Structure of the Bacillus anthracis Sortase A Enzyme Bound to Its Sorting Signal

A FLEXIBLE AMINO-TERMINAL APPENDAGE MODULATES SUBSTRATE ACCESS

Albert H. Chan, Sung Wook Yi, Austen L. Terwilliger, Anthony W. Maresso, Michael E. Jung, and Robert T. Clubb

From the Department of Chemistry and Biochemistry, UCLA-DOE Institute of Genomics and Proteomics, and the Molecular Biology Institute, University of California, Los Angeles, California 90095 and the Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, Texas 77030

The endospore forming bacterium Bacillus anthracis causes lethal anthrax disease in humans and animals. The ability of this pathogen to replicate within macrophages is dependent upon the display of bacterial surface proteins attached to the cell wall by the B. anthracis Sortase A (BaSrtA) enzyme. Previously, we discovered that the class A BaSrtA sortase contains a unique N-terminal appendage that wraps around the body of the protein to contact the active site of the enzyme. To gain insight into its function, we determined the NMR structure of BaSrtA bound to a LPXTG sorting signal analog. The structure, combined with dynamics, kinetics, and whole cell protein display data suggest that the N terminus modulates substrate access to the enzyme. We propose that it may increase the efficiency of protein display by reducing the unproductive hydrolytic cleavage of enzyme-protein covalent intermediates that form during the cell wall anchoring reaction. Notably, a key active site loop (B7/B8 loop) undergoes a disordered to ordered transition upon binding the sorting signal, potentially facilitating recognition of lipid II.

Background: The Bacillus anthracis sortase A (SrtA) enzyme attaches virulence factors to the cell surface.

Results: The structure of BaSrtA bound to a substrate analog reveals key amino acids involved in substrate recognition and catalysis.

Conclusion: BaSrtA modulates substrate access using a unique N-terminal appendage.

Significance: This research could facilitate the design of new anti-infective agents that work by disrupting surface protein display.

Bacillus anthracis is a Gram-positive endospore forming bacterium that causes lethal anthrax disease in humans and animals. The ability of this pathogen to replicate within macrophages is dependent upon the display of bacterial surface proteins attached to the cell wall by the B. anthracis Sortase A (SrtA) enzyme. Previously, we discovered that the class A BaSrtA sortase contains a unique N-terminal appendage that wraps around the body of the protein to contact the active site of the enzyme. To gain insight into its function, we determined the NMR structure of BaSrtA bound to a LPXTG sorting signal analog. The structure, combined with dynamics, kinetics, and whole cell protein display data suggest that the N terminus modulates substrate access to the enzyme. We propose that it may increase the efficiency of protein display by reducing the unproductive hydrolytic cleavage of enzyme-protein covalent intermediates that form during the cell wall anchoring reaction. Notably, a key active site loop (B7/B8 loop) undergoes a disordered to ordered transition upon binding the sorting signal, potentially facilitating recognition of lipid II.

The endospore forming bacterium Bacillus anthracis causes lethal anthrax disease in humans and animals. The ability of this pathogen to replicate within macrophages is dependent upon the display of bacterial surface proteins attached to the cell wall by the B. anthracis Sortase A (SrtA) enzyme. Previously, we discovered that the class A BaSrtA sortase contains a unique N-terminal appendage that wraps around the body of the protein to contact the active site of the enzyme. To gain insight into its function, we determined the NMR structure of BaSrtA bound to a LPXTG sorting signal analog. The structure, combined with dynamics, kinetics, and whole cell protein display data suggest that the N terminus modulates substrate access to the enzyme. We propose that it may increase the efficiency of protein display by reducing the unproductive hydrolytic cleavage of enzyme-protein covalent intermediates that form during the cell wall anchoring reaction. Notably, a key active site loop (B7/B8 loop) undergoes a disordered to ordered transition upon binding the sorting signal, potentially facilitating recognition of lipid II.

Bacillus anthracis is a Gram-positive endospore forming bacterium that causes lethal anthrax disease in humans and animals. The ability of this pathogen to replicate within macrophages is dependent upon the display of bacterial surface proteins attached to the cell wall by the B. anthracis Sortase A (SrtA) enzyme. Previously, we discovered that the class A BaSrtA sortase contains a unique N-terminal appendage that wraps around the body of the protein to contact the active site of the enzyme. To gain insight into its function, we determined the NMR structure of BaSrtA bound to a LPXTG sorting signal analog. The structure, combined with dynamics, kinetics, and whole cell protein display data suggest that the N terminus modulates substrate access to the enzyme. We propose that it may increase the efficiency of protein display by reducing the unproductive hydrolytic cleavage of enzyme-protein covalent intermediates that form during the cell wall anchoring reaction. Notably, a key active site loop (B7/B8 loop) undergoes a disordered to ordered transition upon binding the sorting signal, potentially facilitating recognition of lipid II.
NMR Structure of B. anthracis SrtA-Substrate Complex

by the extracellular membrane-associated \(^{38}\)SrtA (13). Catalysis occurs through a ping-pong mechanism that is initiated when the active site cysteine residue in \(^{38}\)SrtA nucleophilically attacks the backbone carbonyl carbon of the threonine residue within the LPXTG motif, breaking the threonine-glycine peptide bond to create a sortase-protein complex in which the components are linked via a thioacetyl bond (14, 15). The protein is then transferred by \(^{38}\)SrtA to the cell wall precursor lipid II, when the amino group in this molecule nucleophilically attacks the thioacetyl linkage to create a peptide bond-linked protein-lipid II product. Transglycosylation and transpeptidation reactions that synthesize the cell wall then incorporate this product into the peptidoglycan. Other members of the sortase superfamily are believed to operate through a similar enzymatic mechanism as \(^{38}\)SrtA, but in some microbes (primarily enterococal and corynebacterial species), sortases perform equivalent transpeptidation reactions that assemble pili on the bacterial cell surface (16).

Based on their primary sequences members of the sortase superfamily can be partitioned into at least six subfamilies, called class A to F enzymes (6, 17, 18). The genome of \(B.\) anthracis encodes class A, B, and D sortases, known as \(^{88}\)SrtA, \(^{88}\)SrtB, and \(^{88}\)SrtC, respectively (19). \(^{88}\)SrtA, the focus of this study, anchors seven proteins to the cell wall by joining the threonine of the LPXTG sorting signal at the C terminus of the protein to the amine group of \(meso\)-diaminopimelic acid (DAP) within lipid II (20). \(^{88}\)SrtB anchors only one protein to the cell wall, IsdC, which is important for heme binding (21), whereas \(^{88}\)SrtC attaches two proteins involved in sporulation (BasH and BasI) (22).

Atomic structures of class A to D enzymes have now been determined revealing a common eight-stranded \(\beta\)-barrel-fold that contains conserved arginine, cysteine, and histidine active site residues (4–11). However, it is not well understood how sortases recognize their sorting signal and nucleophile substrates, because nearly all sortase structures reported to date have been determined in the apo-state. Previously we discovered that the class A \(^{88}\)SrtA sortase contains a unique N-terminal appendage that wraps around the body of the protein to contact the active site of the enzyme (23). However, the function of this appendage in catalysis remained unknown. In this article, we report the NMR structure of \(^{88}\)SrtA bound to a LPXTG sorting signal analog that reveals for the first time the mechanism of substrate recognition by this enzyme. In the structure of the complex, the appendage partially encapsulates the bound sorting signal and we show that these interactions play a role in the \textit{in vivo} and \textit{in vitro} function of the enzyme. A model of the thioacetyl intermediate constructed from the coordinates of the complex suggests that the N-terminal appendage may obstruct nucleophile access to the active site. We propose that on the cell surface, the N-terminal appendage may increase the efficiency of protein display by reducing the unproductive hydrolytic cleavage of enzyme-protein covalent intermediates that form during the cell wall anchoring reaction.

**Experimental Procedures**

Sample Preparation and NMR Spectroscopy—Wild-type SrtA protein from \(B.\) anthracis containing amino acid residues Asp\(^{57}\)-Lys\(^{210}\) (\(^{88}\)SrtA) was produced as described previously (23). Uniformly \(^{15}\)N- and \(^{13}\)C-labeled \(^{88}\)SrtA protein was covalently attached to an analog of the LPATG sorting signal, Boc-LPAT\(^*\) (where T* is \((2R,3S)-3\)-amino-4-mercapto-2-butanol, and Boc is a tert-butyloxycarbonyl protecting group), by incubating 200 \(\mu\)M \(^{88}\)SrtA with 2 mM Boc-LPAT\(^*\) in modification buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl) for 48 h at room temperature. MALDI-TOF mass spectrometry was used to confirm the production of stable covalent complex. Two samples of the complex dissolved in NMR buffer (50 mM NaPO\(_4\), pH 6.0, 0.01% Na\(_2\)CO\(_3\)) were studied by NMR: 1) 2.6 mM \([^{15}\)N,\(^{13}\)C\]Boc-LPAT\(^*\) bound to the unlabeled peptide dissolved in NMR buffer containing 8% D\(_2\)O; and 2) 2.6 mM \([^{15}\)N,\(^{13}\)C\]\(^{88}\)SrtA bound to the unlabeled peptide dissolved in deuterated NMR buffer.

