BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

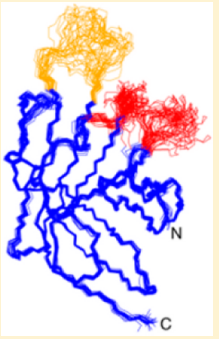
Solution Structure of the Sortase Required for Efficient Production of Infectious *Bacillus anthracis* Spores

Scott A. Robson,[†,‡,∥] Alex W. Jacobitz,[†,‡] Martin L. Phillips,[†] and Robert T. Clubb*[,†,§]

[†]Department of Chemistry and Biochemistry, [‡]UCLA-DOE Institute of Genomics and Proteomics, and [§]Molecular Biology Institute, University of California, Los Angeles, 611 Charles Young Drive East, Los Angeles, California 90095, United States

Supporting Information

ABSTRACT: *Bacillus anthracis* forms metabolically dormant endospores that upon germination can cause lethal anthrax disease in humans. Efficient sporulation requires the activity of the SrtC sortase (BaSrtC), a cysteine transpeptidase that covalently attaches the BasH and BasI proteins to the peptidoglycan of the forespore and predivisional cell, respectively. To gain insight into the molecular basis of protein display, we used nuclear magnetic resonance to determine the structure and backbone dynamics of the catalytic domain of BaSrtC (residues Ser^{56}–Lys^{198}). The backbone and heavy atom coordinates of structurally ordered amino acids have coordinate precision of 0.42 ± 0.07 and 0.82 ± 0.05 Å, respectively. $BaSrtC_{\Delta 55}$ adopts an eight-stranded β-barrel fold that contains two short helices positioned on opposite sides of the protein. Surprisingly, the protein dimerizes and contains an extensive, structurally disordered surface that is positioned adjacent to the active site. The surface is formed by two loops (β2–β3 and β4–H1 loops) that surround the active site histidine, suggesting that they may play a key role in associating BaSrtC with its lipid II substrate. BaSrtC anchors proteins bearing a noncanonical LPNTA sorting signal. Modeling studies suggest that the enzyme recognizes this substrate using a rigid binding pocket and reveals the presence of a conserved subsite for the signal. This first structure of a class D member of the sortase superfamily unveils class-specific features that may facilitate ongoing efforts to discover sortase inhibitors for the treatment of bacterial infections.



Bacterial surface proteins play key roles in microbial physiology and pathogenesis. Gram-positive bacteria display surface proteins using sortase enzymes, a large superfamily of cysteine transpeptidases, which covalently join proteins bearing an appropriate C-terminal cell wall sorting signal (CWSS) to strategically positioned amino groups located on the cell surface.[1−7] Typically, bacteria encode multiple sortases that either attach proteins to the cross-bridge peptide of the cell wall or assemble pili, long proteinaceous structures that extend from the cell surface. Both processes occur through a related transpeptidation reaction that is best characterized for the sortase A enzyme from *Staphylococcus aureus* (SaSrtA). SaSrtA anchors proteins to the cell wall containing a CWSS that consists of an LP*X*TG motif (where *X* denotes any amino acid), followed by a segment of hydrophobic amino acids, and a tail that is comprised primarily of positively charged residues. The C-terminal charged tail presumably retards export, positioning the protein for processing by the extracellular membrane-associated SaSrtA enzyme.

Catalysis occurs through a ping-pong mechanism that is initiated when the active site cysteine residue nucleophilically attacks the backbone carbonyl carbon of the threonine residue within the LP*X*TG motif, breaking the threonine–glycine peptide bond to create a sortase–protein complex in which the components are linked via a thioacyl bond.[8,9] SaSrtA then transfers the protein to the cell wall precursor lipid II, when the amino group in this molecule nucleophilically attacks the thioacyl linkage to create a peptide-linked protein–lipid II product. The transglycosylation and transpeptidation reactions that synthesize the cell wall then incorporate this product into the peptidoglycan, resulting in a surface-displayed protein that is covalently attached to the cross-bridge peptide. Instead of anchoring proteins to the cell wall, some sortases assemble pili.[2,3] These sortases operate through a similar mechanism; however, they link pilin subunits together by joining a lysine amino nucleophile located in one protein to the CWSS of another. A molecular-level understanding of sortase function could lead to new therapeutics for the treatment of bacterial infections, as many clinically significant pathogens are attenuated in their virulence when their sortase enzymes are genetically eliminated.[4,10]

On the basis of primary sequence homology, most sortases in pathogenic Gram-positive bacteria can be grouped into four distinct families, called class A–D enzymes.[7,11,12] Class A enzymes are most closely related to SaSrtA. They appear to perform a housekeeping role in different species of bacteria as members of this group have been shown to anchor a large number of functionally distinct proteins to the cell wall that generally contain an LP*X*TG motif within their CWSS. In contrast, class B–D enzymes appear to have more specialized functions as representative members of these groups display a limited number of proteins that frequently contain non-canonical sorting signals. Class B enzymes are present in

Received: July 2, 2012
Revised: September 12, 2012
Published: September 13, 2012

Firmicutes and can have distinct functions. Some members of this group attach heme receptors to the peptidoglycan,[13] while others function as polymerases that assemble pili. Class C enzymes are also broadly distributed in Gram-positive bacteria and function as pilin polymerases.[14,15] Finally, class D enzymes are present in many bacilli (*Bacillus cereus*, *Bacillus anthracis*, and *Bacillus thuringiensis*). A single member of this family has been characterized, the class D enzyme from *B. anthracis*, which has been shown to attach proteins to the cell wall that facilitate sporulation.[16,17] Atomic structures of representative class A−C enzymes have been determined, revealing a conserved eight-stranded β-barrel fold that houses three essential active site residues: His[120], Cys[184], and Arg[197] (SaSrtA numbering).[1,2,7] The roles of these residues in catalysis have been defined for SaSrtA and are likely conserved in other sortases: Cys[184] acts as a nucleophile that attacks the carbonyl atom in the scissile peptide bond of the CWSS, Arg[197] stabilizes the binding of the sorting signal substrate, and His[120] may act as a general acid or base.

B. anthracis is a Gram-positive facultative anaerobe that causes lethal anthrax disease.[18] Like other bacterial species within the genus *Bacillus*, it forms dormant endospores (spores) that are capable of surviving for long periods of time under harsh conditions. In humans, anthrax caused by the inhalation of aerosolized *B. anthracis* spores has a high rate of mortality and has led to their use as a bioterrorism agent.[19] Efficient sporulation of *B. anthracis* is dependent upon the activity of SrtC (BaSrtC), a class D sortase that attaches the BasH and BasI proteins to distinct cellular structures.[16,17] BasI is attached to the peptidoglycan of predivisional cells, while BasH is exclusively attached to the forespore presumably by BaSrtC inherited from the mother cell before the polar division takes place. *B. anthracis* also encodes a housekeeping class A enzyme (BaSrtA) that anchors a different set of proteins to the cell wall.[20,21] Interestingly, BaSrtC and BaSrtA specifically recognize very closely related sorting signals.[16,17] BaSrtA recognizes a canonical LP*X*TG-type sorting signal present in seven *B. anthracis* proteins (three of these proteins are involved in collagen adhesion, while the functions of the other proteins are not known). In contrast, BaSrtC anchors BasH and BasI that contain LPNTA sorting signals. *B. anthracis* also encodes a third sortase, BaSrtB.[22] This class B enzyme anchors the IsdC heme binding protein to the cell wall by recognizing its unique NPQTN sorting signal. While the sortase enzymes in *B. anthracis* anchor different proteins to the cell wall by recognizing distinct sorting signals, they are believed to attach these proteins to the same chemical group, the *m*-diaminopimelic acid (*m*-DAP) side chain of the cross-bridge peptide.[17,23] In this paper, we describe the structure and dynamics of the *B. anthracis* BaSrtC enzyme, which is the first reported structure of a class D enzyme. The structure provides insight into the mechanism of protein anchoring in pathogenic *B. anthracis*, as well as the evolutionary relationship between different types of enzymes within the sortase enzyme superfamily.

MATERIALS AND METHODS

Cloning, Protein Expression, and Purification. $BaSrtC_{\Delta 19}$ and $BaSrtC_{\Delta 55}$ were amplified via polymerase chain reaction (PCR) from genomic *B. anthracis* DNA (*Sterne* strain) with primers that placed an NdeI restriction site and a BamHI restriction site on the 5′ and 3′ ends of the PCR product, respectively. Each PCR product was digested with NdeI and BamHI restriction enzymes, as was empty vector pET15B (Qiagen). Digested PCR products and pET15B plasmid were ligated together and transformed into *Escherichia coli* XL-1 cells (Stratagene). Successful transformants were confirmed by DNA sequencing. Plasmids where then transformed into *E. coli* BL21(DE3) (Stratagene) for expression.

