecular distances, 94 hydrogen bonds, 88 \(^3\)H\(_{N\alpha}\) couplings, 254 carbon chemical shifts, and 260 dihedral angles. This produced an ensemble of 20 conformers that possess good covalent geometry with no NOE, dihedral angle, or scalar coupling violations greater than 0.5 Å, 5°, or 2 Hz, respectively (Fig. 2a). The amino acids of the peptide analog and residues Pro\(^{61}\)–Ile\(^{87}\) of SrtA\(^{AN59}\) are well defined by the NMR data and have backbone and heavy atom coordinate root mean square deviations to the mean structure of 0.28 ± 0.06 and 0.87 ± 0.08 Å, respectively. Complete structure and restraint statistics are presented in Table 1.

**Structural Basis of LPXTG Binding**—SrtA recognizes the LPXTG sorting signal through a large groove that leads into the active site (Fig. 2b). Residues in strands β4 and β7 form the floor of the groove, whereas the walls are formed by surface loops that connect strand β6 to strand β7 (β6/β7 loop), strand β7 to strand β8 (β7/β8 loop), strand β3 to strand β4 (β3/β4 loop), and strand β2 to helix H1 (β2/H1 loop). The leucine residue of
To determine whether loop closure over the sorting signal analog quenches its mobility, we measured the $R_1$, $R_2$, and $^{15}$N($^{1}$H) NOE relaxation parameters of the protein backbone nitrogen atoms in the complex and interpreted these data using the model-free formalism (60–63). This analysis yields the order parameter ($S^2$), which gives a concise account of the backbone amide’s mobility on the picosecond time scale. It ranges from 0 to 1, with values of 1 indicating that the amide is completely immobilized. The model-free analysis also yields an $R_{ex}$ term that is diagnostic for the presence of slower micro- to millisecond time scale motions. Fig. 4 compares the $S^2$ and $R_{ex}$ values of SrtA$_{AN59}$ in the SrtA$_{AN59}$-LPAT* complex with similar data reported for the apo-form of the enzyme (56). In the SrtA$_{AN59}$-LPAT* complex, the $\beta6/\beta7$ loop is rigid on fast timescales, as evidenced by $S^2$ values that are on average 0.90 ± 0.01 (Fig. 4a, black). It is also immobile on slower timescales, since only a few residues distributed throughout the protein exhibit small magnitude $R_{ex}$ terms (Fig. 4b, black). This is in marked contrast to the apoenzyme, since many of the residues in the $\beta6/\beta7$ exhibit elevated $R_{ex}$ values and/or weak NMR resonances that indicate that they undergo slow micro- to millisecond time scale motions (Fig. 4b, white and gradient bars). Interestingly, in the apoenzyme, several residues in the loop have $S^2$ values of >0.7, demonstrating that they do not participate in large amplitude picosecond time scale motions (Fig. 4a, gray). Combined, the structural and relaxation data suggest that the loop in the apoenzyme adopts a semirigid state that undergoes micro- to millisecond segmental motions that toggle it between open and closed conformations. Substrate binding quenches these motions, locking the loop in a closed state for productive interactions with the sorting signal.

Closure and immobilization of the $\beta6/\beta7$ loop enables extensive enzyme contacts to the leucine residue within the sorting signal. We tested whether these interactions confer specificity for this site by challenging SrtA$_{AN59}$ with a peptide library containing the LPETG sequence randomized at the leucine position (SKRQAXPETGEESTE, where X represents any amino acid except for Ile or Cys). Monitoring product formation using mass spectrometry reveals that only peptides bearing a leucine amino acid at this site are effectively processed by the enzyme (Fig. 5a). This supports the notion that loop closure over the signal plays a key role in substrate recognition.