NMR spectra were acquired at 298 K on Bruker Avance 500-, 600-, and 800-MHz spectrometers equipped with triple resonance cryogenic probes. NMR data were processed using NMRPipe (24) and analyzed using PIPPI (25) and CARA (version 1.8.4) (26). \(^1\)H, \(^{13}\)C, and \(^{15}\)N protein chemical shift assignments were obtained by analyzing the following experiments: \(^{15}\)N-HSQC, \(^{13}\)C-HSQC, HNCA, CBCA(CO)NH, HNCO, HN(CC)CO, HBHA(CO)NH, HNHA, HNHB, CC(CO)NH, H(CC)CO(NH), HCHC-ToCSY, and HCHC-COSY (27, 28). Chemical shift assignments of the unlabeled Boc-LPAT* were obtained by analyzing a two-dimensional (F1,F2) \(^{13}\)C-filtered NOESY spectrum (29). Chemical shift data have been deposited in the Biological Magnetic Resonance Data Bank (accession code 11570).

Structure Determination—Distance restraints to define the structure of the protein were obtained from three-dimensional \(^{13}\)N- and \(^{13}\)C-edited NOESY spectra (mixing time at 120 ms), and intermolecular distance restraints between \(^{88}\)SrtA and Boc-LPAT* were identified in three-dimensional (F1) \(^{13}\)C,\(^{15}\)N-filtered (F2) \(^{13}\)C-edited NOESY-HSQC and (F1) \(^{13}\)C,\(^{15}\)N-filtered (F2) \(^{15}\)N-edited NOESY-HSQC (30), and two-dimensional (F1) \(^{13}\)C-filtered NOESY spectra (29). The majority of \(\varphi\) and \(\psi\) dihedral angle restraints were obtained using the program TALOS (31). Additional \(\varphi\) angle restraints were obtained by measuring \(J_{\text{HNH}}\) values from HNHA spectrum (32), and additional \(\psi\) angle restraints were obtained by analyzing three-dimensional \(^{15}\)N-edited NOESY spectrum (33). \(\chi\) angles for dihedral angle restraints and stereotypical chemical shift assignments of \(\beta\)-methylene protons were obtained by analyzing \(^{15}\)N-ToCSY-HSQC, HNHB, HN(CO)HB, and \(^{15}\)N-ROESY spectra (34, 35).

NOE assignments were obtained automatically using the ATNOS and CANDID algorithms in UNIJO'10 (36, 37), and structure calculations were performed using XPLOR-NIH (38). The three-dimensional \(^{13}\)N-edited NOESY-HSQC spectrum of the complex dissolved in \(H_2\)O, and three-dimensional \(^{13}\)C-edited NOESY-HSQC spectrum of the complex dissolved in D\(_2\)O were used as input for UNIJO'10. These data were supplemented with dihedral angle restraints and carbon chemical shifts for secondary structure predictions. Seven cycles of ATNOS/CANDID and XPLOR-NIH calculations yielded a converged ensemble of the protein in the complex. All of the NOE assignments made by CANDID were subsequently veri-
NMR Structure of B. anthracis SrtA-Substrate Complex

Backbone Dynamics of the \(^{15}\)N Relaxation Data—The \(^{15}\)N relaxation data were collected using a diluted (1.2 mM) \(^{15}\)N- and \(^{13}\)C-labeled sample of the complex dissolved in H\(_2\)O on a Bruker Avance 600-MHz NMR spectrometer equipped with a triple resonance cryogenic probe. Data were analyzed using SPARKY (46) and included: \(^{15}\)N longitudinal relaxation rates (\(R_L\)), transverse relaxation rates (\(R_T\)), and \(^{1}H\)-\(^{15}\)N heteronuclear NOEs. Complete quantifiable relaxation information (\(R_L\), \(R_T\), and NOE) could be measured for 119 of 151 backbone amides. The average \(R_L\), \(R_T\), and NOE values for the complex are 1.47 ± 0.16, 13.14 ± 3.97, and 0.71 ± 0.25 s\(^{-1}\), respectively. The relaxation data were analyzed using the suite of analysis programs kindly provided by Prof. Arthur G. Palmer III, and the strategy used to analyze the relaxation data has been described previously (47).

Briefly, Pdbinertia was used to calculate the principal moments of inertia and yielded relative moments of 1.00±0.91:0.68. R2R1\(_{\text{tm}}\) was used to calculate an approximate correlation time of 9.3 ± 0.4 ns for rigid residues using \(R_2/R_1\) ratios. To ensure only rigid residues were used in the correlation time analysis, only \(R_2/R_1\) ratios that met the following criteria were used in the analysis: 1) the residue had a \(^{1}H\)-\(^{15}\)N NOE value >0.65, and 2) they were within 1 S.D. of the average. The tensor parameters were then calculated using Quadric_Diffusion (48, 49), which indicated that the isotropic model is statistically preferred for the complex over the axially symmetric or anisotropic models of tumbling. The relaxation data were then interpreted using the Lipari-Szabo Model-free formalism (50, 51) using the program FAST-Modelfree (52) to iteratively run the program Modelfree 4.20 (53). Of the 119 amino acids that gave complete quantifiable relaxation information, data from 98 residues could be fit satisfactorily. Forty-five residues were fit to Model 1 (\(S^2\) only), 3 residues were fit to Model 2 (\(S^2\) and \(\tau_J\)), 29 residues were fit to Model 3 (\(S^2\) and \(R_{\phi}\)), 5 residues were fit to Model 4 (\(S^2\), \(\tau_J\), and \(R_{\phi}\)), and 16 residues were fit to Model 5 (\(S^2\), \(S^2_{\phi}\), and \(\tau_J\)). \(S^2\) parameters could be measured for backbone amides within all regions of the protein, with the exception of residues in the loop that connects helix H1 and strand B1 (residues 70–79), and the latter half of the \(\beta8\) loop (residues 192–195).

Modeling of the Thioacyl Intermediate—The procedure used to generate the energy minimized model of the thioacyl intermediate has been described previously (54). Briefly, the methylene group and sulfur of the modified threonine were removed in PyMOL (45), and a carbonyl group was placed between the Ca of the threonine and the sulfur of Cys\(^{187}\), with the oxygen on the carbonyl pointed toward or away from Arg\(^{196}\). The rough models of the thioacyl intermediate were solvated in a periodic water box with a solvent distance of 10 Å, and parametrized in tLEAP (55). Restraints were placed on the initial atom positions of heavy atoms within the protein and the LPAT peptide, except for the side chain atoms of Arg\(^{196}\). Restraints were not placed on the initial coordinates of Arg\(^{196}\) because the position of its side chain was not well defined in the ensemble due to peak broadening. Models were then energy minimized and equilibrated in NAMD (56) by slowly removing restraints over 1 ns using 2-fs steps. Details about NAMD set up can be found in Jacobitz et al. (54). In the ensemble of NMR structures of the [\(^{15}\)SrtA-LPAT*] complex, the side chain of Arg\(^{196}\) adopts multiple conformations and in some conformers its side chain points away from the bound peptide. Importantly, the starting conformation of the arginine side chain prior to energy minimization does not affect its positioning in the model of the thioacyl intermediate, because when test calculations were performed using different starting structures in which the side chain pointed either away or toward the substrate similar results were obtained. Starting models in which the orientation of the oxygen atom within the thioacyl bond pointed toward, or away from, the active site arginine residue (Arg\(^{196}\)) also yielded a similar final model of the intermediate after energy minimization.

Site-directed Mutagenesis and Enzyme Kinetics Measurements—Three single amino acid mutants (S59G, I61A, and C187A) of \(^{15}\)N-SrtA were produced for kinetic measurements. Mutations were made using the QuikChange\textsuperscript{\textregistered} method (Agilent, San Diego, CA) and confirmed by DNA sequencing. Mutants, wild-type \(^{15}\)N-SrtA (amino acid residues Asp57-Lys210), and wild-type \(^{15}\)N-SrtA\(_{\text{Asp}}\) (amino acid residues Asp\(^{57}\)-Lys\(^{210}\)) were purified as described previously (23). In vitro substrate hydrolysis assay was performed as previously described (57). Briefly, the cleavage of a self-quenched fluorescent peptide, Abz-LPETG-Dap(Dnp)-NH\(_2\) (Peptide 2.0 Inc., Chantilly, VA), was monitored by excitation at 335 nm and recording emission at 420 nm on an Infinite\textsuperscript{\textregistered} M1000 PRO spectrofluorometer (Tecan, San Jose, CA). Assay conditions consisted of 20 mM

fied by manually inspecting the NOESY data. In the structure refinement process, additional intra- and intermolecular distance restraints were identified manually, \(^{3}J_{\text{H,N}}\) couplings and hydrogen bond restraints were included in structure calculations. Intermolecular NOEs were assigned in an iterative manner. Initially, the positioning of the peptide in the active site was determined using only unambiguous NOEs (intermolecular NOEs that could be readily identified in the filtered and edited NMR spectra of the complex because they were not overlapped with intramolecular protein NOEs). The initial model of the complex was then was used to guide the assignment of more ambiguous intermolecular NOEs. In all cases, only NOEs present in both the filtered and edited NOESY spectra were employed as distance restraints. Hydrogen bonds were identified by inspecting the NOESY data for characteristic patterns along with deuterium exchange experiments, and restraints for hydrogen bonds were implemented with the HBDB algorithm (39). At the final refinement stage, \(^{1}D_{\text{H,N}}\) and \(^{1}D_{\text{NCO}}\) residual dipolar couplings were included, which were measured using protein samples partially aligned in PEG C12E5/hexanol (40), using two-dimensional \(^{15}\)N-coupled IPAP \(^{1}H\)-\(^{15}\)N HSQC and two-dimensional carbonyl-coupled \(^{1}H\)-\(^{15}\)N HSQC experiments, respectively (41, 42). Axial and rhombic components of the alignment tensor were fitted using Module (43). In the final set of calculations, a total of 200 structures were generated, of which 94 had no NOE, dihedral angle, or scalar coupling violations greater than 0.5 Å, 5°, or 2 Hz, respectively. Of these, 20 structures with the lowest overall energy were chosen to represent the structure of \(^{15}\)N-SrtA-LPAT* and have been deposited in the Protein Data Bank (PDB code 2RUI). The programs MOL-MOL (44) and PyMOL (45) were used to generate figures.