Protein for enzymatic assays was expressed in BL21(DE3) cells in standard Luria-Bertani broth (LB), at 37 °C. Isotopically labeled protein for nuclear magnetic resonance (NMR) studies was expressed in BL21(DE3) cells grown in M9 medium supplemented with $^{15}NH_4Cl$ and/or [$^{13}C_6$]-glucose. At an A_{600} of ∼0.7, expression was induced by the addition of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). In the case of LB cultures, cells were harvested after 4 h by centrifugation at 6000*g* and stored at −80 °C. Cells grown in M9 media where shifted to 22 °C after induction, with induction allowed to progress overnight. Pellets were resuspended in BugBuster (Novagen) with the addition of phenylmethanesulfonyl fluoride (PMSF) and benzamidine (the final concentration of each was 2 mM) and allowed to lyse at room temperature for 30 min. The lysate was cleared by centrifugation at 13000*g* for 30 min at 4 °C. The soluble fraction was incubated with TALON His-affinity resin (Clontech). The resin was washed with 20 mL of lysis buffer followed by 10 mL of 10 mM imidazole followed by a final washing with thrombin cleavage buffer [20 mM Tris (pH 8.0), 150 mM NaCl, and 2.5 mM $CaCl_2$]. Resin was then treated with thrombin at 37 °C for 2 h, releasing untagged BaSrtC proteins. Protein was further purified by Sephacryl-100 gel filtration in either NMR buffer [20 mM HEPES (pH 6.0)] or peptide cleavage assay buffer (see below).

NMR Spectroscopy and Structure Calculations. $BaSrtC_{\Delta 55}$ was concentrated to 1 mM in NMR buffer [20 mM HEPES (pH 6.0) in 7% D_2O]. In addition, PMSF and ethylenediaminetetraacetic acid (EDTA) were added to final concentrations of 1 mM each. A second sample was produced by lyophilizing $BaSrtC_{\Delta 55}$ in NMR buffer and resolubilizing it with an equal volume of 99.999% D_2O. NMR spectra were recorded at 298 K on Bruker Avance 500, 600, and 800 MHz spectrometers equipped with triple-resonance cryogenic probes. All NMR spectra were processed using NMRPipe[24] and analyzed using CARA (version 1.8.4).[25] Chemical shift assignments (^1H, ^{13}C, and ^{15}N) were obtained by analyzing the following experiments: HNCA, HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, HNHA, HNHB, HBHA(CO)NH, CC(CO)NH, HCCH-TOCSY, HCCH-COSY, (HB)CB(CGCDCE)HE, and (HB)CB(CGCD)HD.[26,27] TALOS+[28] was used to obtain φ and ψ dihedral angle restraints. Additional φ and ψ dihedral angle restraints were determined by measuring relative $HNHA_i$ and $HNHA_{i-1}$ NOE intensities. Distance restraints were obtained from three-dimensional ^{15}N- and ^{13}C-edited NOESY spectra. The collection of residual dipolar couplings (RDCs) was attempted in a PEG/hexanol mixture; however, protein alignment was not achieved. Initial NOE assignments were determined using ATNOS/CANDID[29,30] and simulated annealing in NIH-XPLOR.[31] These NOE assignments were subsequently transferred back to CARA for validation and additional manual picking of NOE assignments. Hydrogen bonds were identified initially on the basis of calculated structures and NOE patterns, verified using deuterium exchange of backbone amides, and included in NIH-XPLOR calculations using the HBDB algorithm.[32] A total of 200 final structures were calculated. Structures with no NOE or

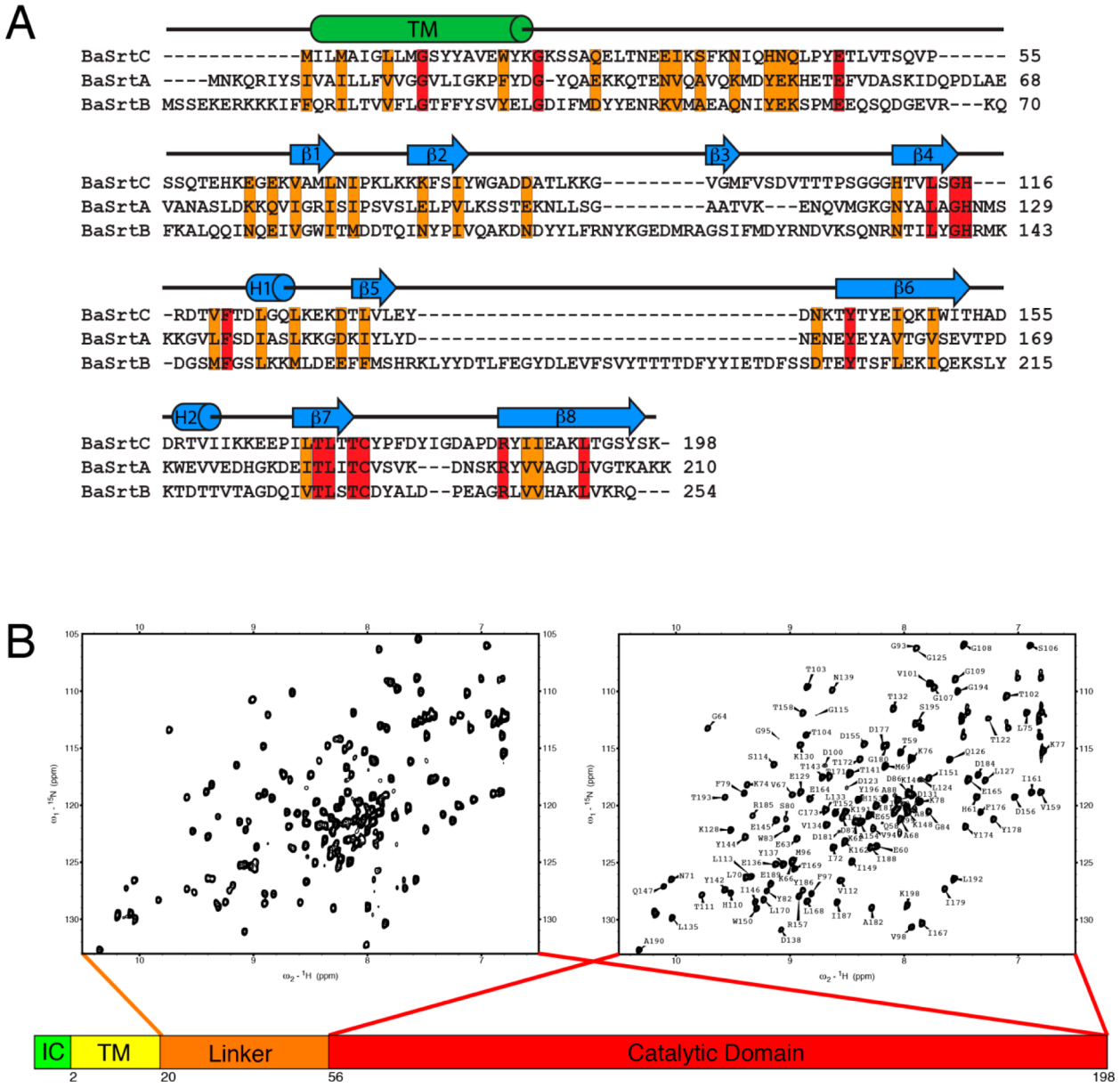


Figure 1. NMR spectra and amino acid sequence of BaSrtC. (A) Sequence alignment of *B. anthracis* (Ames strain) class A sortase (BaSrtA), class B sortase (BaSrtB), and class D sortase (BaSrtC). The sequence alignment was performed by ClustalW.[63] Conserved residues are colored orange, while identical residues are colored red. The predicted transmembrane region (TM) is indicated by a cylinder labeled TM. Secondary structure features of BaSrtC are indicated by cylinders or arrows above the sequence. (B) $^{1}H-^{15}N$ HSQC spectra of $BaSrtC_{\Delta 19}$ (left) and $BaSrtC_{\Delta 55}$ (right). $BaSrtC_{\Delta 19}$ contains approximately 30 more peaks, the majority of which reside toward the center of the spectrum. The $BaSrtC_{\Delta 55}$ HSQC spectrum has its residue assignments indicated.

dihedral angle violations greater than 0.5 Å or 5°, respectively, were selected with the 20 lowest-energy conformers presented. The BaSrtC−LPAT* complex was modeled by aligning the SaSrtA−LPAT* structure [Protein Data Bank (PDB) entry 2KID] with the BaSrtC structure in PyMOL.[33] The enzyme portion of the SaSrtA−LPAT* complex was removed, leaving behind the substrate only. Manual adjustment of the remaining LPAT* substrate was done to remove atom clashes. The NMR structure of the apoenzyme has been deposited as PDB entry 2LN7.