$\text{Ca}^{2+}$ Stabilizes the Closed Conformation of the $\beta6/\beta7$ Loop—

\textit{In vitro,} $\text{Ca}^{2+}$ increases the enzymatic activity of SrtA$_{AN59}$ ~8-fold by lowering the $K_m$ of the enzyme for the sorting signal (34, 56). This adaptation may enable \textit{S. aureus} to increase the number of displayed proteins if elevated concentrations of $\text{Ca}^{2+}$ are encountered at sites of infection. The atomic basis of divalent ion binding is incompletely understood, because all previously reported crystal structures of SrtA$_{AN59}$ were solved in the metal-free state (31). Moreover, the binding mechanism of $\text{Ca}^{2+}$ cannot easily be defined by NMR methods, since $\text{Ca}^{2+}$ lacks proton atoms needed to identify NOE distance restraints. In structure calculations of the complex, we employed three artificial distance restraints between a single $\text{Ca}^{2+}$ ion and the side chains of residues Glu$_{105}$, Glu$_{108}$, and Glu$_{171}$. These were included because mutant proteins that replace these residues with alanine are insensitive to $\text{Ca}^{2+}$ and because the chemical shifts of these residues are significantly perturbed when $\text{Ca}^{2+}$ is added (56). Importantly, structures calculated with the artificial restraints are compatible with all of the experimental data. In the structure of the complex, $\text{Ca}^{2+}$ binds to a pocket formed by the $\beta3/\beta4$ (Glu$_{105}$ and Glu$_{108}$) and $\beta6/\beta7$ loops (Glu$_{171}$) (Fig. 3c). As compared with the crystal structure of the apoenzyme.
His120, Cys184, and Arg197 in the active site are mutated, consistent with their direct participation in the transpeptidation reaction. Outside of the enzyme active site, some of the most persistent with the structure in the SrtA closed state of the substrate contacting 6/7 loop is immobilized when the enzyme binds to the sorting signal, a conformation that is similar to the loop in the S. aureus enzyme. This conclusion is supported by relaxation data, which indicate that Ca2+ binding to the enzyme retards motions in the loop (56).

Compatibility of the SrtAΔN59-LPAT* Complex with Biochemical Data—The structure of the complex explains the enzymatic properties of 35 previously described amino acid mutants of the S. aureus SrtAΔN59 protein (26, 28–30). These data are not fully described here but are summarized in Fig. 5c. As expected, the most severe effects are observed when residues His120, Cys184, and Arg197 in the active site are mutated, consistent with their direct participation in the transpeptidation reaction. Outside of the enzyme active site, some of the most severe effects occur when the β6/β7 loop is mutated, with V168A and L169A mutations reducing enzymatic activity 5.5- and 93-fold, respectively (32). This is consistent with the structure, since in it, these side chains directly contact the sorting signal (Fig. 2c). Mutations within the Ca2+ binding site also disrupt activity, consistent with the ion acting to stabilize the closed state of the substrate contacting β6/β7 loop (Fig. 3c).

To further substantiate the mode of binding observed in the NMR structure, we created A104G, A118G and L97V mutants of SrtAΔN59 and tested their enzymatic activity using a fluorescence resonance energy transfer assay. Based on the structure, the A104G and A118G mutations should remove methyl groups from the floor and the sides of the binding groove that contact the proline ring (Fig. 2c). This is consistent with our results, since these mutants are 27–34-fold less active than the wild-type protein (Fig. 5b). The side chain of Leu97 is conserved in >97% of all sortase enzymes, and in the structure of the complex, its side chain packs against the backbone of the sorting signal (Fig. 2d). Our results indicate that these contacts play an important role in stabilizing the positioning of the sorting signal, since even a conservative L97V mutant of SrtAΔN59 is 10-fold less active than the wild-type protein (Fig. 5b). It is important to note that the structure of the complex reported in this paper is not a perfect mimic of the thio-acyl intermediate, since the peptide contains an extra methylene group that separates the Cys184 thiol from the threonine. However, the spacer only adds ~2 Å, which is unlikely to significantly alter the way in which the peptide is recognized by the enzyme. This is especially true for recognition of the leucine residue of the peptide, which is positioned distal to the active site.

Some Sortase Enzymes May Contain a Preformed Binding Pocket for the LPXTG Sorting Signal—In addition to the SrtA enzyme from S. aureus, the structures of three sortase enzymes that recognize the LPXTG sorting signal have now been determined (Fig. 6) (37, 38). The structures were solved in the absence of the sorting signal and are of enzymes that share only limited sequence homology with S. aureus SrtA. They include S. pyogenes SrtA (26% sequence identity) and the SrtC-1 (25% sequence identity) and SrtC-3 (27% sequence identity) enzymes from S. pneumoniae. Interestingly, these enzymes adopt the same β-barrel fold observed in the S. aureus enzyme; the backbone coordinates of the S. pyogenes SrtAΔN59 and the S. pneumoniae SrtC-1ΔN59 and SrtC-3ΔN59 enzymes can be superimposed with the coordinates of the S. aureus SrtAΔN59 enzyme with an r.m.s. deviation of 1.8, 1.6, and 1.7 Å, respectively. Additionally, the arginine, cysteine, and histidine active site are positioned similarly.

