Structures of *Streptococcus pyogenes* class A sortase in complex with substrate and product mimics provide key details of target recognition

D. Alex Johnson[‡], Isabel M. Piper[‡], Brandon A. Vogel[‡], Sophie N. Jackson[‡], Justin E. Svendsen,
Hanna M. Kodama, Darren E. Lee, Katy M. Lindblom, James McCarty, John M. Antos*, and Jeanine F. Amacher*
From the Department of Chemistry, Western Washington University, Bellingham, Washington, USA

Edited by Joseph Jez

The cell wall is a critical extracellular barrier for bacteria and many other organisms. In bacteria, this structural layer consists of peptidoglycan, which maintains cell shape and structural integrity and provides a scaffold for displaying various protein factors. To attach proteins to the cell wall, Gram-positive bacteria utilize sortase enzymes, which are cysteine transpeptidases that recognize and cleave a specific sorting signal, followed by ligation of the sorting signal–containing protein to the peptidoglycan precursor lipid II (LII). This mechanism is the subject of considerable interest as a target for therapeutic intervention and as a tool for protein engineering, where sortases have enabled sortase-mediated ligation or *sortagging* strategies. Despite these uses, there remains an incomplete understanding of the stereochemistry of substrate recognition and ligation product formation. Here, we solved the first structures of sortase A from *Streptococcus pyogenes* bound to two substrate sequences, LPATA and LPATS. In addition, we synthesized a mimetic of the product of sortase-mediated ligation involving LII (LPAT-LII) and solved the complex structure in two ligand conformations. These structures were further used as the basis for molecular dynamics simulations to probe sortase A-ligand dynamics and to construct a model of the acyl–enzyme intermediate, thus providing a structural view of multiple key states in the catalytic mechanism. Overall, this structural information provides new insights into the recognition of the sortase substrate motif and LII ligation partner and will support the continued development of sortases for protein engineering applications.

Bacterial sortases are cysteine transpeptidase enzymes that play important roles at the cell wall of Gram-positive bacteria. Despite over 20 years since the discovery of the first sortase enzyme in *Staphylococcus aureus*, a complete picture of how these critical enzymes recognize their ligands has remained elusive because of limited structural information involving sortases in complex with their substrates (1–3). This type of characterization is essential to understanding how sortases perform their role of attaching protein factors to the bacterial cell wall (4). A thorough understanding of this process is also relevant to human health and disease in two significant ways; sortases are used in protein engineering, for example, sortase-mediated ligation (SML), *sortagging*, or *sortylation* applications, and sortases are also therapeutic targets for the development of antibiotics (5, 6).

Sortases are widespread in Gram-positive bacteria and are currently grouped into multiple classes (A–F), including several that are considered general housekeeping enzymes (*e.g.*, classes A and E), and those that assemble pili (class C) (4). The sortase mechanism involves two catalytic steps: (i) recognition and cleavage of a target sequence, and formation of an acyl–enzyme, followed by (ii) nucleophilic attack by a second reactant, initiating a ligation reaction that creates a new peptide bond or isopeptide in the case of the bacterial pilus (7, 8). For class A sortases, the general consensus sequence, which is found within the cell wall sorting signal (CWSS), includes a pentapeptide motif, Leu-Pro-X-Thr-Gly (or LPXTG), where X = any amino acid (4). Positions are defined with respect to the location of the cleavage site between the threonine and glycine residues, with P1' = Gly, P1 = Thr, P2 = X, P3 = Pro, and P4 = Leu (9). For protein anchoring to the bacterial cell surface, the nucleophile in the second step of the reaction mechanism is the cell-wall precursor lipid II (LII), thus allowing for incorporation of the protein into the growing peptidoglycan layer (10).

The majority of knowledge to date on sortase structure and mechanism is focused on class A sortases; however, there are available structures of representative sortases from all six classes (A–F), for example, class A (Protein Data Bank [PDB] ID: 2KID), class B (PDB ID: 1NG5), class C (PDB ID: 3O0P), class D (PDB ID: 2LN7), class E (PDB ID: 5CUW), and class F (PDB ID: 5UUS) (9). These structures have revealed that sortases share a conserved core antiparallel eight-stranded β-barrel structure, termed the *sortase fold* (8, 11). This was first identified in the *S. aureus* class A sortase (saSrtA) structure and is consistently found in wildtype and chimeric sortase enzymes (9, 11, 12). As of

---

[‡] These authors contributed equally to this work.
* For correspondence: Jeanine F. Amacher, amachej@wwu.edu; John M. Antos, antosj@wwu.edu.
Present address for D. Alex Johnson, Isabel M. Piper, Justin E. Svendsen: Department of Bioengineering, California Institute of Technology, Pasadena, CA (DAJ), Department of Chemistry, University of California Berkeley, Berkeley, CA (IMP), and Department of Chemistry and Biochemistry, University of Oregon, Eugene, OR (JES).

early 2022, there were over 65 structures of sortases in the PDB, including from all six classes and SrtA structures from 10 different organisms. Despite this, there is a notable lack of structural information about ligand recognition in sortases. Of the three SrtA structures that contain ligands, two approximate the acyl–enzyme intermediates of saSrtA and *Bacillus anthracis* SrtA (baSrtA) using cleverly designed peptidomimetic ligands (PDB IDs: 2KID and 2RUI). However, because it is not present, these structures do not provide information about recognition of the P1′ residue, a position for which SrtA enzymes have shown variable selectivity *in vitro* (9, 13–15). The third structure contains a complex between saSrtA and a noncovalently bound LPETG peptide that is shifted by several Angstroms in the peptide-binding pocket (PDB ID: 1T2W), revealing a geometry that is not consistent with known biochemical data (16).

In this work, we have sought to fill remaining gaps in the understanding of SrtA target recognition through the structural characterization of multiple states in the catalytic mechanism of *Streptococcus pyogenes* class A sortase (spySrtA) (Fig. 1). The apo structure of spySrtA was solved using X-ray crystallography in 2009, and its catalytic triad consists of His142, Cys208, and Arg216 (17). Using similar crystallization conditions, we were able to crystallize and solve the structures of a catalytically inactive C208A spySrtA mutant bound to the peptides LPATA and LPATS, which are sequences that are known to serve as spySrtA substrates *in vitro* (17–23). In addition, we synthesized a model peptide (LPAT-LII) of the ligation product between the LPAT fragment and the *in vivo*

nucleophile LII and solved the structures of two complexes between C208A spySrtA and LPAT-LII where the peptide is in the "Thr-in" and "Thr-out" conformations, terminology previously used to describe the side chain of the P1 Thr as protein interacting ("Thr-in") or solvent interacting ("Thr-out") (8).

Because these are the first solved peptide-bound sortase structures that include the P1′ residue and initial cleavage site, we wanted to investigate the relative dynamics of ligand binding. We ran 900 ns molecular dynamics (MD) simulations using four structures (apo spySrtA [PDB ID: 3FN5], spySrtA–LPATA, spySrtA–LPATS, and spySrtA–LPAT-LII) to assess positional flexibility and the overall dynamics of the sortase–peptide complex. Finally, we used our peptide-bound structures to model the acyl–enzyme intermediate of spySrtA–LPAT. Taken together, this work provides new structural insights for important states in the SrtA catalytic mechanism (Fig. 1), significantly increasing our understanding of target recognition in this important protein family.

Results

Peptide-bound spySrtA crystallization and structure determination

Like other class A sortases, the majority of predicted and verified *in vivo* targets of spySrtA possess LPXTG substrate sequences (24, 25). In addition, prior work from ourselves and others has demonstrated that spySrtA readily accepts LPXTA and LPXTS substrates *in vitro*, despite the fact that these
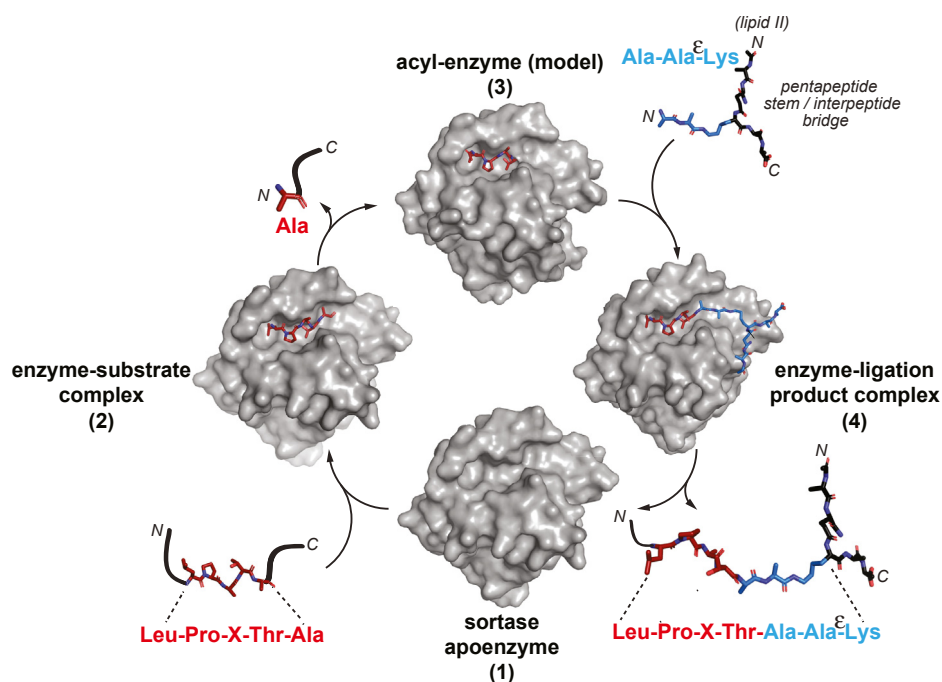