^{15}N Relaxation Measurements. ^{15}N relaxation data (T_1 and T_2) and heteronuclear $\{^{1}H\}-^{15}N$ NOE data were collected on a Bruker Avance 600 MHz spectrometer equipped with a triple-resonance cryogenic probe at an initial protein concentration of 1 mM. Relaxation data and parameters were analyzed in SPARKY,[34] which generated raw relaxation and NOE parameters. Relaxation parameters were analyzed using programs kindly provided by Arthur G. Palmer III (Columbia University, New York, NY) in an attempt to perform "ModelFree" analysis.[35] Using the data from a protein concentration of 1 mM, the ModelFree formalism failed to generate meaningful results. However, we noticed that the R2R1_tm software that utilizes the R_2/R_1 ratio to calculate an approximate correlation time (τ_c) resulted in an unusually high number for a protein of ~16.5 kDa (τ_c at 1 mM ~ 16 ns). We conjectured that monomer−dimer exchange phenomena might result in aberrant tumbling and relaxation behavior, so we measured T_2 and T_1 values at an additional protein

concentration 0.125 mM. To compare our relaxation data to the ModelFree formalism of Lipari and Szabo,[36,37] we used eqs 7a and 7b from ref 36 to plot T_1 versus T_2 for order parameters from $S^2 = 1$ to $S^2 = 0.6$.

Analytical Ultracentrifugation and Enzyme Assays. Runs of sedimentation equilibrium were performed at 25 °C on a Beckman Optima XL-A analytical ultracentrifuge in 12 mm path length double-sector cells. All samples were in 20 mM HEPES (pH 6.0) and 100 mM NaCl with the addition of 5 mM TCEP. Absorption was monitored at 280 nm at sample concentrations of 0.03, 0.015, and 0.0075 mM. Sedimentation equilibrium profiles were measured at 40000 and 50000 rpm. The data were initially fit with a nonlinear least-squares exponential fit for a single ideal species using the Beckman Origin-based software (version 3.01). Analysis of the association behavior used the global analysis software (the multifit option of the software mentioned above) to analyze four scans simultaneously, corresponding to protein at a concentration of 0.03 mM at 40000 and 50000 rpm and a protein at a concentration of 0.015 mM at 40000 and 50000 rpm. Partial specific volumes were calculated from the amino acid composition.

Peptides were synthesized by NEO Biosciences and used without further purification. BaSrtC (10 μM) was incubated with 200 μM peptide substrate in 5 mM $CaCl_2$ and 20 mM HEPES (pH 7.5) at room temperature. Samples (1 μL) were removed and plated with 0.5 μL of 2,5-dihydroxybenzoic acid (DHB) matrix (Acros Organics) dissolved in 50% ethanol and 0.1% trifluoroacetic acid (TFA) and analyzed by MALDI-TOF immediately upon addition of enzyme, and again after 24 h. Each reaction was performed in triplicate, and all samples were spotted and analyzed in triplicate.

RESULTS AND DISCUSSION

We studied the *B. anthracis* BaSrtC enzyme to gain insight into the molecular basis through which it selectively anchors proteins to the cell wall during bacterial sporulation. BaSrtC is 198 amino acids in length and contains a nonpolar N-terminal region (residues Met^1–Tyr^{19}) that likely functions to embed the protein in the membrane. This is followed by a C-terminal region (residues Lys^{20}–Lys^{198}) that has been shown to mediate the in vitro cleavage of an LPNTA peptide that contains the amino acid sequence of the CWSS found in its BasI protein substrate.[16] Inspection of the amino acid sequence of BaSrtC reveals that the C-terminal region contains residues that are homologous to the catalytic domain of SaSrtA as well as a less conserved polypeptide segment that connects the presumed catalytic domain to the N-terminal nonpolar region (Figure 1A). Structural studies have revealed that many sortases can contain ordered appendages that either precede or follow the amino acid sequence of the catalytic domain. We therefore used NMR to study two polypeptides, $BaSrtC_{\Delta 19}$ (residues Lys^{20}–Lys^{198}), which contains the entire C-terminal region, and $BaSrtC_{\Delta 55}$ (residues Ser^{56}–Lys^{198}), which contains only amino acids within the presumed catalytic domain. The ^1H–^{15}N HSQC spectra of these proteins reveal a similar set of well-dispersed backbone amide chemical shifts indicating that these fragments adopt similar three-dimensional structures (Figure 1B). However, the spectrum of $BaSrtC_{\Delta 19}$ also contains an additional ~30 cross-peaks that exhibit narrow line widths and random coil chemical shifts. Surprisingly, this indicates that unlike sortases from other classes, the linker preceding the catalytic domain in BaSrtC is unstructured in the isolated protein.

Previous studies have shown that the intact C-terminal region of BaSrtC selectively cleaves an LPNTA peptide derived from the CWSS of its BasI protein substrate.[16] To ascertain whether only the conserved catalytic domain is required for this enzymatic activity, the ability of $BaSrtC_{\Delta 55}$ to cleave a peptide containing the LPNTA sequence was determined. MALDI-TOF analysis of a reaction mixture containing $BaSrtC_{\Delta 55}$ and the peptide reveals that the isolated catalytic domain cleaves the peptide between the threonine and alanine residues (Figure 2).

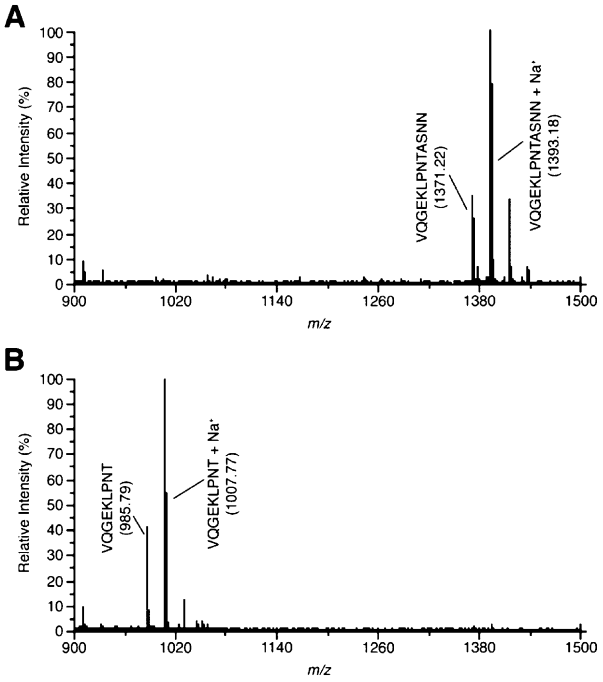


Figure 2. MALDI-TOF demonstrating $BaSrtC_{\Delta 55}$ can cleave the LPNTA-containing peptide. A peptide substrate, VQGEKLPNTASNN, was incubated with $BaSrtC_{\Delta 55}$ for 24 h at room temperature. MALDI spectra of the peptide were taken immediately upon addition of the enzyme (A) and after incubation with $BaSrtC_{\Delta 55}$ for 24 h (B).

Importantly, $BaSrtC_{\Delta 55}$ exhibits specificity for the BasI sorting signal, as it is unable to cleave LPATG and LPETG peptides derived from proteins that are anchored to the cell wall by the BaSrtA sortase (data not shown and ref 16). Combined, these data indicate that in vitro the catalytic domain within BaSrtC is folded and sufficient for hydrolytic activity.