The S. pyogenes and S. pneumoniae enzymes may contain a preformed binding pocket for the sorting signal. In this paper, we have shown that binding of the sorting signal to the S. aureus SrtAΔN59 enzyme causes major changes in the structure and dynamics of the β6/β7 loop. It transitions from a disordered open state to an ordered closed conformation that contains a 310 helix (compared in Fig. 6, a and b). Intriguingly, although the structures of the S. pyogenes and S. pneumoniae enzymes were solved in the absence of the sorting signal, their β6/β7 loops also adopt a closed conformation that contains a 310 helix (Fig. 6, c and d). This conformation is similar to the loop structure in the SrtAΔN59-LPAT* complex, suggesting that in the S. pyogenes and S. pneumoniae enzymes, only modest changes in the structure of the β6/β7 loop are needed to bind the sorting signal.

In the crystal structures of the S. pneumoniae SrtC-1 and SrtC-3 enzymes, an additional polypeptide segment is inserted into the active site and has been proposed to act as a “lid” that opens and closes during pilus assembly (38). Remarkably, two of the residues in the lid are embedded in the sorting signal...
The structure of the SrtA-ΔN59-LPAT* complex. There are also substantial differences in the structure of the non-covalent C184ASrtA-LPAT* complex reported in this paper. In the crystal structure, the loop adopts an extended conformation, whereas in the covalent SrtA-ΔN59-LPAT* complex, a ~90° kink occurs between the alanine and proline residues (Fig. 7a). The peptide in the SrtA-ΔN59-LPAT* complex is also positioned closer to the side chain of Cys184 within the active site. These differences cause the leucine, proline and X positions within the sorting signal substrate to be contacted by different amino acids in each complex. Indeed, none of the contacts between SrtA and the sorting signal substrate are conserved.

The conformations of the β6/β7 and β7/β8 loops differ substantially. In the non-covalent C184A-SrtA-ΔN59-LPETG complex, the loop adopts an “open” conformation in which it extends away from the body of the protein. In contrast, the loop in the covalent SrtA-ΔN59-LPAT* complex is in a “closed” conformation that places it closer to the active site (Fig. 7b). The largest differences in loop positioning occur at residue Val166, whose α-carbon is shifted by ~10 Å toward the active site in the covalent complex. There are also substantial differences in the structure of the β6/β7 loop itself. In the NMR structure, it contains a 3_10 helix at its center and is shorter in length by approximately 6 residues. The length difference is caused by different conformations of residues Thr156-Val161. In the crystal structure, they adopt an unusual bulge structure, whereas in the covalent SrtA-ΔN59-LPAT* complex, they are incorporated into strand β6. The positioning of the β7/β8 loop in the two complexes is different, b, enzyme reaction progress curves for wild-type SrtA-ΔN59 (WT) and L97V, A104G, and A118G mutants. The mutations disrupt protein-sorting signal contacts observed in the solution structure of the SrtA-ΔN59-LPAT* complex. The data show that each mutant has reduced activity relative to the wild-type enzyme and are compatible with the solution structure of the complex (31). The reaction progress curves monitor the transpeptidation of a d-QALPETGEE-e sorting signal substrate using a fluorescence resonance energy transfer assay. Analysis of the kinetic data yielded the following steady-state parameters: wild-type, $k_{cat} = 5.7 \times 10^{-4}$; $K_m = 5.6 \times 10^{-4}$ and $V_{max} = 2.9 \times 10^{-4}$; and L97V, $k_{cat} = 6.7 \times 10^{-4}$ and $V_{max} = 1.8 \times 10^{-4}$; and A104G, $k_{cat} = 9.0 \times 10^{-4}$ and $V_{max} = 7.3 \mu M$. It is important to note that the reported kinetic parameters are for the isolated catalytic domain of sortase catalyzing the transpeptidation of a peptide fragment that mimics the intact protein substrate. Different kinetic parameters may be obtained if the intact protein substrate is used or if sortase is located in its native environment, the bacterial cell surface. c summarizes the compatibility of the structure of the SrtA-ΔN59-LPAT* complex with previously reported biochemical data. The solvent-accessible surface of SrtA-ΔN59 in the complex is color-coded to show the effects of amino acid mutations on enzyme activity. The 10-fold reduction of the catalytic activity of each mutant relative to the wild-type protein is shown. A visual spectrum color code is used and ranges from violet (no effect on catalysis) to red (most severe, >2000-fold reduction in activity).
LPETG complex is non-specifically bound to the enzyme. First, in the structure of the C184A-SrtAΔN59-LPETG complex, the leucine side chain in the sorting signal projects into the solvent and is not contacted by the enzyme (Fig. 7b). This contradicts biochemical data that indicate that the leucine residue of the signal is specifically recognized by the enzyme (Fig. 5a), and it does not explain why the leucine is completely conserved in the sorting signals of S. aureus proteins anchored by SrtA (22, 39). Second, the structure of the non-covalent C184A-SrtAΔN59-LPETG complex is incompatible with the enzymatic properties of several amino acid mutants of SrtA. For example, recently reported E171A and V168A mutations reduce activity 5.4- and 5.5-fold, respectively (32). However, in the C184A-SrtAΔN59-LPETG complex, the side chains of these residues project into the solvent and do not contact the peptide. Another discrepancy is the positioning of the side chain of Gln172. Based on the structure of the C184A-SrtAΔN59-LPETG complex, it should be important for catalysis, because it forms a hydrogen bond to the backbone of the sorting signal, but a Q172A mutant surprisingly retains wild-type activity (32, 56). The reduced enzymatic activities of the A104G, A118G, and L97V mutants of SrtA reported in this paper are also incompatible, since in the crystal structure, the side chains of these residues do not contact the signal (Fig. 5b). Finally, the structure of the C184A-SrtAΔN59-LPETG complex is incompatible with intermolecular NOE data (Fig. 1a). The most significant discrepancies occur at the protein interface that interacts with the proline and leucine residues of the sorting signal. For example, unambiguous NOEs are observed between the proline side chain of the sorting signal and the methyl groups of Ala118 and Val166, indicating that they are separated by <5 Å (Fig. 1c). However, in the crystal structure, these hydrogen atoms are no closer than 11 Å apart. In sum, the previously published crystal structure of the non-covalent C184A-SrtAΔN59-LPETG complex is incompatible with our NMR data and the enzymatic properties of several amino acid...
NMR Structure of the Covalent Sortase-sorting Signal Complex