Mutations were made using the QuikChange\textsuperscript{\textregistered} method (Agilent, San Diego, CA) and confirmed by DNA sequencing. Mutants, wild-type \(^{15}\)N-SrtA (amino acid residues Asp57-Lys210), and wild-type \(^{15}\)N-SrtA\(_{\text{Asp}}\) (amino acid residues Asp\(^{57}\)-Lys\(^{210}\)) were purified as described previously (23). In vitro substrate hydrolysis assay was performed as previously described (57). Briefly, the cleavage of a self-quenched fluorescent peptide, Abz-LPETG-Dap(Dnp)-NH\(_2\) (Peptide 2.0 Inc., Chantilly, VA), was monitored by excitation at 335 nm and recording emission at 420 nm on an Infinite\textsuperscript{\textregistered} M1000 PRO spectrofluorometer (Tecan, San Jose, CA). Assay conditions consisted of 20 mM
HEPES, pH 7.5, and 10 μM enzyme. Abz-LPETG-Dap(Dnp)-NH₂ concentrations between 5 and 320 μM were added to the reaction for a total reaction volume of 50 μl. Fluorescence was recorded for 12 h at 10-min intervals. Initial velocities were calculated in terms of relative fluorescence units/s. Standard curves were used to account for inner filter effects and convert relative fluorescence units to product concentration (14). $K_{m}$ and $k_{cat}$ values were calculated by direct nonlinear fitting three independent sets of data to the Michaelis-Menten equation using SigmaPlot 6.0 (SPSS, Chicago, IL).

Cellular SrtA Activity Measured by BasCFLAG/MH6 Display—pAHG322 plasmids (20) with basCFLAG/MH6 and the WT, S59G, or I61A mutant srtA gene were transformed into the B. anthracis AHG188 strain (20) lacking srtA and transformants selected by plating on LB-chloramphenicol (10 μg/ml). BasCFLAG/MH6 was collected by growing 200 ml of each isolate to $A_{600}$ of 1 in LB (AHG188) or LB-chloramphenicol (10 μg/ml, complemented strains). Cells were harvested by centrifugation and resuspended in 5 ml of TSM-lysozyme buffer (50 mM Tris-HCl, pH 7.5, 0.5 M glucose, 10 mM MgCl₂, 10 mg/ml of lysozyme), shaken at 37 °C for 4 h, then sedimented by centrifugation at 15,000 × g for 30 min. Purification of BasCFLAG/MH6 was achieved by applying cleared lysates to 1 ml of Ni-Sepharose (GE Healthcare, Pittsburgh, PA) and eluting with imidazole as described in Gaspar et al. (20). Eluted proteins were precipitated with acetone, subjected to SDS-PAGE, and transferred to a polyvinylidene difluoride membrane. BasCFLAG/MH6 was detected by chemiluminescence using an anti-FLAG polyclonal rabbit antibody (F7425, Sigma) and an anti-rabbit HRP secondary antibody (A0545, Sigma). Immunoblot images were quantified by measuring integrated densities of the BasC bands in ImageJ. Normalized values for each biological replicate were used for statistical analysis, and significance determined by a two-tailed Student’s t test.

Results

Structure of the B. anthracis SrtA Enzyme Bound to a Substrate Analog—To investigate how the B. anthracis SrtA sortase recognizes its cognate LPXTG sorting signal, we synthesized a peptide analog of this substrate that covalently modifies the enzyme. This peptide, Boc-LPAT*, contains an N-terminal protecting group t-butyloxy carbonyl (Boc) and a modified threonine (T*) that replaces the carbonyl group with -CH₂-SH (Fig. 1A). The thiol group on T* promotes the formation of a disulfide bond between the peptide and the active site Cys187 of the enzyme to generate a covalent complex that structurally mimics the thioacyl intermediate of catalysis (58–60). In this work, we studied a B. anthracis SrtA protein construct containing residues Asp⁵⁷-Lys²¹⁰ (₄₀SrtA). ₄₀SrtA contains the catalytic core of the enzyme as well as its N-terminal appendage. The solution structure of the ₄₀SrtA-LPAT* complex was determined using multidimensional heteronuclear NMR and simulated annealing methods. A total of 2,689 experimental restraints define the structure, including: 1,802 intra-protein NOE distance, 59 intermolecular NOE protein-peptide distance, 318 dihedral angle, 77 ¹H NOE, coupling, 282 ¹³C secondary shift, and 151 residual dipolar coupling restraints. Because the peptide is unlabeled, to generate unambiguous intermolecular NOE protein-peptide distance restraints, we first identified intermolecular NOEs from three-dimensional (F1) ¹³C,¹⁵N-filtered (F2) ¹³C-edited NOESY-HSQC spectrum of the ₄₀SrtA-LPAT* complex dissolved in deuterated buffer. The identity of the proton from ₄₀SrtA that gives rise to the set of NOEs and its chemical shift are shown at the top and bottom of each panel, respectively. On the right side of each cross-peak the sorting signal peptide proton that is proximal to the protein is indicated.

![Structure of the B. anthracis SrtA Enzyme Bound to a Substrate Analog](image)

FIGURE 1. The Boc-LPAT* modifier and NMR data of its complex with ₄₀SrtA. A, the chemical structure of the Boc-LPAT* peptide analog, where T* is (2R,3S)-3-amino-4-mercapto-2-butanol, and Boc is a tert-butyloxy carbonyl protecting group. B, selected panels showing intermolecular NOEs between ₄₀SrtA and the Boc-LPAT* peptide. Panels are taken from a three-dimensional (F1) ¹³C,¹⁵N-filtered, (F2) ¹³C-edited NOESY-HSQC spectrum of the ₄₀SrtA-LPAT* complex dissolved in deuterated buffer. The identity of the proton from ₄₀SrtA that gives rise to the set of NOEs and its chemical shift are shown at the top and bottom of each panel, respectively. On the right side of each cross-peak the sorting signal peptide proton that is proximal to the protein is indicated.
FIGURE 2. NMR solution structure of the BaSrtA-LPAT* complex. A, stereo image showing the ensemble of 20 lowest energy structures of the BaSrtA-LPAT complex. The protein backbone heavy atoms (blue) and the covalently linked peptide (red) are shown. The N-terminal appendage is colored green. The coordinates were superimposed by aligning the backbone N, Cα, and C atoms of Asp⁵⁻Lys²¹⁰ and Leu, Pro, Ala, and Thr* of the Boc-LPAT* peptide. The backbone and heavy atom coordinates of these residues have a r.m.s. deviation to the average structure of 0.42 and 0.89 Å, respectively. B, ribbon drawing of the energy minimized average structure of the BaSrtA-LPAT* complex. The covalently bound peptide is shown in a red ball-and-stick representation with its amino acids labeled. The N-terminal appendage is colored cyan. C, expanded view of the active site showing how the sorting signal peptide is recognized. The side chains (orange) that participate in the interaction are labeled and shown as sticks. D, expanded view of the active site showing how the N-terminal appendage (cyan) positions over the sorting signal peptide (red). The side chains (orange) that participate in the interaction are labeled and shown as sticks. E, overlay of the ¹H-¹⁵N HSQC spectra of apo BaSrtA (red) and the BaSrtA-LPAT* complex (blue). F, histogram plot of the compound chemical shift changes (Δδ) for the backbone amide hydrogen and nitrogen atoms of apo BaSrtA after the addition of Boc-LPAT*. Chemical shift changes are calculated by the equation Δδ = (ΔδH² + ΔδN/6.49)², where ΔδH and ΔδN are, respectively, the amide proton and nitrogen chemical shift difference for a given residue in the presence and absence of Boc-LPAT*. The dashed line represents 1 S.D. above the average Δδ of all amino acids. Amino acids experiencing significant chemical shift changes are labeled, and a schematic of secondary structures of the enzyme is shown above the plot.
The N-terminal Appendage Partially Encapsulates the Bound Sorting Signal—The catalytic domain of SrtA is preceded by a structured N-terminal polypeptide segment that is formed by residues Val73 and Val74 in the 310-helix H4, as well as by the a-protons of nearby residues in the loop (Pro169, Asp169, Lys170, and Trp171) (Fig. 2C). The bottom portion of the S4 pocket is formed by residues Val166, Ile185, and Val198, which together with the 8/3 loop encapsulate the leucine so as to bury ~90% of the accessible surface area of the side chain. The proline residue forms a ~90° kink, redirecting the trajectory of the peptide toward Cys187 within the active site. It is recognized by the S3 site, which is formed by side chains projecting from the floor of the groove (Ala124 and Ile185 in strands B4 and B7, respectively), and by residues within the surrounding walls (Val173, Val110, and Ser108 in B6/B7 loop, B3/B4 loop, and B2/H2 loop, respectively). The proline also makes contacts with the side chain of Ser59 within the N-terminal appendage, which is poised to interact with its carboxyl group. The alanine residue in the peptide is located at the X position of the LPXTG motif. Notably, its side chain methyl does not appear to be recognized as it is positioned on the surface of the enzyme in between Leu103 in helix H2 and Ile185 in the N-terminal appendage. This surface appears to be large enough to accommodate a range of amino acid side chains, consistent with the documented promiscuity of SrtA for this site within the sorting signal (62). The side chain of the threonine residue in the substrate is buried in the S1 pocket formed by the side chains of Ile185 in strand B7, Ala124 in strand B4, and Leu103 in helix H2 (Fig. 2, C and D). The threonine y-methyl group is also packed underneath the imidazole ring of the active site His126 residue. This interaction is substantiated by the intermolecular NOE cross-peaks between the y-methyl group and the e1- and a-protons of His126 (Fig. 1B). For a complete list of intermolecular NOEs identified and used in structure calculations, see supplemental Table S1. Extensive chemical shift assignments for apo-SrtA have been reported previously (23) enabling a comparison to the chemical shifts of the apo-SrtA-LPAT* complex. The chemical shifts of both forms of the protein are presented in supplemental Table S2. Compatible with the structure, the most significant chemical shift changes in the protein that result from peptide binding localize to residues within the binding groove and N-terminal appendage that contact the substrate (Fig. 2, E and F).