Figure 1. Structural model of the spySrtA catalytic mechanism. A summary of the spySrtA catalytic mechanism, as supported by biochemical and structural data in the field, including studies presented here. A portion of the lipid II ligation partner from *Streptococcus pyogenes* is shown, and the structure of this component varies between bacterial species. In this model, LPXTA is shown as the target sequence in the first step of the reaction to reflect the ability of spySrtA to recognize a P1′ Ala residue *in vitro* and to be consistent with the structural data described in this work. The sortase apoenzyme in state 1 is PDB ID: 3FN5. The structures of the enzyme–substrate complex (state 2) and enzyme–ligation product complex (state 4) are the experimental structures presented in this study. As discussed in the main text, the acyl–enzyme (state 3) intermediate is a model generated from experimentally determined structures of the enzyme–substrate complex. PDB, Protein Data Bank; spySrtA, *Streptococcus pyogenes* class A sortase.

particular sequence variants do not appear to be present in naturally occurring spySrtA substrates *in vivo* (20, 23, 26). The spySrtA enzyme also accepts alanine- or serine-based nucleophiles, which is a characteristic that has been exploited for dual-labeling SML strategies and is consistent with the presence of N-terminal alanines in the interpeptide bridge of LII in *S. pyogenes* (17–22, 27). Notably, the ability of spySrtA to recognize non-glycine nucleophiles and to accept substrates that vary at the P1′ position is in stark contrast to saSrtA, which is narrowly selective for glycine at these sites (15, 27, 28).

In order to gain a stereochemical understanding of target recognition by spySrtA and other class A sortases, we sought to cocrystallize a catalytically inactive mutant (C208A) of spySrtA with a range of model peptides containing known

substrate sequences (LPATG/S/A). Briefly, spySrtA protein containing the inactivating C208A mutation was expressed and purified as previously described for the wildtype protein and analyzed *via* SDS-PAGE and LC–electrospray ionization (ESI)–mass spectrometry (MS) (Fig. S1) (23). Purified protein (at ∼1.1 mM) was incubated in a 1:1 ratio with 1 mM peptide for 1 h prior to crystallization by hanging drop vapor diffusion method. Crystallization conditions were optimized from those used for apo spySrtA (PDB ID: 3FN5) and are described in the Experimental procedures section (17). From this, we succeeded in crystallizing and solving two structures of C208A spySrtA bound to the model peptides (P1′ position in **bold**) *Abz*-LPAT**A**GK(Dnp)-*NH₂* and *Ac*-LPAT**S**G-*NH₂* (Abz, 2-aminobenzoyl; Fig. 2A). The former is an example of a FRET
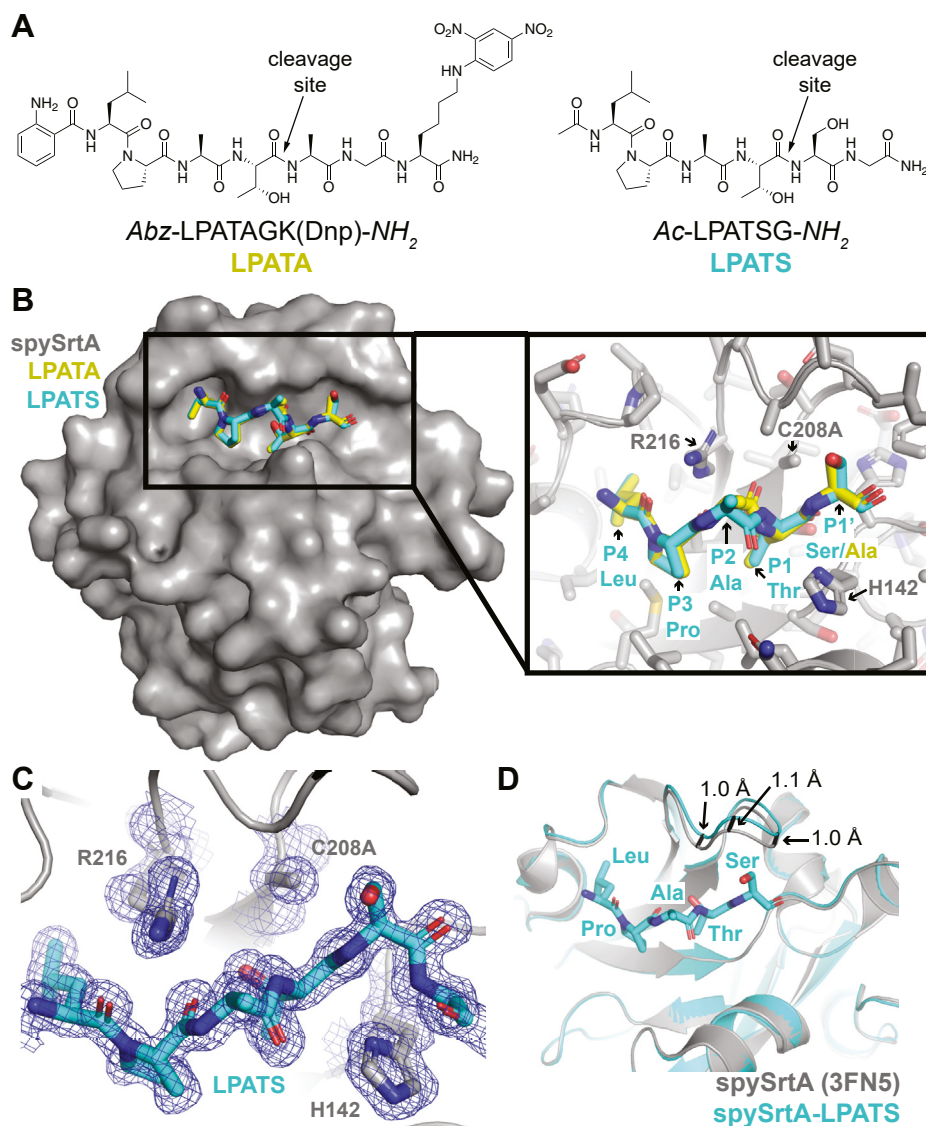

Figure 2. The spySrtA complex structures with LPATS and LPATA peptides. *A*, structures of model peptides cocrystallized with spySrtA. *B*, the spySrtA protein is in *gray* surface representation, and the noncovalently bound LPATS and LPATA peptides are in *sticks* and colored by heteroatom (N = *blue*, O = *red*, and C = *yellow/cyan* as labeled). The N-terminal Abz moiety on the LPATA peptide and the P2′ Gly residue for both peptides are not shown in order to focus on the target recognition sequences. The *inset box* shows a zoomed-in version of peptide binding, with all peptide positions labeled. The spySrtA side chains are shown as *sticks* and colored by heteroatom. The catalytic triad (H142–C208A–R216) are labeled. *C*, The $2F_o–F_c$ electron density map for the LPATS peptide and catalytic triad is shown in *blue mesh* and rendered at 1.0 σ. The structure is shown as in the *inset* of (*B*). *D*, alignment of the the apo spySrtA (PDB ID: 3FN5) and spySrtA–LPATS structures reveals an RMSD = 0.158 Å (508 main-chain atoms). The proteins are shown in *cartoon* representation and colored as labeled. The LPATS peptide is shown as *sticks* and colored by heteroatom. Abz, 2-aminobenzoyl; PDB, Protein Data Bank; spySrtA, *Streptococcus pyogenes* class A sortase.

quencher probe that is commonly used for monitoring sortase enzymatic activity (9, 15, 23, 29–31), whereas the latter is a simplified target containing an acetyl (*Ac-*) cap and C-terminal primary amide (*-NH₂*). For both substrates, LC–ESI–MS was used to confirm that they were cleaved by wildtype spySrtA in a model transacylation reaction (Fig. S2). Notably, we also crystallized C208A spySrtA with peptides containing the canonical LPXTG sequence (*Abz*-LPATGGK(Dnp)-*NH₂* and fluorescently labeled *5-FAM-Ahx*-LPATGG-*NH₂*); however, the crystals obtained were not of suitable diffraction quality.