![Diagram](http://www.jbc.org/content/suppl/2009/07/10/M109.022624.DC1.html)

**FIGURE 7.** The structures of the SrtA<sub>ΔN59</sub>-LPAT<sup>+</sup> and C<sup>184A</sup>SrtA<sub>ΔN59</sub>-LPETG complexes differ substantially. (a) structures of the covalent SrtA<sub>ΔN59</sub>-LPAT<sup>+</sup> complex (left) and the previously reported structure of the C<sup>184A</sup>SrtA<sub>ΔN59</sub>-LPETG complex (right) (31). The solvent-accessible surface of the protein and the structure of the bound peptide are shown. The surface has been colored to indicate the electrostatic properties from acidic (red) to basic (blue). (b) ribbon diagrams of the covalent SrtA<sub>ΔN59</sub>-LPAT<sup>+</sup> complex (left, colored blue) and the C<sup>184A</sup>SrtA<sub>ΔN59</sub>-LPETG complex (right, colored tan). The secondary structural elements, relevant substrate contacting loops and amino acids within each peptide are labeled. The Ca<sup>2+</sup> ion in the covalent SrtA<sub>ΔN59</sub>-LPAT<sup>+</sup> complex is represented by a yellow sphere and labeled. The images have been rotated by 60° relative to a.

mutants, and it does not explain why the leucine residue is highly conserved in LPXTG sorting signals.