NMR Structure of $BaSrtC_{\Delta 55}$. The structure of $BaSrtC_{\Delta 55}$ was determined using multidimensional heteronuclear NMR and simulated annealing methods. A total of 2584 experimental restraints were employed, including 2330 interproton distance, 196 dihedral angle, and 58 hydrogen bond restraints. Figure 3A displays the ensemble of the 20 lowest-energy conformers of $BaSrtC_{\Delta 55}$. Each structure exhibits good covalent geometry and has no NOE or dihedral angle violations greater than 0.5 Å or 5°, respectively (structural and restraint statistics are listed in Table 1). The structure of $BaSrtC_{\Delta 55}$ is generally well defined by the NMR data; the backbone and heavy atom coordinates of residues Lys^{62}–Ala^{85}, Val^{94}–Gly^{115}, Leu^{124}–Thr^{172}, and Arg^{185}–Lys^{198} have root-mean-square deviations (rmsd's) from the average structure of 0.42 ± 0.07 and 0.82 ± 0.05 Å, respectively. The structure of $BaSrtC_{\Delta 55}$ also contains three

dx.doi.org/10.1021/bi300867t | *Biochemistry* 2012, 51, 7953–7963

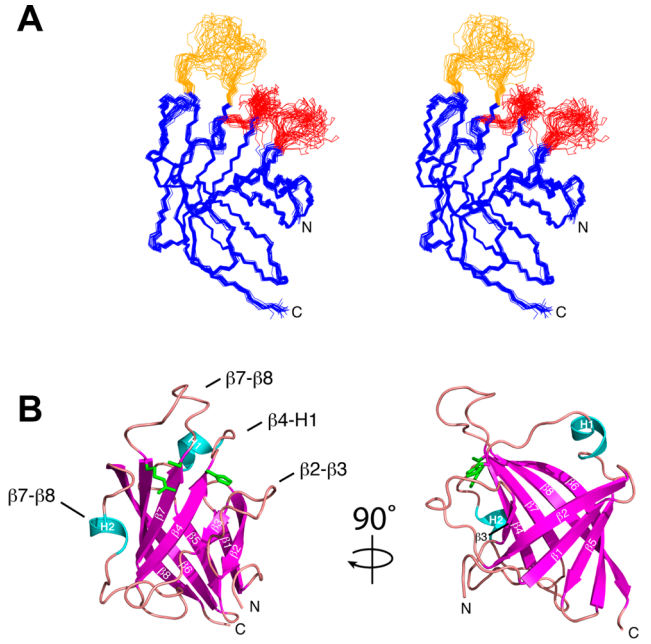


Figure 3. Structures of *B. anthracis* class D sortase. (A) Stereoview of the overlay of the ensemble of NMR structures of $BaSrtC_{\Delta 55}$ (PDB entry 2LN7). The unassigned $\beta 2-\beta 3$ and $\beta 4-H1$ loops are colored red. The assigned but poorly ordered $\beta 7-\beta 8$ loop is colored orange. (B) Ribbon diagram with secondary structure elements labeled. Loop structures are labeled with text. Active site residue side chains for Arg^{185}, Cys^{173}, and His^{86} are colored green.

disordered surface loops that surround the active site (colored red and orange in Figure 3A), two of which (red) likely mediate protein dimerization (vide infra).

$BaSrtC_{\Delta 55}$ adopts an eight-stranded β-barrel fold that contains two short helices positioned on opposite sides of the protein (Figure 3B). The structure is initiated by strand $\beta 1$ ($Val^{67}-Ile^{72}$), which after a short reverse turn interacts with strand $\beta 2$ ($Lys^{77}-Tyr^{82}$) in an antiparallel manner. A poorly ordered 11-amino acid loop then leads into strand $\beta 3$ ($Val^{94}-Met^{96}$), which is positioned parallel and antiparallel to strands $\beta 2$ and $\beta 4$ ($His^{110}-Gly^{115}$), respectively. The active site His^{116} residue is located immediately after strand $\beta 4$ and is followed by a large polypeptide segment containing helix H1 ($Leu^{124}-Gln^{126}$) that stretches to the opposite face of the protein. Strand $\beta 5$ ($Thr^{132}-Glu^{136}$) then joins with $\beta 1$ in an antiparallel manner before the peptide reverses direction to form strand $\beta 6$ ($Thr^{141}-Thr^{152}$). Strand $\beta 6$ wraps around the enzyme toward the active site and contains a sharp point of curvature created by a β-bulge at residues Gln^{147} and Lys^{148}. A long loop containing helix H2 then leads into strand $\beta 7$ ($Ile^{167}-Thr^{172}$), which lies parallel with respect to strand $\beta 4$. A large loop then reverses the directionality of the chain before forming strand $\beta 8$ ($Arg^{185}-Tyr^{197}$) that lies antiparallel with respect to strands $\beta 6$ and $\beta 7$. Strand $\beta 8$ is extensive as it contains a β-bulge at residues Thr^{193} and Gly^{194} that allows it to form a continuous set of hydrogen bonding interactions with strand $\beta 6$ that together contribute residues to both faces of the protein structure.

Disordered Loops Positioned near the Active Site Might Mediate in Vitro Dimerization. On the basis of primary sequence homology, the active site in $BaSrtC_{\Delta 55}$ is formed by three spatially adjacent amino acids: His^{116}, Cys^{173}, and Arg^{185}. These residues are conserved in all sortases, and in

Table 1. Structural Statistics for the Solution Structure of $BaSrtC_{\Delta 55}$[a]

	⟨SA⟩	$(\overline{SA})_r$
rmsd from NOE restraints (Å)[b]		
all (2330)	0.029 ± 0.001	0.046
sequential ($\|i-j\|=1$) (617)	0.032 ± 0.001	0.058
medium-range ($\|i-j\|\leq 4$) (292)	0.039 ± 0.002	0.054
long-range ($\|i-j\|\geq 5$) (977)	0.027 ± 0.001	0.040
intraresidue (444)	0.021 ± 0.001	0.034
rmsd from dihedral angles restraints (deg)[c] (196)	0.182 ± 0.087	0.095
deviation from idealized covalent geometry		
bonds (Å)	0.007 ± 0.003	0.004
angles (deg)	0.414 ± 0.172	0.523
impropers (deg)	0.191 ± 0.081	0.493
PROCHECK results (%)[d]		
most favorable region	85.6 ± 2.6	89.0
additionally allowed region	14.4 ± 2.6	11.0
generously allowed region	0.0 ± 0.0	0.0
disallowed region	0.0 ± 0.0	0.0
coordinate precision (Å)[e,f]		
protein backbone	0.42 ± 0.07	
protein heavy atoms	0.82 ± 0.05	

[a]The notation of the NMR structures is as follows. ⟨SA⟩ represents the final 20 simulated annealing structures. $(\overline{SA})_r$ represents the average energy-minimized structure. The number of terms for each restraint is given in parentheses. [b]None of the structures exhibited distance violations of >0.5 Å or dihedral angle violations of >5°. [c]The experimental dihedral angle restraints were as follows: 95 ϕ, 94 ψ, and 39 χ_1 angular restraints. [d]Determined using PROCHECK.[62] [e]The coordinate precision is defined as the average atomic rmsd of the 20 individual SA structures and their mean coordinates. These values are for residues 63–85, 94–113, 124–172, and 182–198 of $BaSrtC_{\Delta 55}$. Backbone atoms are N, C^{α}, and C′. Assignments were made for residues 58–89, 93–115, 122–183, and 184–198 of $BaSrtC_{\Delta 55}$. [f]Structure calculations also included 58 hydrogen bonds. These bonds were included in xplor structure calculations as HBDB terms as described in ref 32.

$BaSrtC_{\Delta 55}$, they are located at the end of a large cleft formed by strands $\beta 4$, $\beta 7$, and $\beta 8$. Cys^{173} is located at the end of strand $\beta 7$, and in the prototypical SaSrtA enzyme, the analogous residue mediates the nucleophilic attack on the threonine carbonyl carbon within the sorting signal.[38,39] Surrounding the thiol are the side chains of Arg^{185} (strand $\beta 8$) and His^{116} (strand $\beta 4$), which in SaSrtA may stabilize binding of the sorting signal substrate and facilitate acid–base chemistry, respectively.[1] Surprisingly, unlike previously characterized sortases, $BaSrtC_{\Delta 55}$ contains three disordered surface loops that are positioned immediately adjacent to the active site (loops colored red and orange in Figure 3A). The loops are formed by residues that immediately follow the active site cysteine so as to connect strand $\beta 7$ to $\beta 8$ (the $\beta 7-\beta 8$ loop, residues $Tyr^{173}-Asp^{184}$), residues immediately following the active site histidine that connect strand $\beta 4$ to H1 (the $\beta 4-H1$ loop, residues $His^{116}-Asp^{123}$), and residues immediately adjacent to His^{116} that connect strands $\beta 2$ and $\beta 3$ (the $\beta 2-\beta 3$ loop, residues $Trp^{83}-Gly^{93}$). The loops are structurally disordered because an insufficient number of interproton distance restraints were identified in the NOESY spectra to define their position. For the $\beta 2-\beta 3$ and $\beta 4-H1$ loops, few NOEs were observed because many of their residues exhibit resonance line broadening. The broadening is most severe for residues $Leu^{90}-Lys^{92}$ ($\beta 2-\beta 3$ loop) and $His^{116}-Phe^{121}$ ($\beta 4-H1$