For simplicity, we will hereafter refer to the solved enzyme–substrate complexes as spySrtA–LPATA and spySrtA–LPATS (Fig. 2*B*). All diffraction and refinement statistics for these complexes are provided in Table 1. In general, crystals grew stacked and were relatively unstable in traditional cryosolutions (*e.g.*, with 10–20% [w/v] glycerol added). As a likely result of these challenges, the crystal ultimately used for spySrtA–LPATA structure determination contained pseudosymmetry. We predict that this may be due to lattice disruption during crystal harvesting. The space group of this crystal was $P\,2_1\,2_1\,2_1$ and contained two protomers in the asymmetric unit. We refined it to a $R_{\mathrm{work}}/R_{\mathrm{free}} = 0.21/0.24$ at 1.4 Å resolution (Table 1). Relatively high *R*-factors are a consequence of pseudosymmetry in crystal packing (32). Optimization of cryo conditions, namely using PEG 400 as a cryoprotectant, resulted in better quality diffraction data for the crystal used to solve the spySrtA–LPATS structure, as described in the Experimental procedures section. This crystal diffracted to 1.4 Å resolution, and the resulting structure was solved in space group $P\,2_1$ to a $R_{\mathrm{work}}/R_{\mathrm{free}} = 0.17/0.19$, with two spySrtA molecules in the asymmetric unit (Table 1). The unit cell and space group are very similar between spySrtA–LPATS and apo spySrtA (17).

Alignment of chain (or protomer) A of spySrtA–LPATS with the two molecules of spySrtA–LPATA revealed very similar structures, with pairwise RMSD values for main-chain atoms of all protomers of both structures <0.13 Å (Fig. S3*A*). We were able to model all residues of the enzyme in protomer B of spySrtA–LPATA and protomer A of spySrtA–LPATS revealing an additional N-terminal helix not previously seen in the apo structure (Fig. S3*B*). Because of the large degree of similarity between these structures, unless otherwise noted, our analyses will focus on spySrtA–LPATS protomer A.

Stereochemistry of target recognition by spySrtA

We next used our peptide-bound crystal structures to analyze the stereochemistry of target recognition by class A sortases. In both structures, we see clear peptide density and modeled the entire pentapeptide motif for all spySrtA protomers (Fig. 2, *B* and *C*). Unbiased electron density maps, created by omitting the peptide atoms and running a round of refinement, confirm strong electron density for peptide residues (Fig. S3*C*). Alignment of spySrtA–LPATS with the two (A and B) protomers of apo spySrtA revealed RMSD values for main-chain atoms of 0.158 Å (508 atoms) and 0.189 Å (541 atoms), respectively. The largest difference between these structures is an approximately 1 Å displacement in the backbone of the β7–β8[+3], β7–β8[+4], and β7–β8[+5] loop residues (Fig. 2*D*). Here, superscript numbering refers to the residue position with respect to the catalytic C208 residue, as previously defined (9). This suggests that very small structural rearrangements are needed in order to accommodate the target peptide.

We were able to model the Abz moieties in the spySrtA–LPATA protomers, although the 2,4-dinitrophenyl lysine

Table 1
Data collection and refinement statistics

Data collection	spySrtA–LPATA	spySrtA–LPATS	spySrtA-LII "Thr-in"	spySrtA-LII "Thr-out"
Space group	$P\,2_1\,2_1\,2_1$ (19)	$P\,2_1$ (4)	$P\,2_1\,2_1\,2_1$ (19)	$P\,2_1\,2_1\,2_1$ (19)
Unit cell dimensions				
a, b, c (Å)	58.98, 64.57, 75.02	38.49, 59.1, 64.57	34.3, 57.73, 72.25	34.32, 57.68, 71.46
α, β, γ (°)	90, 90, 90	90, 101.7, 90	90, 90, 90	90, 90, 90
Resolution (Å)[a]	48.9–1.4 (1.5–1.4)	43.1–1.4 (1.5–1.4)	45.1–1.8 (1.91–1.8)	44.9–1.9 (2.02–1.9)
R_{sym} (%)[b]	8.1 (40.8)	6.1 (31.5)	7.4 (53.0)	8.3 (54.7)
I/σ_I[c]	12.60 (4.34)	16.86 (4.56)	15.55 (2.47)	16.29 (3.49)
Completeness (%)	99.9 (99.6)	99.6 (98.5)	98.9 (94.0)	98.7 (97.4)
Refinement				
Total no. of reflections	56,989	55,864	13,739	11,555
Reflections in the test set	2838	2709	684	534
R_{work}[d]/R_{free}[e]	21.3/23.8	16.9/19.4	17.5/20.6	18.4/22.9
Number of atoms				
Protein	2684	2683	1303	1311
Water	257	449	129	71
Ramachandran plot (%)[f]	99.12/0.88/0	99.42/0.58/0	100/0/0	98.73/1.27/0
B_{av} (Å²)				
Protein	14.4	14.6	24.3	27.4
Bond length RMSD (Å)	0.007	0.006	0.006	0.007
Bond angle RMSD (°)	0.988	0.872	1.135	1.338
PDB code	7S51	7S40	7T8Y	7T8Z

[a] Values in parentheses are for data in the highest-resolution shell.
[b] $R_{\mathrm{sym}} = \Sigma_h \Sigma_i |\mathrm{I}(h) - \mathrm{I}_i(h)|/\Sigma_h \Sigma_i \mathrm{I}_i(h)$, where $\mathrm{I}_i(h)$ and $\mathrm{I}(h)$ values are the *i*-th and mean measurements of the intensity of reflection *h*.
[c] SigAno = |F(+) - F(−)|/σ.
[d] $R_{\mathrm{work}} = \Sigma ||\mathrm{F}_{\mathrm{obs}}|_h - |\mathrm{F}_{\mathrm{calc}}||_h/\Sigma |\mathrm{F}_{\mathrm{obs}}|_h$, *h* ε {working set}.
[e] R_{free} is calculated as R_{work} for the reflections *h* ε {test set}.
[f] Favored/allowed/outliers.

residue (K(Dnp)) was unresolved. In the A-protomer of spySrtA–LPATA, we see a potential hydrogen bond between the 2-amino group of Abz and the carbonyl of P188 (Fig. S4A). While interesting, we do not consider this interaction to be critical for the binding of this substrate, as it is not observed in the B-protomer of the spySrtA–LPATA complex. This is further supported by the successful binding and cocrystallization of the *Ac*-LPATSG-*NH₂* peptide, which lacks the Abz unit.

We next analyzed position-specific interactions in the LPATX motif of the CWSS. The highly conserved Leu residue at P4 interacts with a hydrophobic pocket formed by V186, V191, and V193 of the β6–β7 loop, as well as V206 in β7 and I218 in β8 (Fig. 3A). A similar pocket was previously identified in the NMR structure of saSrtA with a covalent peptidomimetic (LPAT*), PDB ID: 2KID (14). The proline residue in P3 interacts weakly *via* van der Waals interactions with V206 and A140, residues in the β4 and β7 strands, as well as M125 in the β3–β4 loop (Fig. 3A). The distances between these residues are of equal magnitude or shorter to those seen in the saSrtA–LPAT* structure, where strong intermolecular NOEs were observed that supported P3 Pro interactions with residues in the β4 and β7 strands (Fig. S4B) (14).