It seems likely that the sorting signal in the C<sup>184A</sup>SrtA<sub>ΔN59</sub>-LPETG complex is nonspecifically bound to the enzyme, because the β6/β7 loop failed to undergo a disordered to ordered transition. The x-ray structure of the complex was solved by molecular replacement using crystals of apo-C<sup>184A</sup>SrtA<sub>ΔN59</sub> soaked with the LPETG peptide (31). Three molecules of SrtA<sub>ΔN59</sub> were present in the asymmetric unit of the crystal, and in each, the β6/β7 loop adopted fundamentally distinct conformations. This structural heterogeneity may be caused by protein-protein interactions within the crystal as the β6/β7 loop interacts with other proteins in the unit cell. In addition, the loop may be mobile in the crystal, since in each of the three SrtA<sub>ΔN59</sub> models, residues Lys<sup>162</sup> to Lys<sup>175</sup> have B-factors in excess of 40 Å<sup>2</sup>, or their electron density is missing. Moreover, NMR relaxation studies indicate that the loop is mobile in the absence of the sorting signal (56). Notably, when the peptide is added to the crystal, it binds to only one of the SrtA<sub>ΔN59</sub> enzymes in the asymmetric unit, and the β6/β7 loop in this protein does not change its structure. Thus, it appears that the β6/β7 loop in the C<sup>184A</sup>SrtA<sub>ΔN59</sub>-LPETG complex failed to undergo substrate-induced structural changes required to properly recognize the signal, presumably because SrtA only weakly binds the LPXTG signal and/or because of lattice packing interactions (25). This problem has been overcome in the NMR structure of the SrtA<sub>ΔN59</sub>-LPAT<sup>+</sup> complex, because the peptide is covalently attached to the enzyme.

**Function of Arg<sup>197</sup>**—All sortases contain a highly conserved arginine residue within their active site (28). This residue corresponds to Arg<sup>197</sup> in SrtA, which when mutated significantly reduces catalytic activity (28–30, 32). The function of Arg<sup>197</sup> in catalysis has hitherto remained unclear, since it has been proposed to either deprotonate Cys<sup>184</sup> (28) or lipid II (31) or, alternatively, to stabilize either the binding of the sorting signal (28, 32) or oxyanion intermediates (31, 32). In many of the structures in the ensemble, the side chain of Arg<sup>197</sup> is wedged between Pro<sup>163</sup> of the β6/β7 loop and the sorting signal (Fig. 2d).

Here it presumably stabilizes the binding of the substrate in the active site by donating hydrogen bonds from its guanidine group to the backbone carbonyl oxygens of the leucine and proline residues of the sorting signal. It is important to note that the NMR data does not directly reveal the presence of these hydrogen bonds. Instead, they are inferred from the positioning of Arg<sup>197</sup> relative to the substrate, which is defined by NOEs between the side chains of Arg<sup>197</sup> and Val<sup>161</sup> in the β6/β7 loop as well as NOEs between Arg<sup>197</sup> and the leucine residue within the substrate. Contacts to the substrate require the specific chemical structure of the guanidine group, since mutations that replace Arg<sup>197</sup> with alanine or lysine dramatically decrease activity, whereas replacement with an isosteric citrulline side chain has little impact on catalysis (32).
Location of the Binding Site for the Second Substrate of Catalysis, the Triglycine Portion of Lipid II—SrtA anchors proteins through a ping-pong mechanism in which the second substrate of catalysis, lipid II, nucleophilically attacks the thioacyl-linked complex between SrtA and the LP\textsubscript{X}TG sorting signal (24, 25). The structure of the SrtA\textsubscript{H9004}N59-LPAT* complex mimics this thioacyl intermediate and thereby sheds light onto how lipid II might be recognized by the enzyme (Fig. 1b). A comparison of the substrate-free and -bound forms of the enzyme reveals that attachment of the sorting signal to Cys184 causes a large ~13-Å displacement of the \(\beta7/\beta8\) loop (Figs. 3b and 8a). The displacement is required in order to accommodate the insertion of the threonine side chain of the sorting signal underneath the indole ring of Trp\textsubscript{194} (Fig. 2d). This acts as a wedge at the base of the loop that shifts it as a rigid unit away from helix H1, creating a new groove that leads into the active site. The floor of the exposed groove is formed by residues immediately following strand \(\beta4\), and its walls are formed by residues in helix H1 and the \(\beta7/\beta8\) loop. Interestingly, the center of the exposed groove contains the catalytically important His\textsuperscript{120} side chain, suggesting that it might serve as the entry point into the active site for lipid II.