dx.doi.org/10.1021/bi300867t | *Biochemistry* 2012, 51, 7953–7963

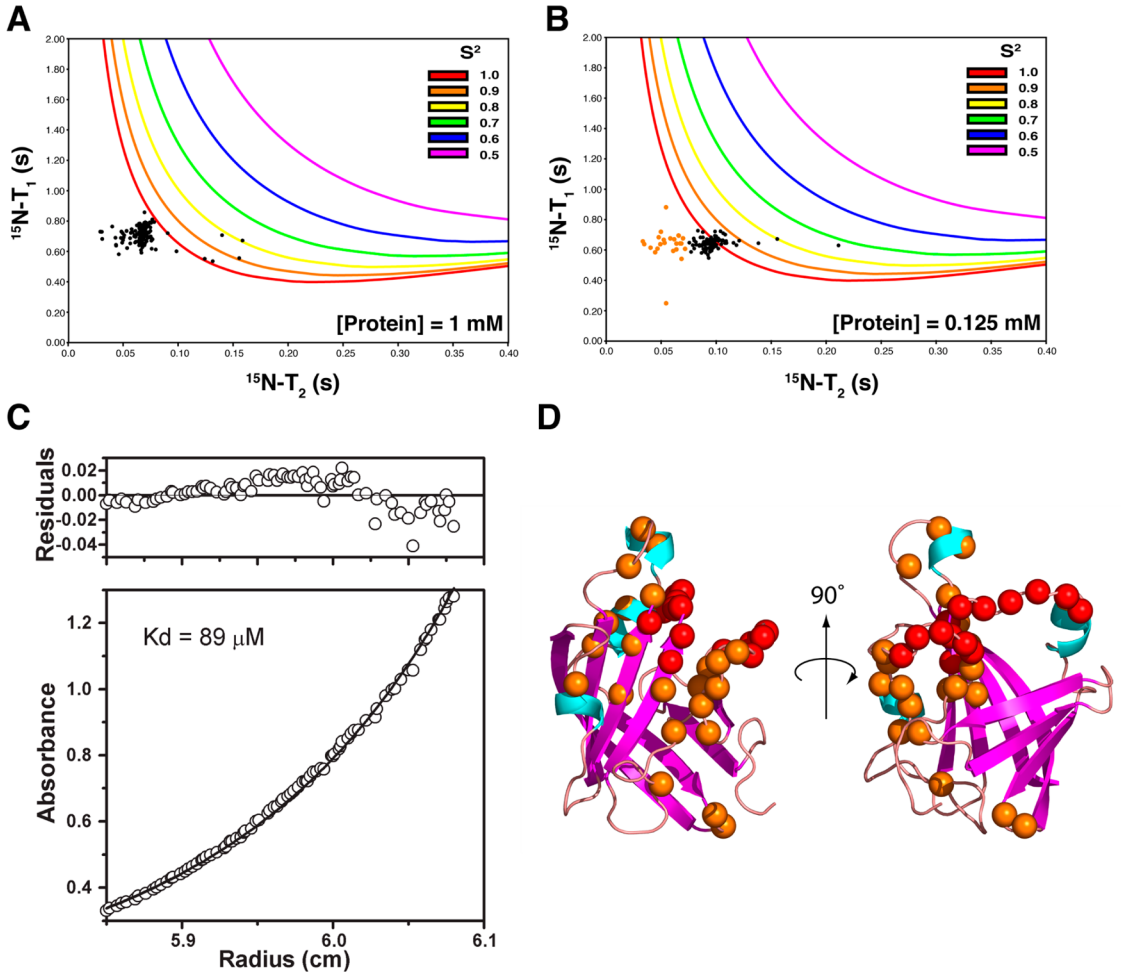


Figure 4. Dynamics and ultracentrifugation analysis. (A) Plot of experimentally determined T_2 vs T_1 data (black dots) for 1 mM BaSrtC$_{\Delta 55}$ (assignable backbone ^{15}N resonances). Also plotted are lines that correspond to calculated T_2 and T_1 values for various S^2 order parameter values and correlation times (correlation times increase along each line in a clockwise fashion). (B) Same as panel A, but T_2 and T_1 data were collected at a protein concentration of 0.125 mM. Residues with T_2 times of <0.07 s are colored orange. (C) Analytical ultracentrifugation equilibrium data demonstrating that BaSrtC is a dimer with a K_D of 89 μM. (D) Structure of BaSrtC with unassigned residue backbone nitrogens colored red. Residue backbone nitrogens with short T_2 times identified in panel B are colored orange.

loop), which are completely absent from the NMR spectra. This suggests that residues within the β2−β3 and β4−H1 loops experience fluctuations in their magnetic environments that occur on a micro- to millisecond time scale, presumably because they are flexible or because they reside within a surface that mediates protein aggregation.

To gain insight into the origin of structural disorder in the three active site loops, we quantitatively probed N−H bond motions in BaSrtC$_{\Delta 55}$ by measuring ^{15}N spin−spin (T_2), ^{15}N spin−lattice (T_1), and {^{1}H}−^{15}N heteronuclear NOE relaxation parameters (Figure S1 of the Supporting Information). In general, the relaxation data are compatible with the structure of BaSrtC$_{\Delta 55}$ as residues within regular secondary structural elements whose coordinates are well defined in the ensemble of conformers have an average {^{1}H}−^{15}N NOE value of 0.79, indicating that they undergo only small amplitude motions on the picosecond time scale. Interestingly, an analysis of the relaxation data suggests that the structurally disordered β2−β3, β4−H1, and β7−β8 loops may participate in transient protein oligomerization. As the chemical shifts of many residues within the disordered β2−β3 and β4−H1 loops are unassigned, the relaxation data do not directly report on their dynamic status. However, several assigned surrounding residues within the β2−β3 and β4−H1 loops, as well as in the adjacently positioned β7−β8 loop, exhibit substantially shortened T_2 values compared to those of other residues in the protein (Figure S1A of the Supporting Information). Intriguingly, this indicates that their backbone amide atoms experience fluctuating magnetic environments on the micro- to millisecond time scale that could result from motions within the loops themselves and/or because the loops reside within or near a molecular surface involved in transient oligomerization.

A more detailed analysis of the NMR relaxation data indicates that active site loops likely form a surface that mediates protein oligomerization. Initially, we attempted to model the dynamics of BaSrtC using the Lipari−Szabo model-free formalism, which yields for each backbone N−H bond vector a generalized order parameter (S^2).[35] S^2 ranges from 0 to 1 and describes the rigidity of the amide bond on the picosecond time scale, with a value of 1 indicating that it is completely immobilized. Surprisingly, fitting of our data consistently yielded many S^2 values of precisely 1.0, which is suggestive of protein aggregation. This is illustrated in Figure 4A, which shows a scatter plot of experimentally derived T_1 and

dx.doi.org/10.1021/bi300867t | Biochemistry 2012, 51, 7953−7963

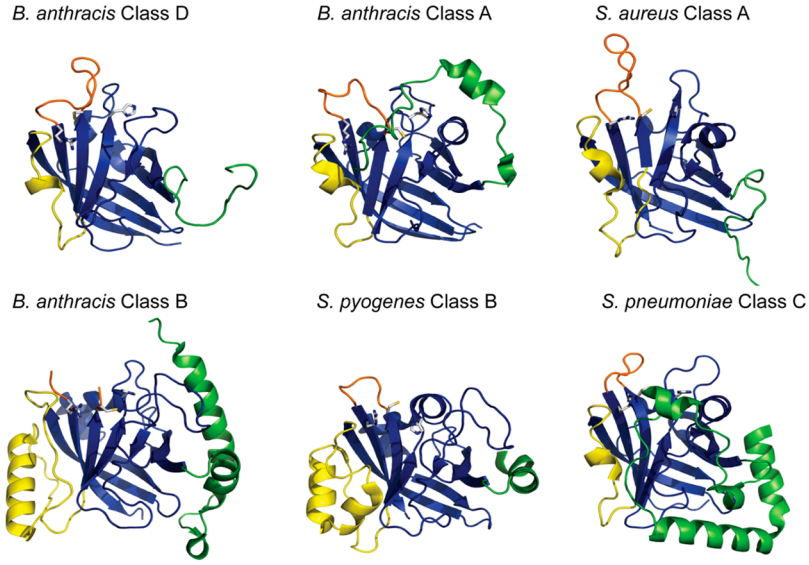


Figure 5. Comparison of representative structures from class A–D sortases. *B. anthracis* class D enzymes are structurally similar to class A enzymes. BaSrtC lacks the long, structured, $\beta6$–$\beta7$ loop (yellow) present in class B enzymes, but the appearance of a short 3_{10}-helix in this region likely indicates the existence of a rigid substrate-binding pocket, similar to BaSrtA. BaSrtC also lacks the structured N-terminal lid present in class C enzymes (green).