### TABLE 1

Statistics for the NMR structure of SrtA-LPAT* complex

| R.m.s. deviations from NOE interproton distance restraints (Å) | (S4) | (SA) |
|---------------------------------------------------------------|------|------|
| All (17161)                                                   | 0.031 ± 0.002 | 0.025 |
| Intermolecular (59)                                           | 0.054 ± 0.006 | 0.045 |
| R.m.s. deviations from dihedral angle restraints (degrees) (318) | 0.501 ± 0.055 | 0.970 |
| R.m.s. deviations from γH2O coupling constant (Hz) (77)       | 0.701 ± 0.021 | 0.648 |
| R.m.s. deviations from secondary 1C shifts (ppm) (141)        | 1.166 ± 0.021 | 1.188 |
| 1Cα (ppm)                                                    | 1.187 ± 0.021 | 1.196 |
| Residual dipolar coupling R-factors (%)                      | 4.3 ± 0.2 | 6.4 |
| Dorder (67)                                                  | 23.1 ± 0.8 | 27.9 |
| Deviations from idealized covalent geometry Bonds (Å)         | 0.0084 ± 0.0001 | 0.0082 |
| Angles (degrees)                                             | 0.748 ± 0.013 | 0.769 |
| Improper (degrees)                                           | 0.814 ± 0.013 | 0.872 |
| PROCHECK-NMR data                                            | Most favorable region (%) | 88.5 ± 1.2 | 87.8 |
| Additionally allowed region (%)                              | 10.6 ± 1.3 | 11.5 |
| Generously allowed region (%)                                | 1.0 ± 0.5 | 0.7 |
| Disallowed region (%)                                        | 0.0 ± 0.0 | 0.0 |
| Coordinate precision%                                        | 0.4 ± 0.06 | |
| Protein backbone (Å)                                        | 0.89 ± 0.06 | |
| Protein heavy atoms (Å)                                      | 0.89 ± 0.06 | |
| a (SA) represents an ensemble of the 20 best structures calculated by simulated annealing. (SA) represents the average energy-minimized structure. The number of terms for each restraint is given in parentheses. None of the structures exhibited distance violations greater than 0.5 Å, dihedral angle violations greater than 5°, or coupling constant violations greater than 2 Hz.
| b Experimental dihedral angle restraints comprised 150 φ, 144 ψ, and 24 χ1 angles.
| c The dipolar coupling R-factor ranges between 0 and 100% and is defined as the ratio of the r.m.s. deviation between observed and calculated values to the expected r.m.s. deviation if the vectors were randomly distributed, given by (2D)/(4πφ3/2), where Dorder is the magnitude of the principle component of the alignment tensor, and φ is the rhombicity, calculated to be 10.6 Hz and 0.19, respectively.
| d PROCHECK-NMR data includes residues Asp57 to Lys110 of SrtA, and the substrate analog Boc-LPAT*.
| e The coordinate precision is defined as the average atomic r.m.s. deviation of the 20 individual SA structures and their mean coordinates. The reported values are for residues Asp57 to Lys110 of SrtA, and the substrate analog Boc-LPAT*.
| f The coordinate precision is defined as the average atomic r.m.s. deviation of the 20 individual SA structures and their mean coordinates. The reported values are for residues Asp57 to Lys110 of SrtA, and the substrate analog Boc-LPAT*.

respectively. Complete structure and restraint statistics are presented in Table 1.

In the structure of the SrtA-LPAT* complex, the catalytic core of the enzyme recognizes the LPAT portion of the sorting signal via a large groove that leads into the active site, and the N-terminal appendage that precedes the catalytic core partially shields the peptide from the solvent (Fig. 2B). On the catalytic domain, the floor of the LPAT binding groove is formed by residues in strands B4 and B7, and its walls are formed by surface loops that connect strand B6 to B7 (B6/B7 loop), strand B3 to strand B4 (B3/B4 loop), and strand B2 to helix H2 (B2/H2 loop). Within the groove the peptide adopts an L-shaped conformation, and it is positioned so as to join the C-terminal end of the signal via a disulfide bond to the active site cysteine residue. The leucine, proline, and threonine residues within the substrate are recognized by the enzyme. Following convention, residues N-terminal to the Thr-Gly scissile peptide bond in the bound peptide are referred to as P1 to P4, and the corresponding subsites that they interact with on the enzyme are referred to as S1 to S4, respectively (61). The S4 binding site is formed by a pocket created by residues at the N-terminal end of the B6/B7 loop, and strands B6 and B8. The leucyl side chain within the pocket is contacted by the non-polar side chains of Val173 and Val174 in the 310-helix H4, as well as by the a-protons of nearby residues in the loop (Pro169, Asp169, Lys170, and Trp171) (Fig. 2C). The bottom portion of the S4 pocket is formed by residues Val166, Ile185, and Val198, which together with the 8/3 loop encapsulate the leucine so as to bury ~90% of the accessible surface area of the side chain. The proline residue forms a ~90° kink, redirecting the trajectory of the peptide toward Cys187 within the active site. It is recognized by the S3 site, which is formed by side chains projecting from the floor of the groove (Ala124 and Ile185 in strands B4 and B7, respectively), and by residues within the surrounding walls (Val173, Val110, and Ser108 in B6/B7 loop, B3/B4 loop, and B2/H2 loop, respectively). The proline also makes contacts with the side chain of Ser59 within the N-terminal appendage, which is poised to interact with its carboxyl group. The alanine residue in the peptide is located at the X position of the LPXTG motif. Notably, its side chain methyl does not appear to be recognized as it is positioned on the surface of the enzyme in between Leu103 in helix H2 and Ile185 in the N-terminal appendage. This surface appears to be large enough to accommodate a range of amino acid side chains, consistent with the documented promiscuity of SrtA for this site within the sorting signal (62). The side chain of the threonine residue in the substrate is buried in the S1 pocket formed by the side chains of Ile185 in strand B7, Ala124 in strand B4, and Leu103 in helix H2 (Fig. 2, C and D). The threonine y-methyl group is also packed underneath the imidazole ring of the active site His126 residue. This interaction is substantiated by the intermolecular NOE cross-peaks between the y-methyl group and the e1- and a-protons of His126 (Fig. 1B). For a complete list of intermolecular NOEs identified and used in structure calculations, see supplemental Table S1. Extensive chemical shift assignments for apo-SrtA have been reported previously (23) enabling a comparison to the chemical shifts of the apo-SrtA-LPAT* complex. The chemical shifts of both forms of the protein are presented in supplemental Table S2. Compatible with the structure, the most significant chemical shift changes in the protein that result from peptide binding localize to residues within the binding groove and N-terminal appendage that contact the substrate (Fig. 2, E and F).
the peptide. These include NOE cross-peaks between: the α-proton of Ser59 and β-methyl group of alanine in the peptide, the β-protons of Ser59 and γ-protons of the proline residue in the peptide, and the α-proton of Ile61 and β-methyl group of alanine in the peptide (Fig. 1B).