There are several backbone atoms in the LPATX motif that form noncovalent interactions with residues in spySrtA (Fig. 3B). In both the LPATA and LPATS structures, the carbonyl oxygens of the P4 Leu and P3 Pro residues are hydrogen bonded with nitrogen atoms in R216, the catalytic arginine residue. In the LPATA complex, R216 also interacts with the P2 Ala carbonyl, whereas in the LPATS structure, this carbonyl is rotated ~180° and interacting with solvent (Fig. 3B). In all structures, the orientation of the P2 and P1 residue side chains (AT, respectively) observed is rotated ~180° as compared with the saSrtA–LPAT* structure, agreeing more closely with the structure of baSrtA–LPAT* from the same group (Fig. S4C) (13, 14). As described previously, the conformation observed in spySrtA–LPATA and spySrtA–LPATS is referred to as "Thr-in" to describe the P1 Thr side chain oriented toward the enzyme (8). The carbonyl oxygen of P1 Thr further interacts with the amide of C208A and side chain hydroxyl of T207 as well as the amide of H143, the residue immediately C terminal to the catalytic histidine, H142 (Fig. 3B). The methyl group of the P1 Thr is oriented toward the side chain atoms of A140 and V206, and the side chain hydroxyl interacts with the amide of the catalytic C208A
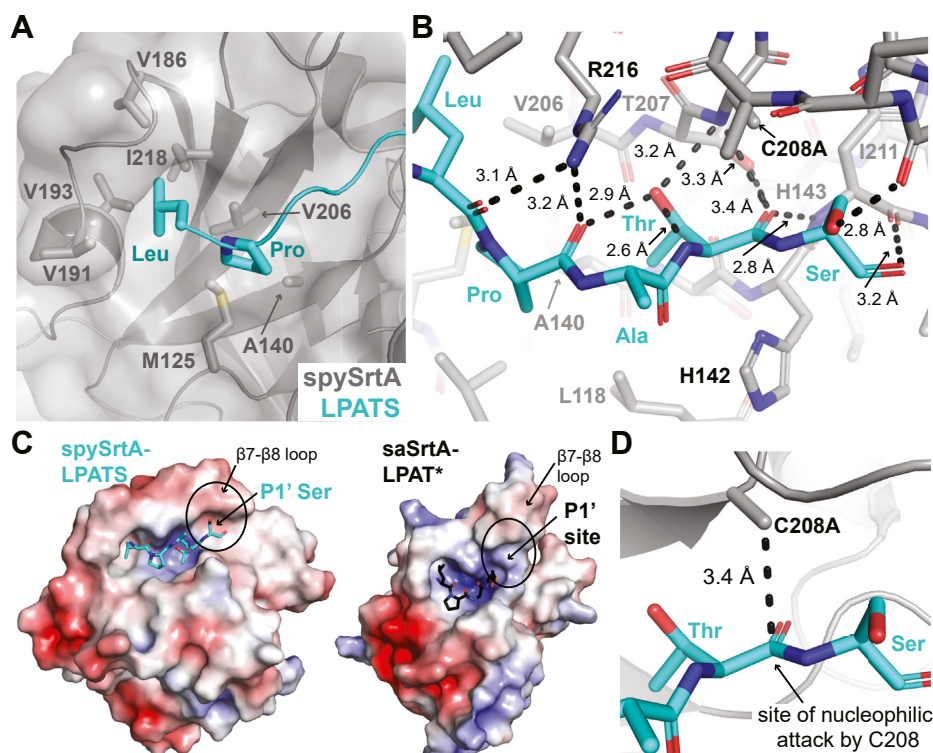

Figure 3. Stereochemistry of the spySrtA–LPATS interaction. *A*, the interactions of the P4 Leu and P3 Pro ligand residues with spySrtA are highlighted. The spySrtA enzyme is in cartoon and surface representation, with residues that interact directly with the Leu-Pro shown as sticks and colored by heteroatom (C = *gray*, S = *golden yellow*). The ligand is shown as a *cyan cartoon*, with the side chain sticks of Leu-Pro shown and colored by heteroatom (C = *cyan*, N = *blue*). *B*, there are several noncovalent interactions between the LPATS ligand and spySrtA enzyme, shown as *black dashed lines* with distances labeled. There are also intramolecular interactions between the P1 Thr sidechain and its own amide, as well as the P3 Pro carbonyl oxygen, as labeled. The ligand is in stick representation and colored by heteroatom (C = *cyan*, N = *blue*, O = *red*). The spySrtA enzyme is shown as sticks and colored by heteroatom. The catalytic triad (H142-C208A-R216) is labeled. *C*, the electrostatic potential surface maps for SrtA in spySrtA-LPATS (left) and saSrtA-LPAT* (right, PDB ID 2KID) were created using APBS in PyMOL and are shown from ± 5 eV, with *red* = negative and *blue* = positive. The ligands are shown as sticks and colored by heteroatom (as in *A-B*). The location (spySrtA, left) or predicted location (saSrtA, right) of the P1′ site is circled, and the β7-β8 loop is labeled. *D*, the distance between the CB atom of C208A and the C of the P1 Thr is shown as *black dashed lines* and labeled. This is the site of nucleophilic attack by C208. The structures are rendered as in *B*, with the exception that spySrtA is in cartoon representation with only the side chain sticks of C208A shown.

residue, as well as forms intrapeptide hydrogen bonds with its own amide and the carbonyl of the P3 Pro (Fig. 3*B*).

Finally, the P1′ Ser in spySrtA–LPATS interacts with a weakly negative ridge formed by the β7–β8 loop, specifically because of E212, the β7–β8^{+4} residue (Fig. 3*C*). A spatially analogous P1′ binding site, albeit with some differences in morphology and overall charge, was predicted in the previously reported saSrtA–LPAT* structure (PDB ID: 2KID) (Fig. 3*C*). In our spySrtA–LPATS structure, we also observe a hydrogen bond with the hydroxyl group of the P1′ Ser and the carbonyl of I211 (Fig. 3*B*). This interaction is necessarily absent from the spySrtA–LPATA complex, and therefore, we do not consider it a requirement for substrate binding. In general, the binding site for the P1′ position in spySrtA does not appear to be particularly selective, which is consistent with our previous work on *S. pneumoniae* SrtA (9, 15). Because of the observed similarities in these *Streptococcus* SrtA proteins, as well as our previous work investigating the β7–β8 loop in these proteins, we hypothesize that spySrtA is also nonselective at this position and can accommodate a wide variety of P1′ amino acids (9, 23).

Overall, the observed location for the P1′ Ser, as well as the adjacent P1 Thr, renders the LPATS peptide ideally positioned for nucleophilic attack by the catalytic cysteine residue. Specifically, the methyl group of C208A in the spySrtA–LPATS structure is 3.4 Å from the P1 Thr carbonyl carbon (the corresponding distance in spySrtA–LPATA is 3.2 Å) (Fig. 3*D*). The scissile amide bond of the P1–P1′ linkage is also held in close proximity to the catalytic histidine (His142), which is consistent with the suggested role of this residue in facilitating proton transfers to the excised P1′ fragment and from the incoming LII nucleophile (8). Taken together, these observations support the validity of the spySrtA–LPATS and spySrtA–LPATA complexes as reasonable models for target recognition by class A sortases that are consistent with the current understanding of the sortase catalytic mechanism (4, 8, 33, 34).

Model of the acyl–enzyme intermediate

Next, we used our spySrtA–LPATS complex structure to model the acyl–enzyme intermediate (Fig. 4*A*). The model was constructed as described in the Experimental procedures section. Briefly, coordinates for the cleaved peptide were determined and fit into the experimental electron density for spySrtA–LPATS. In addition, C208A was mutated *in silico* to the wildtype cysteine, and a round of refinement was run to validate the peptide geometry. We then performed a steepest descent energy minimization of the acyl–enzyme model to obtain the final geometry (Fig. S8*D*). The resulting acyl–enzyme model is therefore very similar to the spySrtA–LPATS structure, including nearly identical positions for the P4–P2 residues of the LPATS substrate (Fig. 4*B*). Slight differences were observed, however, in the case of the P1 Thr residue. As discussed previously, the P1 Thr carbonyl in spySrtA–LPATS appears to be stabilized by the amides of H143 and C208A as well as the side chain hydroxyl of T207. These interactions were largely maintained in our model;
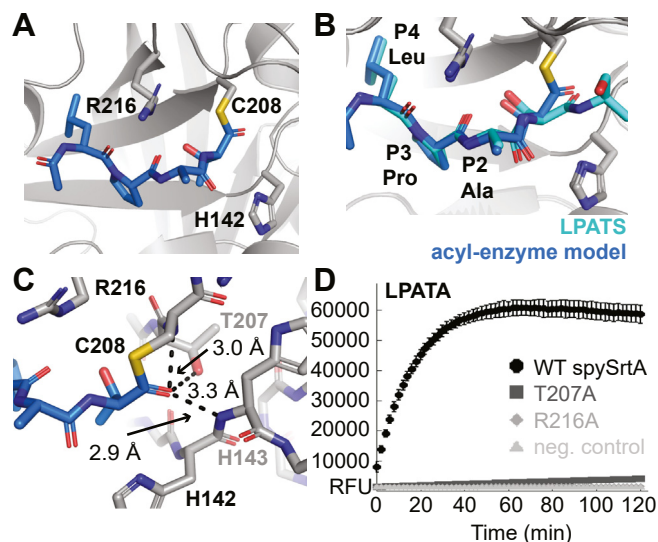
Figure 4. A model of the acyl–enzyme intermediate of spySrtA–LPAT. *A–C,* the spySrtA protein for the acyl–enzyme intermediate model is in *gray cartoon* (*A* and *B*) or *stick* representation (*C*), with catalytic residue side chains (H142, C208, and R216) and other relevant positions shown as *sticks*, colored by heteroatom (C = *blue/cyan*, O = *red*, N = *blue*, and S = *golden yellow*), and labeled. The spySrtA protein from the spySrtA–LPATS structure (*B*) is rendered similarly but is *darker gray*. The peptides are in *stick* representation, with carbons colored as labeled. *C,* predicted interactions between the P1 Thr carbonyl and spySrtA T207 side chain, C208A amide, and H143 amide are shown as *dashed black lines*, with distances labeled. *D,* triplicate fluorescence data (in relative fluorescence units, RFU) for the reaction of *Abz*-LPATAGK(Dnp)-*NH₂* and H₂NOH in the presence of WT (*black circles*), T207A (*dark gray squares*), and R216A (*gray diamonds*) spySrtA protein, and a no protein control (*light gray triangles*). Data for the *Abz*-LPATGGK(Dnp)-*NH₂* and *Abz*-LPATSGK(Dnp)-*NH₂* peptides are in Fig. S4*E*. Abz, 2-aminobenzoyl; spySrtA, *Streptococcus pyogenes* class A sortase.