T_2 values for each residue overlaid with model-free predicted values of T_1 and T_2 for a range of S^2 values [colored lines indicate values of T_1 and T_2 expected for backbone amide nitrogens containing S^2 values ranging from 0.5 to 1 (see the key)]. The plot reveals that data for most residues are inconsistent with all reasonable values of S^2 because their T_2 and T_1 times are conjointly too short; this results in the software attempting a "best fit" by assigning values of 1 to S^2. We hypothesized that these anomalous results are caused by protein aggregation. As protein aggregation is concentration-dependent, we remeasured the T_1 and T_2 times using a sample of $BaSrtC_{\Delta55}$ that was diluted to 0.125 mM. As shown in Figure 4B, the data for the dilute sample are in better agreement with values predicted by the model-free formalism primarily because the residues now have longer T_2 values. This is most likely because in the dilute sample a smaller fraction of the protein forms higher-molecular weight oligomers that influence the measured relaxation parameters.

To measure the stoichiometry and affinity of protein aggregation, we performed sedimentation equilibrium ultracentrifugation experiments (Figure 4C). The resultant data are best fit by a monomer–dimer equilibrium model that yields a dimer dissociation constant (K_D) of 89 μM. Thus, at the protein concentrations used in our NMR studies, ~90% of $BaSrtC_{\Delta55}$ likely exists in its dimeric form. This explains our difficulty in fitting the relaxation data with the model-free formalism and why analysis of the relaxation data yielded an overall molecular correlation time of ~16 ns, a value that is significantly longer than expected for a protein with $BaSrtC_{\Delta55}$'s molecular mass. Taken together, the NMR and centrifugation data indicate that residues within the $\beta2$–$\beta3$ and $\beta4$–H1 loops likely form a dimerization interface and that these residues are broadened beyond detection in the NMR data because of monomer–dimer exchange that occurs at a rate that is intermediate on the chemical shift time scale. This notion is compatible with the positioning of many residues that exhibit anomalously short T_2 times at high protein concentrations as they are positioned proximal to the $\beta2$–$\beta3$ and $\beta4$–H1 loops (orange points in Figure 4B and orange spheres in Figure 4D). Apparently, these residues also experience monomer–dimer exchange broadening, however to a lesser extent, allowing us to still detect them. In addition to forming a dimer interface, residues within the three disordered loops may also exhibit elevated mobility as several residues within the $\beta2$–$\beta3$ and $\beta7$–$\beta8$ loops exhibit depressed $\{^1H\}$–^{15}N heteronuclear NOEs indicating that they undergo faster picosecond motions.

The biological significance of the weak in vitro dimerization surface detected by NMR and centrifugation is unclear. As BaSrtC is embedded in the membrane, it may indeed dimerize on the cell surface because it undergoes limited two-dimensional diffusion. If this is the case, then class D enzymes might be unique, as previously reported structures of class A–C sortases have shown that they are monomeric[40−55] (it should be noted that biochemical studies have shown that the *S. aureus* SaSrtA enzyme dimerizes with weak affinity, but this finding is controversial[56]). Alternatively, it is also possible that the in vitro dimerization we have observed occurs only because BaSrtC is being studied in isolation. In this scenario, dimerization via the large structurally disordered surface occurs because the appropriate binding partner that normally interacts with this surface is missing. On the cell surface, the missing binding partner could be, among others, protein factors involved in secretion (e.g., the SEC translocon), a component of the cell wall, or BaSrtC's intact lipid II and protein substrates. This is supported by microscopy studies that have shown that the SrtA and SecA proteins in *Streptococcus pyogenes* colocalize at the cross wall compartment where the cell wall is synthesized, and at polar sites where surface protein anchoring also occurs.[57] It would also explain the in vitro catalytic properties of BaSrtC as we and others have been able to demonstrate only that it can hydrolyze its sorting signal (Figure 2 and ref 16). The full transpeptidation reaction using *m*-DAP as a nucleophile has never been demonstrated biochemically. This is presumably because to be fully active, BaSrtC needs to either associate with other factors on the cell surface or bind to its intact lipid II and protein substrates. The notion that the disordered surface

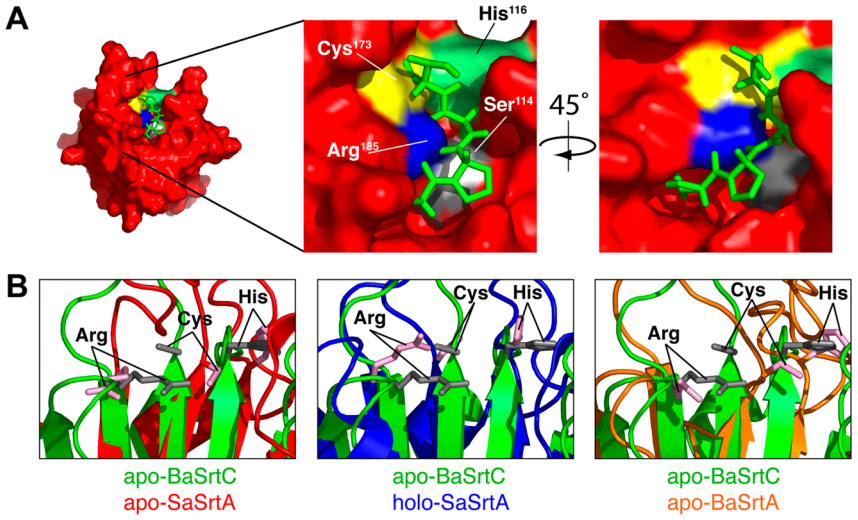


Figure 6. (A) Model of the CWSS peptide structure from the SaSrtA−LPAT* complex on the substrate binding surface of BaSrtC. The peptide (green) is taken from the SaSrtA−LPAT* structure without modification. Active site residues are indicated along with the potential Ser^{114} binding site. The hydroxyl oxygen of Ser^{114} (Ser^{114} is colored gray, while its OH group is colored white) comes into close contact with the side chain of position *X*. (B) Close-up overlay of the BaSrtC structure (green) with apo-SaSrtA (red, left), holo-SaSrtA (blue, center), and apo-BaSrtA (orange, right). Active site arginine, cysteine, and histidine residues are denoted. The active site residues of apo-BaSrtC overlay best with those of holo-SaSrtA (SaSrtA−LPAT*).

mediates substrate recognition on the cell surface is supported by NMR and X-ray studies of SaSrtA and SaSrtB that have implicated the $\beta7-\beta8$ loop in binding the cross-bridge peptide.[49,53] Furthermore, it is interesting to note that of all the sortase enzymes whose structures have been determined only the sortases from *B. anthracis* (BaSrtA, BaSrtB, and BaSrtC) contain a disordered $\beta7-\beta8$ loop and only these enzymes attach proteins to lipid II molecules with a *m*-DAP group.[50,51] Thus, on the cell surface, the large disordered loop regions identified by NMR may play a key role in associating BaSrtC with protein factors and substrates that facilitate cell surface protein anchoring.

Class A and D Sortases Are Structurally Related. Structures of representative class A−C sortases have been determined. Despite a limited degree of sequence homology, BaSrtC is most structurally similar to class A enzymes based on a DALI analysis. Three structures of class A enzymes have been reported: SaSrtA, BaSrtA, and the SrtA enzyme from *St. pyogenes* (SpSrtA).[41,48−50,52] The sequence of BaSrtC is only 23−24% identical with the sequences of these enzymes; however, the coordinates of its carbon atoms that participate in regular secondary structures can be superimposed with rmsd's of 2.4, 1.9, and 2.0 Å, respectively. Notably, both class A and D sortases contain a short 3_{10}-helix located within the $\beta6-\beta7$ loop (Figure 5, yellow). In SaSrtA, NMR studies have shown that this loop interacts with the leucine-proline portion of the LP*X*TG sorting signal, suggesting that BaSrtC will also recognize its distinct LPNTA sorting signal through a generally similar mechanism (vide infra).[49] Interestingly, class A enzymes also exhibit structural heterogeneity in their N-termini, because BaSrtA possesses a long N-terminal appendage that contacts the active site histidine. This N-terminal appendage is disordered in the structures of BaSrtC, SaSrtA, and SpSrtA (Figure 5, green). Although the BaSrtC enzyme adopts a canonical sortase fold, its structure differs markedly from class B and C enzymes. For example, in contrast to class B enzymes that possess a large structured $\beta6-\beta7$ loop that contains an α-helix, the analogous loop in BaSrtC is substantially shorter and lacks a similarly positioned helix (Figure 5, yellow).[42,51,53] Unlike class B enzymes, $BaSrtC_{\Delta55}$ is also missing the N-terminal α-helices that precede the catalytic domain (Figure 5, green). The structure of $BaSrtC_{\Delta55}$ also differs substantially from that of class C enzymes as it is missing an N-terminal "lid" appendage that has been proposed to regulate access to the active site in several members of this group.[44,46,54,58]