**15N Relaxation Measurements: the N-terminal Appendage Transiently Detaches from the Enzyme**—To investigate backbone motions within the BaSrtA-LPAT* complex, we measured the 1H-15N NOEs, spin-lattice ($R_1$), and spin-spin ($R_2$) relaxation rates of the protein backbone nitrogen atoms, and interpreted the data using the Model-free formalism (63). Of the 151 non-proline amino acids in the protein, complete quantifiable relaxation information (NOE, $R_1$, and $R_2$) can be obtained for 119 residues, of which data from 98 of these residues could be interpreted using the Model-free approach. This analysis yields three parameters that describe the motion of the protein backbone: the general order parameter $S^2$, the effective correlation time for internal motion $\tau_e$, and $R_{ex}$. $S^2$ reports the backbone mobility of the amide on the picosecond time scale and is characterized by the internal correlation time $\tau_e$. The value of $S^2$ ranges from 0 to 1, with a value of 1 indicating that the amide bond is completely immobilized. $R_{ex}$ is the chemical exchange contribution to $R_2$ and is diagnostic for the presence of slower micro- to millisecond time scale motions. The $R_1$, $R_2$, 1H-15N NOE, and $S^2$ data are shown in Fig. 3, A–D. The complete Model-free analysis results are presented as supplemental Table S3). Notably, similar $R_2$ values are obtained when data are acquired using a more dilute protein sample (180 μM BaSrtA-LPAT* complex), indicating that the complex is not aggregating.

**FIGURE 3. Mobility of BaSrtA-LPAT* as defined by NMR relaxation data.** Panels A–D are scatter plots of the relaxation data for the complex: $R_1$ (A), $R_2$ (B), 1H-15N heteronuclear NOE (C), and general order parameter $S^2$ of the backbone 15N atoms (D) as a function of residue number. The value of $S^2$ ranges from 0 to 1, with a value of 1 indicating that the amide bond is completely immobilized. The secondary structures present in the protein are displayed on top of each plot. E, ribbon drawing of BaSrtA-LPAT* showing the location of residues in which the backbone amide resonances could not be assigned (black spheres), or residues that have $R_{ex}$ values greater than 10 Hz (gray spheres).
NMR Structure of B. anthracis SrtA-Substrate Complex

The relaxation data indicate that the N-terminal appendage transiently detaches from the enzyme to adopt a disordered state. This is evident from $^{1}H-^{15}N$ NOE data that shows a gradient of increasing mobility toward the N terminus, presumably because the polypeptide progressively unzips ("unzips") from the enzyme starting at its N terminus (Fig. 3C). Because the Model-free analysis only indirectly reveals the presence of $R_{ex}$ contributions to relaxation, we also performed constant-relaxation-time Carr-Purcell-Meiboom-Gill (CPMG) experiments (64) that probe motions occurring on time scales ranging from 0.1 to 10 ms (63, 65, 66). However, analysis of $^{15}N$-CPMG relaxation dispersion curves did not reveal the presence of $R_{ex}$ in the N-terminal appendage (data not shown). This indicates that the motions that detach the amino-terminal residues in the appendage from the enzyme occur on a time scale that is faster than 100 µs (66).

In contrast to the N-terminal appendage that partially encapsulates the signal, the peptide binding surface on the catalytic domain adopts a rigid conformation. In particular, both the aforementioned $\beta_7/\beta_8$ and $\beta_6/\beta_7$ active site loops are rigid on the picosecond time scale according to their $S^2$ data ($S^2$ values for residues in these regions are >0.87). Overall, few residues exhibit significant $R_{ex}$ values in the signal binding pocket. A notable exception are backbone amide atoms of Asp$^{169}$ and Glu$^{172}$ within the $\beta_6/\beta_7$ loop that are located next to the N-terminal Boc protecting group of the bound peptide that is likely in motion (Fig. 3E). In addition, several residues near the active site exhibit resonance line broadening that prevented their chemical shift assignment: Ser$^{97}$ within the $\beta_2/H_2$ loop, His$^{126}$ and Asn$^{127}$ in the $\beta_4/\beta_5$ loop, Cys$^{187}$ in the end of strand $\beta_7$, and residues Val$^{188}$, Asn$^{193}$, and Lys$^{195}$ in the $\beta_7/\beta_8$ loop. These residues may be broadened as a result of localized protein motions in the active site, binding of the N-terminal appendage and/or changes in the tautomeration state of the imidazole side chain of His$^{126}$.

The N-terminal Appendage Slows Peptide Hydrolysis and Contributes to Protein Display in Intact Cells—To understand the role of the N-terminal appendage in catalysis, a FRET-based assay was used to monitor the ability of wild-type and mutant $^{15}N$SrtA proteins to hydrolyze a self-quenched fluorescent peptide substrate (Abz-LPETG-Dap(Dnp)-NH$_2$). Notably, we were unable to reproduce the transpeptidation reaction in vitro using diaminopimelic acid as a nucleophile, which is the moiety on lipid II to which surface proteins are attached. This suggests that sortase recognizes additional features on lipid II.

The kinetics of hydrolysis were determined for wild-type $^{15}N$SrtA, and three mutants: BaSrtA$^{S59G}$ and BaSrtA$^{I61A}$ single amino acid mutants that based on the structure should alter interactions between the appendage and the substrate (Fig. 4A). Wild-type $^{15}N$SrtA cleaves the peptide with $k_{cat}$ and $K_m$ values of $3.6 \times 10^{-5} \pm 2 \times 10^{-6}$ s$^{-1}$ and $306 \pm 23$ µM, respectively (Table 2). In contrast, removal of residues Asp$^{57}$-Pro$^{64}$ ($^{15}N$SrtA$^{64}$), Asp$^{57}$-Pro$^{64}$ and Ile$^{61}$-Asp$^{64}$ ($^{15}N$SrtA$^{64}$), and removal of residues Ser$^{97}$ within the $\beta_2/H_2$ loop, His$^{126}$ and Asn$^{127}$ in the $\beta_4/\beta_5$ loop, Cys$^{187}$ in the end of strand $\beta_7$, and residues Val$^{188}$, Asn$^{193}$, and Lys$^{195}$ in the $\beta_7/\beta_8$ loop. These residues may be broadened as a result of localized protein motions in the active site, binding of the N-terminal appendage and/or changes in the tautomeration state of the imidazole side chain of His$^{126}$.

Proteins contain mutations that should disrupt appendage-substrate interactions, but they have $K_m$ values similar to the wild-type protein. As described under "Discussion," the finding that the $^{15}N$SrtA$^{I61A}$ and $^{15}N$SrtA$^{64}$ mutants have increased $k_{cat}$ values relative to the wild-type protein is compatible with the amino-terminal appendage acting to slow hydrolysis of the thioacyl intermediate.

To investigate the biological importance of the N-terminal appendage we determined how mutating this segment in $^{15}N$SrtA affected its ability to display proteins in intact B. anthracis cells. Protein display was monitored in a ΔsrtA strain of B. anthracis. Sterne 34F2 strains lacking srtA (Δ) or complemented with the indicated srtA and basCFLAG/MH6 on pAHG322 were grown to an A$_{600}$ of 1, and the harvested cells digested with lysozyme to liberate BasCFLAG/MH6. Cleared lysates were purified via affinity chromatography and the resulting eluates were analyzed for the presence of BasCFLAG/MH6 by immunoblotting with anti-FLAG polyclonal antibodies. A representative blot is shown at the top. Biological replicates were quantified and normalized to the BasCFLAG/MH6 anchored by WT SrtA. Means are plotted with standard error at the bottom panel. * $p < 0.01$, n = 3.
B. anthracis (strain AHG188) that was transformed with plasmid (pAHG322) that expresses wild-type or mutated \(^{18}\)SrtA, as well as the BasC_{FLAG/MH6} reporter protein. BasC_{FLAG/MH6} has previously been shown to be anchored to the cell wall by sortase harboring an active cell wall sorting signal, as well as His6 and FLAG tags and contains the full-length BasC substrate of sortase harboring previously been shown to be anchored to the cell wall by sortase.

In this work, we report the structure of the SrtA enzyme from B. anthracis (\(^{18}\)SrtA) that anchors surface proteins required for bacterial escape from macrophages, a key step in the progression of lethal anthrax disease (3). In contrast to previously studied sortases, \(^{18}\)SrtA contains a unique N-terminal polypeptide appendage that directly interacts with its active site histidine residue (His\(^{126}\)). However, the function of the appendage and its role in signal recognition remained unclear. Here we report NMR studies that show that the appendage in \(^{18}\)SrtA directly contacts the bound sorting signal, potentially increasing the efficiency of sortase-mediated cell wall protein anchoring by retarding the spurious hydrolysis of a key thioacyl reaction intermediate.