however, a slight rotation of the P1 carbonyl toward T207 was observed (Fig. 4*B*). Specifically, in the geometry of the acyl–enzyme model, the T207 hydroxyl is 3.3 Å from the P1 Thr carbonyl (Fig. 4*C*). This distance is 3.4 Å in the solved structure of spySrtA–LPATS (Fig. 3*B*).

With respect to catalytic mechanism, a feature of the acyl–enzyme model that was also shared by both the spySrtA–LPATS and spySrtA–LPATA structures was the absence of a clear interaction between the P1 Thr carbonyl group and the putative catalytic arginine (R216) side chain. This is significant as this arginine has been proposed to stabilize high-energy oxyanion intermediates generated during the sortase ligation reaction (8, 14, 35). The P1 Thr carbonyl in our acyl–enzyme model and solved structures was actually observed to point away from the R216 side chain, and the distance between these sites is >6 Å (Fig. 4*C*). Nonetheless, R216 was found to be essential for spySrtA function, as mutating it to an Ala residue resulted in complete loss of enzyme activity when tested with model LPATG/S/A peptide substrates (Figs. 4*D* and S4*E*).

In terms of oxyanion stabilization, our structures are more consistent with a key role for the side-chain hydroxyl of T207. This residue, along with the amides of H143 and C208, is ideally positioned to bind to the P1 Thr carbonyl and potentially stabilize tetrahedral oxyanion intermediates formed immediately prior to the acyl enzyme state and following nucleophilic attack by LII (Fig. 4*C*). This type of role for the

Thr immediately preceding the catalytic Cys has indeed been suggested in previous computational studies (36). Moreover, sequence analysis of 400 sortase A enzymes in the National Center for Biotechnology Information database reveals that over 90% (363 total) contain a Thr residue immediately preceding the catalytic Cys, which suggests a fundamentally important role for this Thr such as stabilization of key reaction intermediates. Consistent with this hypothesis, we found that a T207A mutant of spySrtA exhibited a near total loss of enzymatic activity (Figs. 4D and S4E). Notably, a dramatic drop in enzyme activity has also been reported when mutating the corresponding Thr residue (T183) of saSrtA (37).

Structure and biochemical analyses of spySrtA bound to an LII mimetic

Building from our peptide-bound structures, we next explored the nature of the interaction between spySrtA and its *in vivo* nucleophile, LII. The LII molecule has been identified as the anchor for sortase-catalyzed attachment of many proteins to the bacterial cell wall and serves as a key precursor for the production of peptidoglycan. The nature of this peptidoglycan layer and the cell exterior as a whole is what differentiates Gram-positive and Gram-negative bacteria. Whereas Gram-negative bacteria have an inner membrane surrounded by a relatively thin peptidoglycan layer, followed by a second lipoprotein outer membrane, Gram-positive bacteria lack the outer membrane and contain a relatively thick peptidoglycan layer (38). Although there are exceptions and possible modifications, the main glycan moiety of the peptidoglycan layer consists of alternating GlcNAc and *N*-acetylmuramic acid (MurNAc) residues that are further crosslinked by peptide subunits (38, 39).

The LII building block itself consists of the GlcNAc–MurNAc disaccharide attached to a polyisoprenoid membrane anchor and a pentapeptide stem that is linked *via* an amide bond to the C-3 D-lactyl ether of MurNAc (39). While the structure of the pentapeptide stem varies, a common sequence in Gram-positive bacteria such as *S. pyogenes* is L-alanine, D-isoglutamine, L-lysine, D-alanine, and D-alanine (38–40). In many of these organisms, the L-lysine is subsequently modified by peptidyltransferases to create an *interpeptide bridge*, which are the residues that ultimately serve as the nucleophile for SML of surface proteins to the peptidoglycan layer. The nature of this interpeptide bridge is variable but commonly includes L-Gly/Ala/Ser residues, for example, for *S. aureus* = Gly$_5$, *Enterococcus faecalis* = Ala–Ala, and *Streptococci* = Ala/Ser–Ala (38, 39).

To visualize the interaction of spySrtA with LII and its related ligation products, we synthesized a model branched peptide representing the ligation of an LPATX substrate to the interpeptide bridge/pentapeptide stem portion of LII from *S. pyogenes* (Fig. 5A). Synthesis and characterization are described in the Experimental procedures and Supporting information sections. Specifically, this structure (LPAT-LII) possesses an *Abz*-LPAT fragment derived from the *Abz*-LPATAGK(Dnp)-*NH$_2$* substrate described previously

covalently linked to an LII mimetic *via* a dialanine interpeptide bridge. To our knowledge, there is no evidence of specific interactions between the glycan residues of LII and the sortase enzyme; therefore, those portions were omitted and replaced with a simple acetyl group on the terminal L-alanine residue. We also note that some structural heterogeneity in the interpeptide bridge/pentapeptide stem of *S. pyogenes* is likely. Examples of this include variable numbers of alanine residues in the interpeptide bridge and even low levels of hydroxylysine (38, 41). However, our LPAT-LII model is consistent with structural features reported in the literature and should therefore be representative of a significant fraction of LII structures in *S. pyogenes* (17, 38, 42).

As a preliminary assessment of whether our LPAT-LII model was recognized by the enzyme, it was used in a model of spySrtA-catalyzed reaction and found to be efficiently cleaved at the expected site between the Thr and Ala residues (Figs. 5B and S5). Indeed, we found LPAT-LII to react more rapidly than the related *Abz*-LPATAGK(Dnp)-*NH$_2$* peptide, suggesting that the added interpeptide bridge/pentapeptide stem portion may be enhancing binding and recognition by spySrtA (Fig. S5).

We next crystallized and solved the structure of C208A spySrtA noncovalently bound to our LPAT-LII mimetic. Two distinct conformations were observed, with the peptide Thr residue in both the "Thr-in" and "Thr-out" conformations previously observed in other SrtA structures (Fig. 5, C and D) (13, 14). These structures will be referred to as spySrtA–LPAT-LII "Thr-in" and spySrtA–LPAT-LII "Thr-out." Crystallization was performed similarly to the peptide-bound structures described previously, and as in the Experimental procedures section. Microseeding was used in this case to obtain crystals of suitable diffraction quality. Both the "Thr-in" and "Thr-out" structures crystallized in the space group $P\,2_1\,2_1\,2_1$ with one molecule in the asymmetric unit and to a resolution of 1.8 Å and 1.9 Å, respectively. The spySrtA–LPAT-LII "Thr-in" structure was refined to a final $R_{work}/R_{free} = 0.18/0.21$ and the spySrtA–LPAT-LII "Thr-out" structure to a final $R_{work}/R_{free} = 0.18/0.23$ (Table 1). Overall, the structures are very similar, and the main chain atoms of spySrtA align with an RMSD = 0.082 Å (559 atoms).