To investigate whether the exposed groove serves as a binding site for lipid II, we performed chemical shift perturbation experiments. In vivo, SrtA joins the terminal amine group of the pentaglycine branch of lipid II to the threonine carbonyl carbon.
NMR Structure of the Covalent Sortase-sorting Signal Complex

of the LPXTG sorting signal (18–20). A Gly3 peptide mimics this portion of lipid II and can be effectively used by the enzyme as a substrate in vitro (23). To locate the surface on the enzyme that interacts with Gly3, we titrated a sample of the SrtAΔNS9-LPAT* complex with the peptide and used NMR to monitor the chemical shifts of the backbone amide atoms of the protein. A histogram plot of the chemical shift differences in the presence and absence of Gly3 reveals that the peptide selectively perturbs the NMR spectrum of the protein (Fig. 8b, black bars). Interestingly, when the most significantly perturbed residues are mapped onto the structure of the protein in the SrtAΔNS9-LPAT*, they define a continuous surface that encompasses the groove that is unmasked when the sorting signal binds (Fig. 8c).

When the apo-form of the enzyme is titrated with the Gly3 peptide, the backbone amide chemical shifts of residues within this surface are not perturbed (Fig. 8b, white bars). Combined, these data suggest that sorting signal-induced displacement of the β7/β8 loop unmasks the binding site for the Gly3 portion of lipid II and may thereby direct catalysis toward product formation.

Mechanism of Catalysis—The structure of the complex enables a more detailed mechanism of the SrtA-catalyzed transpeptidation reaction to be proposed (Fig. 8d).

The apoenzyme is dynamic, with residues in the β6/β7 loop undergoing motions that periodically displace them from the active site (56) (step i in Fig. 8d). Substrate binding nucleates the folding and closure of the loop over the sorting signal, enabling recognition (step ii). Based on NMR and enzyme kinetic studies of the apo-form of sortase, the pKa values of His120 and Cys184 are 6.3–7 and 9.4, respectively (25, 33). Therefore, at physiological pH values, the predominant form of the enzyme contains His120 and Cys184 in their uncharged states.

Interestingly, solvent isotope effect measurements have led to the suggestion that this form of the enzyme is inactive. Instead, the active form of sortase has been proposed to be sparsely populated (0.06% of the total protein) and to contain His120 and Cys184 in their ionized states (25). Therefore, the incoming threonine carboxyl carbon of the sorting signal is probably attacked by the thiolate of Cys184, resulting in the concerted displacement of the β7/β8 loop (step iii). This opens a large groove leading out of the active site that accommodates residues positioned on the COOH-terminal side of the threonine residue of the sorting signal in the surface protein precursor. The groove contains the imidazolium side chain of His120 at its center, which is poised to protonate the amide leaving group as the scissile bond is broken (67). NMR chemical shift experiments indicate that the groove also functions as the entry point for the Gly3 portion of lipid II, suggesting that the newly protonated His120 side chain activates the incoming terminal amine of lipid II for nucleophilic attack on the enzyme-linked thioacyl intermediate (step iv).

The dual function of His120 as a general acid and base is consistent with its measured pKa value of 6.3–7 and is functionally similar to histidine residues in a number of other enzymes, including serine proteases (33, 71). The reaction is then completed by formation of the second tetrahedral intermediate and subsequent breakage of the enzyme-substrate bond to liberate the protein-lipid II-linked product (step v). In all steps, Arg197 plays a key role in catalysis by stabilizing the positioning of the substrate through direct hydrogen bonding to its polypeptide backbone. In addition, the side chain of Arg197 is properly positioned in several conformers in the NMR ensemble to stabilize both tetrahedral intermediates of catalysis by interacting with the oxyanion.

SrtA is required for the virulence of multidrug-resistant methicillin resistant S. aureus, which in the United States kills an estimated 18,000 people annually (67). In addition, a number of other important human pathogens contain SrtA homologs that when genetically eliminated cause defects in virulence. The data presented here should therefore aid in the design of small molecule sortase inhibitors that are useful in treating a range of bacterial infections.

Acknowledgments—We thank Dr. Robert Peterson for assistance with the NMR experiments, Dr. Joseph A. Loo for mass spectrometry experiments, Drs. Mandar T. Naik and Rosemarie L. Pilpa for technical assistance with the NMR relaxation studies, and Drs. James U. Bowie and Scott A. Robson for useful discussions.

REFERENCES

25. Frankel, B. A., Kruger, R. G., Robinson, D. E., Kelleher, N. L., and
NMR Structure of the Covalent Sortase-sorting Signal Complex


Supplemental Material can be found at: http://www.jbc.org/content/suppl/2009/07/10/M109.022624.DC1.html

Downloaded from www.jbc.org at UCLA Louise Darling Biobibliography on September 25, 2009