Model of the $BaSrtC_{\Delta55}$−Sorting Signal Complex. To gain insight into how $BaSrtC_{\Delta55}$ recognizes its sorting signal, we modeled the structure of its substrate complex by superimposing the backbone coordinates of $BaSrtC_{\Delta55}$ onto the coordinates of our previously determined solution structure of the SaSrtA enzyme covalently bound to an LPAT peptide analogue.[49,59] Inspection of the model reveals that BaSrtC and SaSrtA can bind to their respective sorting signals in a similar manner as there is minimal atomic overlap between the peptide and BaSrtC atoms in the model of the BaSrtC−peptide complex (Figure 6A). In the model, the sorting signal peptide adopts a kinked structure and rests in a groove whose base is formed by residues in strands $\beta4$ and $\beta7$ and whose walls are formed by the $\beta6-\beta7$, $\beta7-\beta8$, $\beta3-\beta4$, and $\beta2-$H2 surface loops. Because the experimentally determined structure of the apo form of BaSrtC can readily accommodate the peptide, this suggests that the enzyme may recognize its sorting signal substrate through a lock-and-key mechanism. This is in marked contrast to SaSrtA, which undergoes major changes in the structure and dynamics of the $\beta6-\beta7$ loop upon signal binding.[49,60] This can be seen in Figure 6B. The left panel shows apo-SaSrtA (red) aligned with our BaSrtC structure (green). While the active site residues of apo-SaSrtA (pink) align well with the active site residues of BaSrtC (gray), the agreement is significantly better than that of the holo-SaSrtA (SaSrtA−LPAT*) structure in Figure 6B [middle panel (blue)]. Furthermore, the BaSrtC active site residues more closely resemble the position of holo-SaSrtA than that of the closely related apo form of the class A sortase from *B. anthracis*, apo-BaSrtA (Figure 6B, right panel, orange). This result further

dx.doi.org/10.1021/bi300867t | *Biochemistry* 2012, 51, 7953−7963

suggests that BaSrtC exists in an "active-ready" form prior to binding its cognate sorting signal.

To facilitate further discussion of the recognition process, hereafter we refer to amino acids within the signal as P and P′ if they precede and follow the scissile peptide bond, respectively [e.g., the LPNTA signal-recognized BaSrtC is Leu (P4)-Pro (P3)-Asn (P2)-Thr (P1)-Ala (P′1)]. Corresponding binding sites on the enzyme for these residues are termed S and S′, respectively. The model clearly defines enzyme subsites that can accommodate the leucine (P4) and proline (P3) residues in the LPNTA signal. The leucine residue is recognized by a small pocket on the enzyme that is formed by the $\beta6-\beta7$ loop and residues in the underlying β-sheet. Within the S4 subsite, the leucine is contacted by the side chains of Val[166], Pro[168], Val[173], and Val[174] on the $\beta6-\beta7$ loop and Val[198] on strand $\beta8$. The S3 subsite, which recognizes the LPNTA signal's proline residue, is formed by residues in the underlying β-sheet [Ala[124] ($\beta4$) and Ile[185] ($\beta7$)] and Val[110] within the adjacent $\beta3-\beta4$ loop. The structures of the class A BaSrtA and SpSrtA enzymes also contain similarly shaped nonpolar subsites,[48,50] which is compatible with bioinformatics studies that have predicted that the majority of sorting signals processed by class A and D enzymes contain proline and leucine amino acids at positions P3 and P4, respectively.[11] In *B. anthracis*, the class A BaSrtA and class D BaSrtC enzymes anchor proteins to the cell wall that contain closely related LP*X*TG and LPNTA sorting signals, respectively. Our modeling studies suggest both enzymes recognize the conserved leucine-proline element of these signals through structurally conserved subsites that do not require substrate binding to form. How each distinguishes differences that occur at the P1′ position cannot be predicted as the coordinate precision of amino acids that contact this residue is poorly defined in the NMR ensembles of both enzymes. However, BaSrtC may in part distinguish the unique signal present in its BasH and BasI substrates by interacting with the P2 amino acid that is invariably an asparagine, as the bottom of BaSrtC's S2 subsite contains Ser[114], which can presumably favorably interact with the asparagine amide group through hydrogen bonding (Figure 6A, gray), whereas BaSrtA contains a hydrophobic residue (Ala[124]).

At present, more than 800 sortase genes have been identified in nearly 300 species of bacteria. The functions of the vast majority of these enzymes are not known because most bacteria contain multiple sortases with unknown substrate specificities. Work described here lays the foundation for more detailed structure–function studies that will elucidate the molecular basis of substrate specificity and could facilitate future efforts to predict sortase function and to discover therapeutically useful sortase inhibitors. Structural data could aid inhibitor discovery efforts in several ways.[61] First, atomic structures can be used to discover new chemical scaffolds that are likely to bind and inhibit sortases. In this approach, an in silico screen of compound libraries could be performed to identify small molecules that have the appropriate physicochemical properties to efficiently interact with the enzyme active site. These molecules serve as new leads for further development after their inhibitory properties have been confirmed experimentally. Second, the structures can be used to optimize the binding affinity and selectivity of established sortase inhibitors. In this approach, computational methods are used to dock the lead inhibitor molecule to the structure of the enzyme. Analogues that are likely to have increased potency are then identified in silico by modifying the bound inhibitor and calculating an algorithm-dependent binding energy. The most promising of these analogues are then tested biochemically to ascertain whether they have improved activity. Finally, if the structure of the enzyme–substrate complex is known, computational methods can be used to identify small molecules with physicochemical features similar to those of the substrate. These molecules presumably bind and inhibit the enzyme that can be tested experimentally. As several sortases with known structures recognize sorting signals with distinct amino acid sequences, their unique active site features in combination with the aforementioned computational methods may allow class- and/or bacterial species-specific inhibitors to be developed. Such molecules are urgently needed as infections caused by MRSA and other multidrug-resistant bacteria are a major health concern.

■ ASSOCIATED CONTENT

Supporting Information

Plots of all NMR relaxation data collected in this study. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Address: 611 Charles E. Young Dr. E., Boyer Hall, Rm. 602, Los Angeles, CA 90095-1570. E-mail: rclubb@mbi.ucla.edu. Phone: (310) 206-2334.

Present Address

[‖]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Ave., Boston, MA 02115.

Funding

This work was supported by National Institutes of Health Grant AI52217 (R.T.C.) and the Ruth L. Kirschstein National Research Service Award (Grant GM007185, UCLA Cellular and Molecular Biology Training Grant for A.W.J.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We express great appreciation to Dr. Joseph Loo and members of the Clubb laboratory for their assistance.

■ ABBREVIATIONS

CWSS, cell wall sorting signal; *m*-DAP, *meso*-diaminopimelic acid; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; NMR, nuclear magnetic resonance.

■ REFERENCES

(1) Clancy, K. W., Melvin, J. A., and McCafferty, D. G. (2010) Sortase transpeptidases: Insights into mechanism, substrate specificity, and inhibition. *Biopolymers 94*, 385–396.

(2) Hendrickx, A. P., Budzik, J. M., Oh, S. Y., and Schneewind, O. (2011) Architects at the bacterial surface: Sortases and the assembly of pili with isopeptide bonds. *Nat. Rev. Microbiol. 9*, 166–176.

(3) Mandlik, A., Swierczynski, A., Das, A., and Ton-That, H. (2008) Pili in Gram-positive bacteria: Assembly, involvement in colonization and biofilm development. *Trends Microbiol. 16*, 33–40.