\(^{18}\)SrtA binds its sorting signal via a preformed binding groove located in the catalytic domain, which surprisingly triggers a disordered to ordered transition in the adjacent \(\beta7/\beta8\) active site loop. The vast majority of enzyme interactions with the signal are mediated by residues located in a structurally ordered binding groove that is formed by residues in \(\beta6/\beta7\) loop (Val\(^{196}\)-Val\(^{197}\)), strand \(\beta4\) (Ala\(^{124}\)), strand \(\beta7\) (Ile\(^{183}\), \(\beta3/\beta4\) loop (Val\(^{110}\), \(\beta2/\beta2\) loop (Ser\(^{89}\)), and helix H2 (Leu\(^{103}\)) (Fig. 2A). In the apo-form of \(^{18}\)SrtA, residues in the groove adopt a fixed structure and undergo only small positional changes upon ligand binding (the backbone atoms of the residues in the complex that form the pocket are aligned with an r.m.s. deviation of 0.76 Å to the structure of the apo-form of \(^{18}\)SrtA). In contrast to this preordered binding groove, the \(\beta7/\beta8\) loop in the active site experiences large changes in its dynamics and structure when the signal binds (Fig. 5A). In the apo-form of \(^{18}\)SrtA, the \(\beta7/\beta8\) loop is structurally disordered and its residues undergo motions on the micro to millisecond time scales; the backbone amide resonances of nearly all of its residues exhibited substantial motion-induced line broadening (only the backbone amide of Lys\(^{196}\) could be assigned). In contrast, in the \(^{18}\)SrtA-LPAT* complex, the coordinates of the \(\beta7/\beta8\) loop becomes well defined and more immobilized. The substrate dependent disorder-to-order transition is evidenced by the NMR relaxation data, our ability to assign the majority of backbone amide resonances of the loop in the NMR spectra of the complex (the only backbone amides of Val\(^{188}\), Asn\(^{193}\), and Lys\(^{195}\) could not be assigned), and by the observation of numerous NOEs in the NOESY spectra that define the structure of the loop. It is also supported by NOESY spectra of the free and bound forms of the protein, because key long range and intra-loop NOE cross-peaks that are present in the spectra of the bound enzyme are either absent or significantly less intense (~6-fold less intense) in the spectra of the apo-form. Collectively, these data strongly suggest that binding of the substrate causes the \(\beta7/\beta8\) loop to become more structurally ordered and immobilized.

Interestingly, substrate binding indirectly causes conformational ordering of the \(\beta7/\beta8\) loop by displacing the polypeptide backbone containing the active site Cys\(^{187}\) residue. In the complex, the peptide forms a covalent bond with the thiol of Cys\(^{187}\) located at the end of strand \(\beta7\) immediately preceding the \(\beta7/\beta8\) loop. Peptide binding causes the backbone atoms Cys\(^{187}\) to shift by ~7 Å away from the signal, which in turn displaces the \(\beta7/\beta8\) loop (Fig. 5A). As a result, the \(\beta7/\beta8\) forms stabilizing interactions with residues in the \(\beta4/\beta5\) loop; the side chain of Val\(^{190}\) within the \(\beta7/\beta8\) loop packs against Met\(^{128}\) and Ser\(^{129}\) within the \(\beta4/\beta5\) loop. The structural change in the \(\beta7/\beta8\) loop may play an important role in catalysis as this loop has been postulated to form a surface that recognizes the second substrate of catalysis, lipid II (68, 69). This may be a common feature of class A enzymes, as similar positional changes in the \(\beta7/\beta8\) loop have been observed in the S. aureus SrtA enzyme as a result of substrate binding.

Modeling studies suggest that the appendage may transiently mask the thioacyl intermediate. We modeled the thioacyl enzyme-substrate reaction intermediate using the NMR structure of the \(^{18}\)SrtA-LPAT* complex. The coordinates of the methylene group and sulfur atom in the modified threonine residue in the sorting signal were replaced by a carbonyl group, and a thioacyl bond to the active site cysteine was established. The coordinates were then energy minimized to yield a model of the thioacyl reaction intermediate. In test calculations, a range of starting structures were energy minimized in which the active site arginine and carbonyl oxygen were positioned differently. In all cases, these calculations yielded the same basic model shown in Fig. 6 (see “Experimental Procedures” for

---

TABLE 2

| Enzyme hydrolysis kinetics of the \(^{18}\)SrtA enzyme | \(k_{cat}\) | \(K_{m}\) | \(k_{cat}/K_{m}\) |
|---|---|---|---|
| S59G | \(3.3 \times 10^{-8} \pm 1 \times 10^{-8}\) | \(280 \pm 16\) | \(1.2 \times 10^{-7} \pm 8 \times 10^{-9}\) |
| I61A | \(5.8 \times 10^{-7} \pm 5 \times 10^{-8}\) | \(265 \pm 38\) | \(2.2 \times 10^{-7} \pm 4 \times 10^{-8}\) |
| C187A | NA | NA | NA |
| \(^{18}\)SrtA | \(3.6 \times 10^{-7} \pm 2 \times 10^{-6}\) | \(306 \pm 23\) | \(1.2 \times 10^{-7} \pm 1 \times 10^{-8}\) |
| \(^{18}\)SrtA_{mut} | \(5.7 \times 10^{-7} \pm 2 \times 10^{-8}\) | \(173 \pm 11\) | \(3.3 \times 10^{-7} \pm 2 \times 10^{-8}\) |

NA, not active.

---

NMR Structure of B. anthracis SrtA-Substrate Complex

---
In the model of the thioacyl intermediate, the side chain of Arg196 makes extensive contacts with the LPAT peptide. Its side chain is wedged into a pocket that is enclosed by the L-shaped backbone of the peptide, with its guanidino group hydrogen bonding to the backbone carbonyl oxygen of proline, and to the side chain hydroxyl group of the threonine. The latter interaction is compatible with the existence of a substrate-stabilized oxyanion hole as previously postulated for the S. aureus SrtB enzyme (54). This is because the oxygen atom presumably moves by only ~0.5 Å during the transition between sp² and sp³ hybridization of the carbonyl carbon atom, such that the cationic guanidino group can be expected to remain near the oxyanion of the tetrahedral for favorable interactions, and to interact with the γ-hydroxyl group of the threonine. The latter interaction is compatible with the existence of a substrate-stabilized oxyanion hole as previously postulated for the S. aureus SrtB enzyme (54). This is because the oxygen atom presumably moves by only ~0.5 Å during the transition between sp² and sp³ hybridization of the carbonyl carbon atom, such that the cationic guanidino group can be expected to remain near the oxyanion of the tetrahedral for favorable interactions, and to interact with the γ-hydroxyl group of the threonine.

This shields the electrophilic carbon atom in the thioacyl bond from the solvent, thus the appendage may obstruct the nucleophilic attack of the intermediate by lipid II or water. However, the obstruction is transient in nature and does not completely inhibit the isolated enzyme because the N-terminal residues in the appendage frequently detach from the enzyme based on the NMR relaxation data.

FIGURE 5. Comparison of the structure of ⁸⁸SrtA in its apo and LPAT* bound states. A. ribbon drawing showing the superposition of the structures of apo ⁸⁸SrtA (PDB code 2KW8; red) and the ⁸⁸SrtA-LPAT* complex (green). Key residues are represented as sticks and colored orange (⁸⁸SrtA-LPAT*) or magenta (apo ⁸⁸SrtA). An expanded view of the β7/β8 loop is shown. The loop transitions from a disordered state (represented by dotted line) to an ordered state after binding the substrate. Cys187 is displaced by ~7 Å away from the substrate binding pocket, facilitating new contacts between the side chains of Val190, Met128 and Ser129. The N-terminal appendage and Boc-LPAT* are removed for clarity. B, as in panel A, but an expanded view of the N-terminal appendage and its interactions with the active site is shown. Ile61 is displaced when the substrate binds such that it only partially buries His126. The backbone atoms of Ile61-Thr186 and Arg196-Lys104 of the two structures align to a r.m.s. deviation of 0.73 Å. C, scatter plot showing the change in the general order parameter S² after the addition of Boc-LPAT*. A positive number indicates the backbone amide becomes more rigid, whereas a negative number indicates it becomes more disordered. A schematic of secondary structures in the enzyme is shown above the plot. D, histogram plot of the difference in RCI between Boc-LPAT* bound ⁸⁸SrtA and apo ⁸⁸SrtA. The dashed line represents 1 S.D. above the average difference in RCI (random coil index) for all amino acids. A positive value indicates that protein likely becomes less structured upon binding the substrate. A schematic of secondary structures of the enzyme is shown above the plot.

FIGURE 6. Model of the thioacyl intermediate. Expanded view of the active site in the energy minimized model of the ⁸⁸SrtA-LPAT thioacyl intermediate. The peptide substrate LPAT (orange) and active site residues Cys187 and Arg196 (green) are shown as sticks. Hydrogen bonds are indicated by yellow dotted lines. The distance between Thr carbonyl oxygen and Arg196 guanidino C is 3.6 Å.
In addition to the structure of the \( \text{BaSrtA-LPAT}^* \) complex, two structures of sortase enzymes covalently bound to their sorting signal have been reported: the NMR and crystal structures of \( S. \text{ aureus SrtA and SrtB} \) enzymes bound to \( \text{LPAT}^* \) and \( \text{NPQT}^* \) substrate analogs, respectively (54, 68). A detailed comparison of these structures with the \( \text{BaSrtA-LPAT}^* \) complex is beyond the scope of this article. However, as in the \( \text{BaSrtA-LPAT}^* \) complex, the proline in these structures causes the sorting signal to adopt an L-shaped conformation that is critical for positioning of the P4 residue. Moreover, in both the \( \text{BaSrtA-LPAT}^* \) and \( S. \text{ aureus SrtB-NPQT}^* \) structures, the threonine residue in the bound signal points “in” toward the floor of the substrate binding site where it is poised to interact with active site arginine (Arg\( ^{190} \) and Arg\( ^{233} \), in \( \text{BaSrtA and SrtB} \), respectively). As described in detail in Jacobitz \textit{et al.} (54) this suggests \( \text{BaSrtA} \) catalysis is facilitated by the formation of a substrate-stabilized oxyanion hole that may be a common feature in sortase enzymes.