In the spySrtA–LPAT-LII "Thr-in" structure, the stereochemistry of the LPATA portion is consistent with our peptide-bound structures (Fig. S6A). The main-chain atoms align to the A- and B-protomer of spySrtA–LPATA with an RMSD = 0.218 Å (518 atoms) and 0.205 Å (495 atoms), respectively. Values are almost identical for spySrtA–LPAT-LII "Thr-out," at 0.218 Å (497) and 0.207 Å (489) for the spySrtA–LPATA A- and B-protomers. The positions of the interpeptide bridge dialanine and ε-amine/ε-carbon of the L-lysine residue are also well conserved between the "Thr-in" and "Thr-out" structures (*gray arrow* in Fig. 5D). These sites make contacts with residues of the β7–β8 loop and appear to be stabilized by a hydrophobic pocket in spySrtA formed by four amino acids (I119 in α1, I144 and I147 in the β4–α2 loop, and V247 at the C terminus) (Fig. 5E). Moving beyond the ε-carbon of L-lysine, there is more variability in the
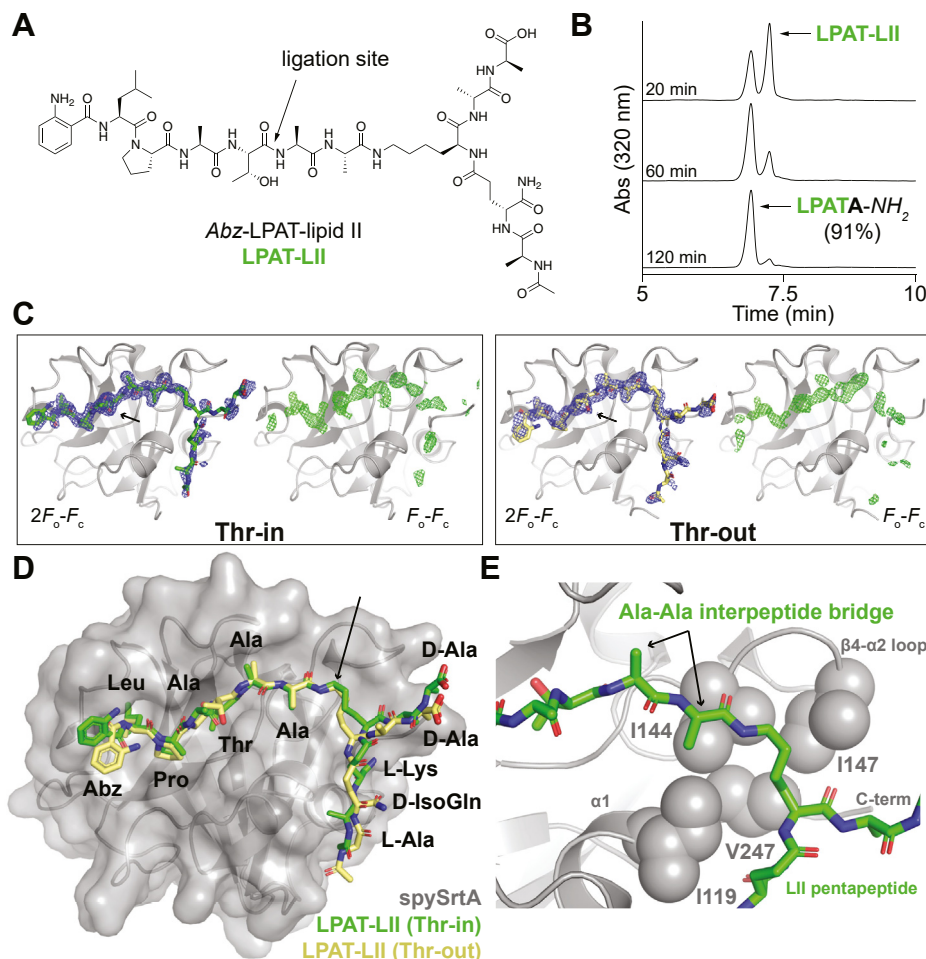
Figure 5. The structure of spySrtA bound to a peptide model of the LPAT-lipid II ligation product. *A*, chemical structure of the *Abz*-LPAT-lipid II (LPAT-LII) model with the ligation site between the LPAT fragment and lipid II highlighted. *B*, reverse phase-HPLC chromatograms showing efficient cleavage (91% conversion) of the LPAT-LII model in the presence of spySrtA and an excess of alanine amide (A-NH₂) nucleophile. Percent conversion was estimated by comparing reverse phase-HPLC peak areas for the unreacted LPAT-II substrate and the LPATA-NH₂ product. *C*, the two $F_o–F_c$ electron density maps for the "Thr-in" and "Thr-out" LPAT-LII peptides are shown at 0.6 σ (*blue mesh*, *left figures*), highlighting the specific, but relatively weak, density in the pentapeptide stem region. The LPAT-LII ligand is in *sticks* and colored as in (*D*). Unbiased $F_o–F_c$ maps at 2.0 σ are also shown (*green mesh*, *right figures*), which were created by deleting the peptide density and running a round of refinement. The *black arrows* indicate the Thr side chain in both structures. *D*, the structure of spySrtA and the LPAT-LII molecule. This complex was crystallized in both P1 "Thr-in" and "Thr-out" ligand conformations, and both peptides are shown here. SpySrtA is very similar for both structures, and alignment reveals RMSD = 0.082 Å (559 main-chain atoms). Therefore, only the spySrtA protein for the "Thr-in" structure is shown (*gray cartoon* and in surface representation). The LPAT-LII residues are colored as labeled and by heteroatom (N = *blue* and O = *red*). The *black arrow* highlights the Cε atom of the Lys residue of LII, which is the atom at which the conformations of the two solved structures begin to vary, indicating flexibility in the LII pentapeptide. *E*, interactions at the interpeptide bridge of the LPAT-LII ligand are highlighted. Side-chain atoms in spySrtA that form a hydrophobic interaction surface in the vicinity are shown as *spheres*, and the residues are labeled. The "Thr-In" peptide is rendered as in (*D*), and the *arrows* point to the Ala–Ala residues of the interpeptide bridge, as labeled. Abz, 2-aminobenzoyl; spySrtA, *Streptococcus pyogenes* class A sortase.

conformation of the pentapeptide stem between the two structures; this reflects the weaker electron density for these residues (Fig. 5, *C* and *D*). Indeed, in both structures, there is only one observed noncovalent interaction with the LII pentapeptide and spySrtA enzyme, a hydrogen bond formed between the spySrtA α1 Y120 hydroxyl and the amide of the LII D-isoglutamine residue (Fig. S6B). In each, there are also multiple interactions with the LII pentapeptide and spySrtA enzyme of molecules related by symmetry (Fig. S6C).

Taken together, our crystallographic findings suggest that while the interpeptide bridge likely plays an important role in SrtA recognition of LII, the pentapeptide stem does not substantially interact with the enzyme. As noted previously, the electron density for the pentapeptide stem was weaker than that of the LPAT segment and interpeptide bridge dialanine,

suggesting flexibility in the stem region of the LPAT-LII ligand (Fig. 5C). Nonetheless, the clear electron density for the dialanine interpeptide bridge revealed a discrete binding site with potential implications for substrate binding outside the standard LPXTG substrate motif, specifically at the P2′ position. Interestingly, several predicted *in vivo* substrates of *S. pyogenes* and other streptococcal species possess LPXTGE motifs, with glutamic acid occupying this P2′ position (25). In our hands, preliminary experiments suggest that spySrtA recognizes LPATGG and LPATGE peptides similarly, but additional work is ongoing to investigate P2′ specificity (data not shown).

MD simulations of spySrtA bound to target peptides

During structure refinement and model building for spySrtA–LPAT-LII, we observed reduced electron density for

the pentapeptide stem as compared to the LPAT sequence and interpeptide bridge, which suggested variations in conformational dynamics for different segments of the LPAT-LII ligand (Fig. 5C). To probe this further, as well as investigate the MD of our other spySrtA substrate complexes, we ran ∼900 ns MD simulations of apo spySrtA (PDB ID: 3FN5), spySrtA–LPATA, spySrtA–LPATS, and spySrtA-LII "Thr-in" structures (Table S1). Briefly, MD simulations were performed in full atomistic detail with explicit water using the AMBER99SB*-ILDN force fields (43). The starting structures were solvated with ∼10,000 TIP3P water molecules in a cubic box with periodic boundary conditions. The system was neutralized with an ionic concentration of 150 mM. These simulations are

described in further detail in the Experimental procedures and Supporting information sections.