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

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

(6) Scott, J. R., and Barnett, T. C. (2006) Surface proteins of Gram-positive bacteria and how they get there. *Annu. Rev. Microbiol. 60*, 397−423.
(7) Spirig, T., Weiner, E. M., and Clubb, R. T. (2011) Sortase enzymes in Gram-positive bacteria. *Mol. Microbiol. 82*, 1044−1059.
(8) 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.
(9) Huang, X., Aulabaugh, A., Ding, W., Kapoor, B., Alksne, L., Tabei, K., and Ellestad, G. (2003) Kinetic mechanism of *Staphylococcus aureus* sortase SrtA. *Biochemistry 42*, 11307−11315.
(10) 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.
(11) Comfort, D., and Clubb, R. T. (2004) A comparative genome analysis identifies distinct sorting pathways in Gram-positive bacteria. *Infect. Immun. 72*, 2710−2722.
(12) 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.
(13) Mazmanian, S. K., Ton-That, H., Su, K., and Schneewind, O. (2002) An iron-regulated sortase anchors a class of surface protein during *Staphylococcus aureus* pathogenesis. *Proc. Natl. Acad. Sci. U.S.A. 99*, 2293−2298.
(14) Ton-That, H., Marraffini, L. A., and Schneewind, O. (2004) Protein sorting to the cell wall envelope of Gram-positive bacteria. *Biochim. Biophys. Acta 1694*, 269−278.
(15) Ton-That, H., and Schneewind, O. (2003) Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol. Microbiol. 50*, 1429−1438.
(16) Marraffini, L. A., and Schneewind, O. (2006) Targeting proteins to the cell wall of sporulating *Bacillus anthracis*. *Mol. Microbiol. 62*, 1402−1417.
(17) Marraffini, L. A., and Schneewind, O. (2007) Sortase C-mediated anchoring of BasI to the cell wall envelope of *Bacillus anthracis*. *J. Bacteriol. 189*, 6425−6436.
(18) Koch, R. (1876) Die Ätiologie der Milzbrand-Krankheit, begründet auf die Entwicklungsgeschichte des *Bacillus anthracis*. *Beitr. Biol. Pflanz. 2*, 277−310.
(19) Brookmeyer, R., and Blades, N. (2002) Prevention of inhalational anthrax in the U.S. outbreak. *Science 295*, 1861.
(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) Aucher, W., Davison, S., and Fouet, A. (2011) Characterization of the sortase repertoire in *Bacillus anthracis*. *PLoS One 6*, e27411.
(22) 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.
(23) Budzik, J. M., Oh, S. Y., and Schneewind, O. (2008) Cell wall anchor structure of BcpA pili in *Bacillus anthracis*. *J. Biol. Chem. 283*, 36676−36686.
(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) Keller, R. (2004) *The Computer Aided Resonance Assignment Tutorial*, 1st ed., CANTINA Verlag, Goldav, Switzerland.
(26) Palmer, A. G., III (1995) *Protein NMR Spectroscopy*, Academic Press, San Diego.
(27) Teng, Q. (2005) *Structural Biology: Practical NMR Applications*, Springer, Berlin.
(28) Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR 44*, 213−223.
(29) Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J. Biomol. NMR 24*, 171−189.
(30) Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol. 319*, 209−227.
(31) Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson. 160*, 65−73.
(32) Grishaev, A., and Bax, A. (2004) An empirical backbone-backbone hydrogen-bonding potential in proteins and its applications to NMR structure refinement and validation. *J. Am. Chem. Soc. 126*, 7281−7292.
(33) Delano, W. L. (2006) *The PyMOL Molecular Graphics System*, version 0.99, DeLano Scientific, South San Francisco, CA.
(34) Goddard, T. D., and Kneller, D. G. (2001) *Sparky NMR Analysis Software*, University of California, San Francisco.
(35) Palmer, A. G., III (2001) NMR probes of molecular dynamics: Overview and comparison with other techniques. *Annu. Rev. Biophys. Biomol. Struct. 30*, 129−155.
(36) Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc. 104*, 4559−4570.
(37) Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc. 104*, 4546−4559.
(38) 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.
(39) Ton-That, H., Liu, G., Mazmanian, S. K., Faull, K. F., and Schneewind, O. (1999) Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc. Natl. Acad. Sci. U.S.A. 96*, 12424−12429.
(40) Cozzi, R., Malito, E., Nuccitelli, A., D'Onofrio, M., Martinelli, M., Ferlenghi, I., Grandi, G., Telford, J. L., Maione, D., and Rinaudo, C. D. (2011) Structure analysis and site-directed mutagenesis of defined key residues and motives for pilus-related sortase C1 in group B *Streptococcus*. *FASEB J. 25*, 1874−1886.
(41) Ilangovan, U., Ton-That, H., Iwahara, J., Schneewind, O., and Clubb, R. T. (2001) Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U.S.A. 98*, 6056−6061.
(42) Kang, H. J., Coulibaly, F., Proft, T., and Baker, E. N. (2011) Crystal structure of Spy0129, a *Streptococcus pyogenes* class B sortase involved in pilus assembly. *PLoS One 6*, e15969.
(43) Lu, G., Qi, J., Gao, F., Yan, J., Tang, J., and Gao, G. F. (2011) A novel 'open-form' structure of sortaseC from *Streptococcus suis*. *Proteins: Struct., Funct., Bioinf. 79*, 2764−2769.
(44) Manzano, C., Contreras-Martel, C., El Mortaji, L., Izore, 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.
(45) Manzano, C., Izore, T., Job, V., Di Guilmi, A. M., and Dessen, A. (2009) Sortase activity is controlled by a flexible lid in the pilus biogenesis mechanism of Gram-positive pathogens. *Biochemistry 48*, 10549−10557.
(46) Neiers, F., Madhurantakam, C., Falker, S., Manzano, C., Dessen, A., Normark, S., Henriques-Normark, B., and Achour, A. (2009) Two crystal structures of pneumococcal pilus sortase C provide novel insights into catalysis and substrate specificity. *J. Mol. Biol. 393*, 704−716.
(47) Persson, K. (2011) Structure of the sortase AcSrtC-1 from *Actinomyces oris*. *Acta Crystallogr. D67*, 212−217.
(48) Race, P. R., Bentley, M. L., Melvin, J. A., Crow, A., Hughes, R. K., Smith, W. D., Sessions, R. B., Kehoe, M. A., McCafferty, D. G., and Banfield, M. J. (2009) Crystal structure of *Streptococcus pyogenes*

sortase A: Implications for sortase mechanism. *J. Biol. Chem. 284*, 6924−6933.
(49) 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.
(50) 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.
(51) Zhang, R., Wu, R., Joachimiak, G., Mazmanian, S. K., Missiakas, D. M., Gornicki, P., Schneewind, O., and Joachimiak, A. (2004) Structures of sortase B from *Staphylococcus aureus* and *Bacillus anthracis* reveal catalytic amino acid triad in the active site. *Structure 12*, 1147−1156.
(52) Zong, Y., Bice, T. W., Ton-That, H., Schneewind, O., and Narayana, S. V. (2004) Crystal structures of *Staphylococcus aureus* sortase A and its substrate complex. *J. Biol. Chem. 279*, 31383−31389.
(53) 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.
(54) Khare, B., Fu, Z. Q., Huang, I. H., Ton-That, H., and Narayana, S. V. (2011) The crystal structure analysis of group B *Streptococcus* sortase C1: A model for the "lid" movement upon substrate binding. *J. Mol. Biol. 414*, 563−577.
(55) Khare, B., Krishnan, V., Rajashankar, K. R., I-Hsiu, H., Xin, M., Ton-That, H., and Narayana, S. V. (2011) Structural differences between the *Streptococcus agalactiae* housekeeping and pilus-specific sortases: SrtA and SrtC1. *PLoS One 6*, e22995.
(56) Lu, C., Zhu, J., Wang, Y., Umeda, A., Cowmeadow, R. B., Lai, E., Moreno, G. N., Person, M. D., and Zhang, Z. (2007) *Staphylococcus aureus* sortase A exists as a dimeric protein *in vitro*. *Biochemistry 46*, 9346−9354.
(57) Raz, A., and Fischetti, V. A. (2008) Sortase A localizes to distinct foci on the *Streptococcus pyogenes* membrane. *Proc. Natl. Acad. Sci. U.S.A. 105*, 18549−18554.
(58) Lu, G., Qi, J., Gao, F., Yan, J., Tang, J., and Gao, G. F. (2011) A novel "open-form" structure of sortaseC from *Streptococcus suis*. *Proteins 79*, 2764−2769.
(59) 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.
(60) Naik, M. T., Suree, N., Ilangovan, U., Liew, C. K., Thieu, W., Campbell, D. O., Clemens, J. J., Jung, M. E., and Clubb, R. T. (2006) *Staphylococcus aureus* Sortase A transpeptidase. Calcium promotes sorting signal binding by altering the mobility and structure of an active site loop. *J. Biol. Chem. 281*, 1817−1826.
(61) Taboureau, O., Baell, J. B., Fernández-Recio, J., and Villoutreix, B. O. (2012) Established and Emerging Trends in Computational Drug Discovery in the Structural Genomics Era. *Chem. Biol. 19*, 29−41.
(62) Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR 8*, 477−486.
(63) Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2002) Multiple sequence alignment using ClustalW and ClustalX. *Current Protocols in Bioinformatics*, Chapter 2, Unit 2, 3, Wiley, New York.