A unique feature of the \( \text{BaSrtA-LPAT}^* \) complex is the presence of an N-terminal appendage that directly interacts with the bound peptide. The N-terminal appendage contacts the active site histidine residue and is poised to hydrogen bond to the backbone of the bound signal via the side chain of Ser\( ^{59} \). The appendage spans residues Asp\( ^{57} \) to Val\( ^{79} \) and contains helix H1. In the structure of the apo-form of the enzyme, the portion preceding helix H1 packs against the imidazole ring of His\( ^{126} \) via contacts from the side chains of Ile\( ^{61} \) and Pro\( ^{64} \). These interactions, along with contacts from Leu\( ^{191} \) located in helix H2, shield the histidine side chain from the solvent (Fig. 5B). In the \( \text{BaSrtA-LPAT}^* \) complex, these interactions are largely preserved, except for contacts between His\( ^{126} \) and Ile\( ^{61} \) that are disrupted because the backbone of Ile\( ^{61} \) and residues that precede it in the primary sequence are slightly displaced (Fig. 5B). This movement positions the side chain of Ser\( ^{59} \) near the bound signal, potentiating hydrogen bonding interactions. The NMR relaxation data indicate that in both apo- and signal-bound forms of \( \text{BaSrtA} \) the entire N-terminal appendage toggles between two distinct conformers, a closed conformation in which it contacts the catalytic core, and an open conformation in which residues including helix H1 of the appendage are detached from the body of the protein and become structurally disordered. The presence of the equilibrium is supported by the \( S^2 \) values of residues in the appendage, because in both the free and substrate-bound forms of the enzyme they exhibit small magnitudes (Fig. 3D). Detachment of the appendage presumably occurs on nanosecond or shorter time scales, because we were unable to detect significant \( R_{\text{ex}} \) values. Moreover, as the \( S^2 \) values progressively become smaller toward the N terminus of the protein the appendage presumably unlatches in a “peeling” motion. As compared with the full-length protein, the construct studied here by NMR is N terminally truncated (it is missing residues 1–56). It is therefore possible that on the cell surface these missing residues could also influence the structure of the N-terminal appendage and/or substrate access to the active site.

Interestingly, binding of the sorting signal may slightly destabilize latching of the N-terminal appendage to the core of enzyme, because the \( S^2 \) values of residues preceding helix H1 in the primary sequence positioned near the active site (Ile\( ^{61} \), Leu\( ^{66} \), and Ala\( ^{67} \)) become smaller when the sorting signal binds to the enzyme (Fig. 5C). Three lines of evidence further support this assertion. First, a comparison of the NOESY spectra of the apo- and complexed forms of the protein reveals that substrate binding causes a 3-fold drop in intensity in the NOE cross-peaks that occur between residues located in the N-terminal appendage and the rest of the protein. Second, an analysis of the random coil index (RCI), a method to measure protein flexibility by comparing backbone chemical shifts to reference random coil shifts (70, 71), shows that residues in the N-terminal appendage have slightly higher RCI values when the \( \text{LPAT}^* \) substrate analog is bound (Fig. 5D). Although the change in RCI is small, they occur in regions flanking helix H1 suggesting that when \( \text{LPAT}^* \) binds there is an increased likelihood that the entire N-terminal appendage detaches from the protein. Last, a comparison of the structures of the apo- and complexed forms of the protein reveals that the substrate disrupts interactions between the side chains of Ile\( ^{61} \) in the appendage and His\( ^{126} \) in the active site, which presumably destabilize the closed form of the appendage. The biological significance of the destabilizing effects of signal binding on the attachment of the appendage is unclear, because the appendage is dynamic in both the free and bound states, and because the destabilizing effects are small.

Transient closure of the appendage over the sorting signal retards nucleophile access to the thioacyl reaction intermediate in vitro. On the cell surface, the first step in protein anchoring involves the formation of an enzyme-protein intermediate in which the threonine in the sorting signal of the protein substrate is connected to Cys\( ^{187} \) via a thioacyl bond. This intermediate is then nucleophilically attacked by lipid II to form the protein-lipid II linked product of the sortase-catalyzed reaction. Our kinetics data probes the first steps of catalysis and monitors the rate at which the thioacyl bond is hydrolyzed by a water molecule that functions as a surrogate for lipid II (Table 2). Previously reported enzyme kinetic studies of the \( \text{BaSrtA} \) enzyme have shown that nucleophilic attack by the water molecule is rate-limiting in the hydrolytic reaction (14, 15). Therefore, our observation that \( k_{\text{cat}} \) increases when a portion of the N-terminal appendage is removed suggests that its presence hinders nucleophile attack of the thioacyl bond. This notion is compatible with the increased \( k_{\text{cat}} \) of the I61A mutant as this alteration presumably disrupts contacts that stabilize positioning of the appendage over the active site. Importantly, the time scale of latching and unlatching of the appendage over the bound sorting signal is significantly shorter than the rate-limited hydrolysis step. Thus, in the isolated enzyme studied here, the N-terminal appendage repeatedly latches and unlatches over the thioacyl-linked sorting signal to hinder, but not completely prevent nucleophile access to the substrate.

The results of computational docking and whole cell display studies suggest that on the cell surface the N-terminal appendage may modulate lipid II access to the thioacyl intermediate. How sortases recognize lipid II remains unknown. To investigate the mechanism of binding we used the program FTSite to predict the binding site for meso-DAP, the amino group con-
containing moeity in the *B. anthracis* lipid II molecule that is joined to the sorting signal by sortase. Two binding calculations were performed, one using the intact *B*.*SrtA-LPAT* complex, and a second in which the coordinates of the N-terminal appendage are removed. Only when a portion of the appendage is removed from *B*.*SrtA* does FTSite predict the existence of a binding site for DAP that is located near the active site (compare Fig. 7, A and B). This site is located adjacent to His126 and the β7/β8 and β4/β3 loops. Attractively, the site is near the electrophilic carbonyl carbon atom in the thioacyl bond, which could be accessed by the nucleophile with only minor rearrangements in the active site. Interestingly, only a few residues in the N-appendage (residues Asp57-Pro84) need to be displaced from the enzyme to unmask the DAP binding site such that helix H1 in the appendage could in principle remain attached during the transpeptidation reaction. The notion that residues Asp57-Pro84 in the N-terminal appendage modulate substrate access is consistent with our whole cell studies, which revealed that S59G and I61A mutations in this segment of the appendage reduced display by ~50 and 40% as compared with the wild-type enzyme, respectively (Fig. 4B).

Similar to *B*.*SrtA*, many class C enzymes that polymerize pili contain an N-terminal appendage that interacts with the active site. This appendage, called a “lid,” is structurally distinct from the *B*.*SrtA* appendage, and it approaches the active site from a different direction. Moreover, unlike the appendage in *B*.*SrtA*, the lid more fully occludes the sorting signal binding site (74). When *B*.*SrtA* is located in its natural context on the extracellular membrane, it is conceivable that its N-terminal appendage forms more extensive interactions with the body of the enzyme similar to the lid structure, which could cause it to fully occlude the binding site for the sorting signal. In this scenario, the N-terminal appendage might have a function similar to the lid, holding the sortase in an inactive state until factors on the cell surface detach it to promote sorting signal binding. Our modeling and biochemical studies also suggest that closure of the N terminus over the bound sorting signal hinders nucleophile access to the thioacyl bond that joins sortase to its protein substrate during the anchoring reaction. Thus, on the cell surface, the appendage may prevent unproductive hydrolysis of the thioacyl-linked enzyme-protein intermediate by masking the thioacyl bond from the solvent until lipid II, or other factors become available to complete the transpeptidation reaction. A deeper understanding of how sortase enzymes are regulated and the mechanism of catalysis will require structure determination of additional reaction intermediates that contain the substrate nucleophile, as well as a systematic investigation of how the enzyme operates on the cell surface.