Overall, the RMSD of backbone atom positions for spySrtA indicated that the enzyme remained stable over the course of all four simulations. (Fig. S7A). The root mean square fluctuation (RMSF) of backbone atoms for spySrtA were also consistent with a well-defined eight-stranded antiparallel β-barrel sortase core structure, in that these regions are relatively inflexible over the course of the simulation, as compared with some α-helical and all loop regions (Fig. 6A). In all the peptide ligands, the LPATX sequences were also relatively inflexible. This was clearly evident in the visualization of representative frames taken over the course of each simulation (Fig. 6B) as
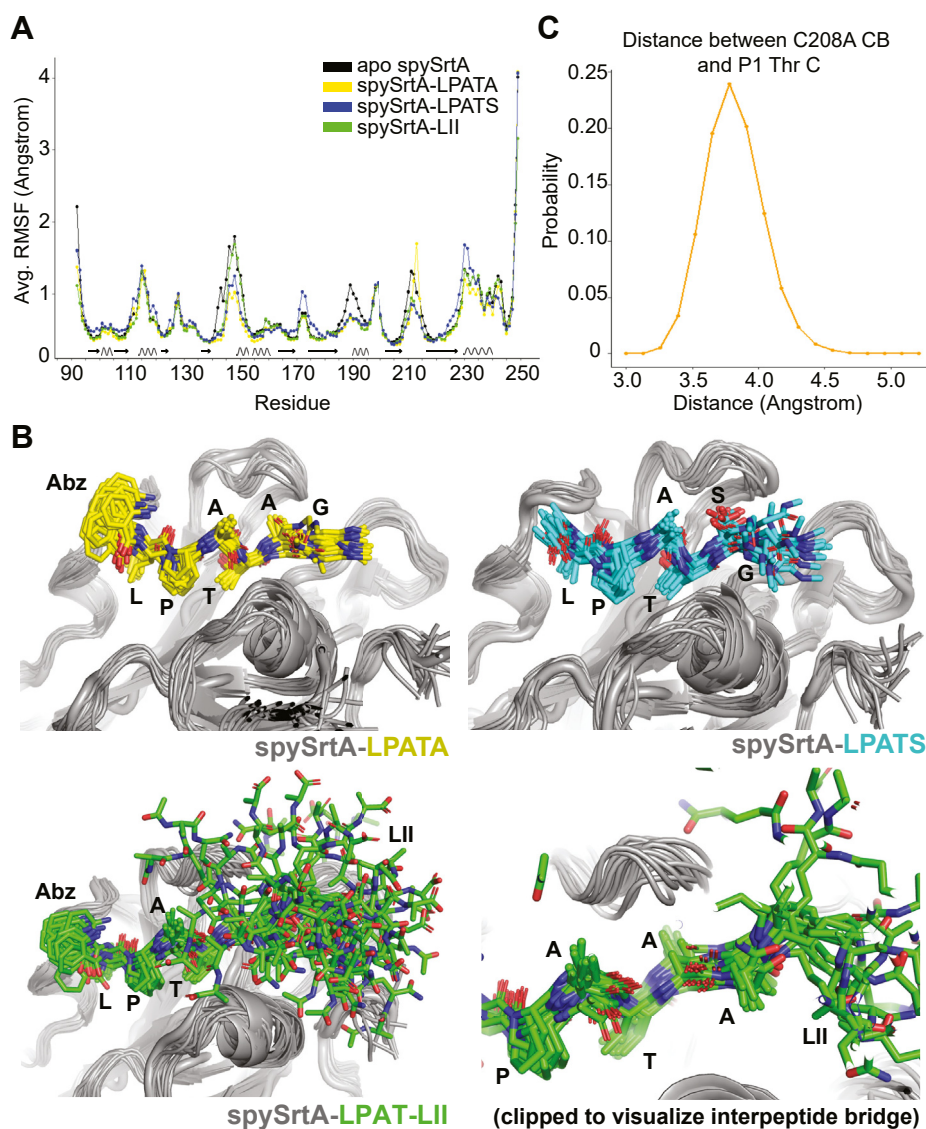

Figure 6. Molecular dynamics simulations of spySrtA structures. *A*, the average root mean square fluctuation (RMSF) of backbone atoms in each residue are shown and colored as labeled. The secondary structure elements are indicated under the curve by *arrows* for β-strands and *curved lines* for α-helices. These are based on the apo spySrtA structure (PDB ID: 3FN5), and residue boundaries fluctuate slightly amongst the structures. *B*, representative frames from every approximately 45 ns of simulation time are aligned to the original model for spySrtA–LPATA (*top left, yellow peptide*), spySrtA–LPATS (*top right, cyan peptide*), and spySrtA–LPAT-LII (*bottom, green peptide*) simulations. A clipped version of the spySrtA–LPAT-LII simulation is shown in the *bottom right* to highlight the relative inflexibility of the interpeptide bridge dialanine residues as compared with the rest of the pentapeptide stem. For all images in (*B*), the spySrtA protein is in *gray cartoon*, and the peptides are shown in *stick* representation and colored by heteroatom (N = *blue* and O = *red*). *C*, the distribution of observed distances between the C208A CB (C$_\beta$) and P1 Thr C (main-chain carbonyl carbon) is shown as a function of its probability. The highest probability distance is approximately 3.8 Å. PDB, Protein Data Bank; spySrtA, *Streptococcus pyogenes* class A sortase.

well as the average RMSF of backbone atoms in each peptide (Fig. S7B). The RMSF of the P1′ Ser in the LPATS peptide was also similar to that of the P1′ Ala in either the LPATA or LPAT-LII simulations.

We also analyzed the distance distribution between the C208A methyl group (or C_β atom) and that of the P1 Thr carbonyl C in the spySrtA–LPATA simulation, revealing that the most often sampled distance equals 3.8 Å (Fig. 6C). Surprisingly, our experimentally observed distance of 3.4 Å (Fig. 3D) was observed less than 5% of the time in the simulation; however, considering the C208A mutation and the standard C-S bond length of ~1.8 Å, this distribution of distances still positions the P1 Thr C in an ideal position for nucleophilic attack by the thiol group of the catalytic cysteine.

Finally, analysis of the average RMSF of every nonhydrogen atom in the LPAT-LII ligand was consistent with increased conformational dynamics for the pentapeptide stem portion (Fig. S7C). We see a dramatic increase in flexibility in atoms in the LII pentapeptide, as compared with the LPAT and interpeptide bridge sequences (Fig. S7C). Specifically, this increase begins at the Cε atom of the lysine side chain and gets progressively larger moving down the lysine side chain toward the pentapeptide stem. This was also clearly evident in the alignment of representative frames (taken every 45 ns) from the MD trajectory of the spySrtA–LPAT-LII system (Fig. 6B). Taken together, these MD simulations strongly support our described structure-based conclusions.

Discussion

As we highlight in Figure 1, there are multiple key states in the SrtA catalytic cycle when attaching a protein to the cell surface of Gram-positive bacteria. Facilitated by a conserved Cys-His-Arg triad, the apo enzyme (state 1) recognizes a motif within the CWSS on the C terminus of a target protein (state 2) and cleaves the peptide between the P1 Thr and P1′ Gly residues (or other P1′ residues *in vitro*), presumably forming a tetrahedral oxyanion that resolves to generate a thioacyl–enzyme intermediate (state 3). Nucleophilic attack by the N-terminal amine of the interpeptide bridge of LII on the carbonyl carbon of the P1 Thr residue leads to a second tetrahedral oxyanion intermediate that collapses into the final ligation product and completes the transpeptidation reaction, whereby the initial target sequence (minus the P1′ residue and all residues C terminal to this position) is covalently attached to LII (state 4) (8, 10, 44). Using spySrtA as a model, we solved structures that experimentally show how the full LPXTX substrate is recognized by the enzyme (state 2) as well as how the final ligation product is accommodated within the enzyme active site (state 4). In addition, we used our peptide-bound structures and energy minimization to model the acyl–enzyme intermediate (state 3); thus, providing a nearly comprehensive structural view of the spySrtA catalytic mechanism.

Considered alongside other SrtA structures that contain bound substrate mimetics, the studies reported here both reaffirm certain common structural features and reveal new

insights. As described previously, we observe several position-specific interactions similar to those first reported for the peptidomimetic-bound structures of saSrtA and baSrtA (13, 14). Our observed interactions at the P4 Leu, P3 Pro, P2 Ala, and P1′ Ala/Ser positions also support additional data on substrate selectivity in class A sortases (9, 15, 27, 45). Finally, in our LPAT-LII structure, we see both "Thr-in" and "Thr-out" conformations, molecular orientations that have also been previously described (8, 13, 14).

However, apart from the orientation of the P1 Thr side chain, other aspects of the positioning of this residue reveal unique attributes of our spySrtA complexes that differ from prior work with baSrtA and saSrtA (13, 14). It was previously suggested that the carbonyl group of the P1 Thr may be stabilized by contacts with the highly conserved Arg residue that forms part of the ubiquitous Cys-His-Arg triad found in sortases. This proposed interaction would further allow the Arg side chain to stabilize tetrahedral oxyanion intermediates generated during the sortase-catalyzed transpeptidation reaction. While we do find that the conserved Arg (R216) of spySrtA is critical for enzyme function (Figs. 4D and S4E), and appears to play a role in positioning the LPXTX motif through direct contacts with the P4 and P3 carbonyl groups (Fig. 3B), we see no evidence for interactions with the P1 carbonyl. Indeed, the P1 carbonyl in our complexes is projected away from the R216 side chain, and instead forms interactions with a series of other sites (Figs. 3B and 4C), including the hydroxyl group of a conserved Thr residue (T207) adjacent to the active site Cys (C208). These same contacts would also appear to provide a suitable oxyanion hole for stabilizing high-energy reaction intermediates, which is supported by our finding that a T207A mutant of spySrtA was essentially inactive (Figs. 4D and S4E).