**Acknowledgment**—We thank Dr. Olaf Schneewind for kindly providing cells and plasmids for our in vivo experiments.

**References**

1. Mock, M., and Fouet, A. (2001) Anthrax. *Annu. Rev. Microbiol.* 55, 647–671
2. Holty, J. E., Bravata, D. M., Liu, H., Olshen, R. A., McDonald, K. M., and Owens, D. K. (2006) Systematic review: a century of inhalational anthrax cases from 1900 to 2005. *Ann. Intern. Med.* 144, 270–280
3. Zink, S. D., and Burns, D. L. (2005) Importance of srtA and srtB for growth of *Bacillus anthracis* in macrophages. *Infect. Immun.* 73, 5222–5228
4. Cascioferro, S., Totsika, M., and Schillaci, D. (2014) Sortase A: an ideal target for anti-virulence drug development. *Microb. Pathog.* 77, 105–112
5. Schneewind, O., and Missiakas, D. (2014) Sec-secretion and sortase-mediated anchoring of proteins in Gram-positive bacteria. *Biochim. Biophys. Acta* 1843, 1687–1697
6. Spirig, T., Weiner, E. M., and Clubb, R. T. (2011) Sortase enzymes in Gram-positive bacteria. *Mol. Microbiol.* 82, 1044–1059
7. Clancy, K. W., Melvin, J. A., and McCafferty, D. G. (2010) Sortase transpeptidases: insights into mechanism, substrate specificity, and inhibition. *Biopolymers* 94, 385–396
8. Hendrickx, A. P., Budzik, J. M., Oh, S. Y., and Schneewind, O. (2011) Architecture at the bacterial surface: sortases and the assembly of pili with isopeptide bonds. *Nat. Rev. Microbiol.* 9, 166–176
9. Maresco, A. W., and Schneewind, O. (2008) Sortase as a target of anti-infective therapy. *Pharmacol. Rev.* 60, 128–141
10. Scott, J. R., and Zähner, D. (2006) Pili with strong attachments: Gram-positive bacteria do it differently. *Mol. Microbiol.* 62, 320–330
11. Suree, N., Jung, M. E., and Clubb, R. T. (2007) Recent advances towards new anti-infective agents that inhibit cell surface protein anchoring in *Staphylococcus aureus* and other Gram-positive pathogens. *Mini. Rev. Med. Chem.* 7, 991–1000
12. Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin,
NMR Structure of B. anthracis SrtA-Substrate Complex

O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R., and Bateman, A. (2010) The Pfam protein families database. *Nucleic Acids Res.* 38, D211–D222

13. Schneewind, O., Model, P., and Fischetti, V. A. (1992) Sorting of protein A to the staphylococcal cell wall. *Cell* 70, 267–281

14. Huang, X., Aulabaugh, A., Ding, W., Kapoor, B., Alksne, L., Tabei, K., and Elledstad, G. (2003) Kinetic mechanism of *Staphylococcus aureus* sortase SrtA. *Biochemistry* 42, 11307–11315

15. 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

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

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

18. Dramsi, S., Trieu-Cuot, P., and Bierne, H. (2005) Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria. *Res. Microbiol.* 156, 289–297

19. Aucher, W., Davison, S., and Fouet, A. (2011) Characterization of the sortase repertoire in *Bacillus anthracis*. *PLoS ONE* 6, e27411

20. Gaspar, A. H., Marraffini, L. A., Glass, E. M., Debord, K. L., Ton-That, H., and Schneewind, O. (2005) *Bacillus anthracis* sortase A (SrtA) anchors LPXTG motif-containing surface proteins to the cell wall envelope. *J. Bacteriol.* 187, 4646–4655

21. Maresso, A. W., Chapa, T. J., and Schneewind, O. (2006) Surface protein IsdC and Sortase B are required for heme-iron scavenging of *Bacillus anthracis*. *J. Bacteriol.* 188, 8145–8152

22. Marraffini, L. A., and Schneewind, O. (2007) Sortase C-mediated anchoring of Bas to the cell wall envelope of *Bacillus anthracis*. *J. Bacteriol.* 189, 6425–6436

23. 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

24. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293

25. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) A- to the staphylococcal cell wall. *Cell* 70, 267–281

26. Viet, S., Aulabaugh, A., Ding, W., Kapoor, B., Alksne, L., Tabei, K., and Elledstad, G. (2003) Kinetic mechanism of *Staphylococcus aureus* sortase SrtA. *Biochemistry* 42, 11307–11315

27. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., and Skelton, N. J. (2006) Protein NMR spectroscopy. *CATINA Verlag, Goldau, Switzerland*

28. Vuister, G. W., and Bax, A. (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bond (JHNHa) coupling constants in 15N-enriched proteins. *J. Am. Chem. Soc.* 115, 7772–7777

29. Gagné, S. M., Tsuda, S., Li, M. X., Chandra, M., Smillie, L. B., and Sykes, B. D. (1994) Quantification of the calcium-induced secondary structural changes in the regulatory domain of tropolin-3C-PepSci. 3, 1961–1974

30. Clore, G. M., Bax, A., and Gronenborn, A. M. (1991) Stereospecific assign-
NMR Structure of B. anthracis SrtA-Substrate Complex

Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Goetz, A. W., Kolosvary, I., Wong, K. F., Paesani, F., Vanicek, J., Wolf, R. M., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M. J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Salomon-Ferrer, R., Sagui, C., Babkin, V., Luchko, T., Gusaov, S., Kovalenko, A., and Kollman, P. A. (2012) AMBER 12, University of California, San Francisco, CA

56. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. J. Comput. Chem. 26, 1781–1802

57. Suree, N., Yi, S. W., Thieu, W., Marohn, M., Damaiseaux, R., Chan, A., Jung, M. E., and Clubb, R. T. (2009) Discovery and structure-activity relationship analysis of Staphylococcus aureus sortase A inhibitors. Bioorg. Med. Chem. 17, 7174–7185

58. Jung, M. E., Clemens, J. J., Suree, N., Liew, C. K., Pilpa, R., Campbell, D. O., and Clubb, R. T. (2005) Synthesis of (2R,3S)-3-amino-4-mercapto-2-butanol, a threonine analogue for covalent inhibition of sortases. Bioorg. Med. Chem. Lett. 15, 5076–5079

59. Jung, M. E., and Yi, S. W. (2012) Synthesis of threo-/H9252-aminoalcohols from aminoaldehydes via chelation-controlled additions: total synthesis of L-threo-sphingosine and safingol. Tetrahedron Lett. 53, 4216–4220

60. Liew, C. K., Smith, B. T., Pilpa, R., Suree, N., Ilangovan, U., Connolly, K. M., Jung, M. E., and Clubb, R. T. (2004) Localization and mutagenesis of the sorting signal binding site on sortase A from Staphylococcus aureus. FEBS Lett. 571, 221–226

61. Schechter, I., and Berger, A. (1967) On the size of the active site in proteases: I. papain. Biochem. Biophys. Res. Commun. 27, 157–162

62. Aucher, W., Davison, S., and Fouet, A. (2011) Characterization of the sortase repertoire in Bacillus anthracis. PLoS ONE 6, e27411

63. Palmer, A. G., 3rd, Kroenke, C. D., and Loria, J. P. (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. Methods Enzymol. 339, 204–238

64. Toller, M., Skrynnikov, N. R., Mulder, F. A., Forman-Kay, J. D., and Kay, L. E. (2001) Slow dynamics in folded and unfolded states of an SH3 domain. J. Am. Chem. Soc. 123, 11341–11352

65. Kleckner, I. R., and Foster, M. P. (2011) An introduction to NMR-based approaches for measuring protein dynamics. Biochim. Biophys. Acta 1814, 942–968

66. Loria, J. P., Berlow, R. B., and Watt, E. D. (2008) Characterization of enzyme motions by solution NMR relaxation dispersion. Acc. Chem. Res. 41, 214–221

67. Schneewind, O., and Missiakas, D. M. (2012) Protein secretion and surface display in Gram-positive bacteria. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367, 1123–1139

68. Suree, N., Liew, C. K., Villareal, V. A., Thieu, W., Fadeev, E. A., Clemens, J. J., Jung, M. E., and Clubb, R. T. (2009) The structure of the Staphylococcus aureus sortase-substrate complex reveals how the universally conserved LPXTG sorting signal is recognized. J. Biol. Chem. 284, 24465–24477

69. Zong, Y., Mazmanian, S. K., Schneewind, O., and Narayana, S. V. (2004) The structure of sortase B, a cysteine transpeptidase that tethers surface protein to the Staphylococcus aureus cell wall. Structure 12, 105–112

70. Berjanskii, M. V., and Wishart, D. S. (2005) A simple method to predict protein flexibility using secondary chemical shifts. J. Am. Chem. Soc. 127, 14970–14971

71. Berjanskii, M. V., and Wishart, D. S. (2008) Application of the random coil index to studying protein flexibility. J. Biomol. NMR 40, 31–48

72. Ngan, C. H., Hall, D. R., Zerbe, B., Grove, L. E., Kozakov, D., and Vajda, S. (2012) FTSite: high accuracy detection of ligand binding sites on unbound protein structures. Bioinformatics 28, 286–287

73. Brenke, R., Kozakov, D., Chuang, G. Y., Beglov, D., Hall, D., Landon, M. R., Mattos, C., and Vajda, S. (2009) Fragment-based identification of drugable “hot spots” of proteins using Fourier domain correlation techniques. Bioinformatics 25, 621–627

74. Manzano, C., Contreras-Martel, C., El Mortaji, L., Joz̆e, T., Fenel, D., Vernet, T., Schoehn, G., Di Guilmi, A. M., and Dessen, A. (2008) Sortase-mediated pilus fiber biogenesis in Streptococcus pneumoniae. Structure 16, 1838–1848