Our observations with the P1 Thr indicate that further work on the exact role of the conserved Arg residue in sortase catalysis is warranted. Along these lines, intriguing results from a recent directed evolution study suggest that the conserved Arg in sortase A enzymes may primarily be responsible for substrate positioning and binding, as opposed to stabilization of catalytic intermediates. Specifically, an engineered variant of saSrtA was reported that is selective for an LMVGG substrate motif (46). Remarkably, in this enzyme, the conserved Arg of wildtype saSrtA was mutated to Ser, and yet it remained an efficient transpeptidase. While it is possible that this highly mutated saSrtA variant acquired a series of compensatory mutations that negated the need for Arg to stabilize high-energy oxyanion intermediates, we would argue an alternate interpretation that the wildtype Arg is not critical for creating an oxyanion hole and rather its primary function is substrate binding and controlling substrate selectivity. The Arg to Ser mutation in the LMVGG-specific saSrtA variant is thus understood as contributing to a change in substrate selectivity as opposed to representing a fundamental change in the catalytic mechanism.

We anticipate that our work will also prove useful in the continued development of SML protein modification strategies. Structure-guided engineering efforts have already seen

success in generating sortases with altered substrate selectivity or increased activity, as well as a saSrtA mutant that no longer requires a Ca^{2+} cofactor (9, 47–49). Moving forward, further optimization of spySrtA and other class A sortases for use in SML can be envisioned based on the molecular characteristics elucidated in the spySrtA complexes and related structures presented here. In addition, the extended target binding cleft revealed in our spySrtA–LPAT-LII structures suggests that portions of the substrate outside the LPXTX motif could make specific contacts with the spySrtA enzyme, for example, residues in the P2′ site. A systematic exploration of how these positions impact enzymatic activity *in vitro* may therefore be helpful in optimizing SML using spySrtA and other class A sortases. A similar approach has already proven beneficial for saSrtA, where it is known that a P2′ Gly residue generally provides superior reactivity *in vitro* (50).

In summary, this work reports the first crystal structures of spySrtA bound to an LPXTX substrate as well as a model of the *in vivo* ligation product involving LII. These structures reveal new details on substrate recognition by bacterial sortases, which may prove valuable for the use of sortases as tools for protein engineering. More broadly, this work improves our understanding of the fundamental enzymology of this large and clinically relevant class of bacterial enzymes.

Experimental procedures

Expression and purification of spySrtA protein

Wildtype spySrtA, C208A spySrtA, T207A spySrtA, and R216A spySrtA genes were recombinantly expressed using *Escherichia coli* BL21 (DE3) cells in the pET28a(+) vector (Genscript), as previously described (23). The wildtype sequence used matches that of the published spySrtA structure, PDB ID: 3FN5 (17). Briefly, transformed cells were grown at 37 °C in LB media to an absorbance of 0.6 to 0.8 at 600 nm, followed by induction using 0.15 mM IPTG for 18 to 20 h at 18 °C. The cells were harvested in lysis buffer (0.05 M Tris [pH 7.5], 0.15 M NaCl, and 0.5 mM EDTA], and whole-cell lysate was clarified using centrifugation, followed by filtration of the supernatant. Initial purification was conducted using a 5 ml HisTrap HP column (Cytiva) and wash (0.05 M Tris [pH 7.5], 0.15 M NaCl, 0.02 M imidazole, and 0.001 M Tris(2-carboxyethyl)phosphine) and elution (wash buffer with 0.3 M imidazole) buffers.

Following immobilized metal affinity chromatography, the His tag was proteolyzed off the N terminus of the C208A spySrtA protein using tobacco etch virus protease overnight at 4 °C and a ratio of ∼1:100 (tobacco etch virus:protein). The proteins used for activity assays (wildtype, T207A, and R216A) were not cleaved, consistent with our previous work (9, 23). After collecting the flow-through of a second 5 ml HisTrap HP column (wash buffer identical to that described previously), size-exclusion chromatography (SEC) was conducted using a HiLoad 16/600 Superdex 75 column (Cytiva) in SEC running buffer (0.05 M Tris [pH 7.5], 0.15 M NaCl, and 0.001 M Tris(2-carboxyethyl)phosphine). Purified protein fractions corresponding to the monomeric peak were pooled and concentrated using an Amicon Ultra-15 Centrifugal Filter Unit (10,000 NWML). Protein concentrations were determined using theoretical extinction coefficients calculated using ExPASy ProtParam (Swiss Institute of Bioinformatics; https://web.expasy.org/protparam/) (51). Protein not immediately used was flash frozen in SEC running buffer and stored at –80 °C.

The purity, monomeric state, and identity of purified enzymes were confirmed by SDS-PAGE, analytical SEC, and LC–ESI–MS, respectively. For LC–ESI–MS, analyses were performed on an Agilent 6545XT AdvanceBio Q-TOF system interfaced with an Agilent 1290 HPLC system. Separations upstream of the Q-TOF were achieved with a Phenomenex Aeris 3.6 μM WIDEPORE C4 200 Å column (100 × 2.1 mm) (H_2O [0.1% formic acid]/MeCN [0.1% formic acid] mobile phase at 0.3 ml/min, method: hold 10% MeCN 0.0 to 1.0 min, linear gradient of 10 to 90% MeCN 1.0 to 9.0 min, hold 90% MeCN 9.0 to 11.0 min, linear gradient of 90 to 10% MeCN 11.0 to 11.1 min, re-equilibrate at 10% MeCN 11.1 to 15.0 min). Deconvolution of protein charge ladders was achieved using Agilent MassHunter BioConfirm software (version 10.0). The expected and observed molecular weights (MWs) for all proteins in this study were as follows: wildtype spySrtA (calculated average MW = 20,657.5 Da, observed = 20,657.6 Da), C208A spySrtA (calculated average MW = 18,573.3 Da, observed = 18,573.4 Da), T207A spySrtA (calculated average MW = 20,627.3 Da, observed = 20,627.5 Da), and R216A spySrtA (calculated average MW = 20,572.3 Da, observed = 20,572.5 Da). Representative MS data for wildtype and C208A spySrtA are also provided in Fig. S1.

Peptide synthesis

Model peptide substrates used in crystallization and/or enzyme assays with the general structure *Abz*-LPATXGK(Dnp)-NH_2 (*Abz* = 2-aminobenzoyl; *Dnp* = 2,4-dinitrophenyl; NH_2 = C-terminal primary amide) were synthesized and purified as previously described (9). The *Ac*-LPATSG-NH_2 peptide (*Ac* = acetyl; NH_2 = C-terminal primary amide) used for spySrtA–LPATS cocrystallization was purchased from Biomatik. *5-FAM-Ahx*-LPATGG-NH_2 (5-FAM-Ahx = 5-carboxyfluorescein linked *via* an aminohexanoic acid linker; NH_2 = C-terminal primary amide) used in attempted cocrystallization studies was also purchased from Biomatik. Finally, the synthesis of *Abz*-LPAT-LII was achieved *via* manual Fmoc solid phase peptide synthesis. Full experimental details for the preparation of LPAT-LII are provided in the Supporting information and Fig. S8.

HPLC and LC–MS characterization of spySrtA-catalyzed reactions

LPATS/LPATA/LPAT-LII peptide substrates (50 μM), alanine amide nucleophile (5 mM), and wildtype spySrtA enzyme (5 μM in the reaction with LPATS, otherwise 1 μM) were combined at room temperature and incubated for the times indicated. All reactions contained 10% (v/v) sortase reaction buffer (500 mM Tris [pH 7.5] and 1500 mM NaCl) as

well as ≤1.1% (v/v) residual dimethyl sulfoxide from the peptide substrate stock solutions. Reactions were analyzed using a Dionex Ultimate 3000 HPLC system interfaced with an Advion CMS expression[L] mass spectrometer. Separations were achieved with a Phenomenex Kinetix 2.6 µM C18 100 Å column (100 × 2.1 mm) (aqueous [95% H_2O, 5% MeCN, 0.1% formic acid]/MeCN [0.1% formic acid] mobile phase at 0.3 ml/min, gradients adjusted for each substrate to achieve separation between relevant reaction components).

Fluorescence assay for sortase activity

Enzyme assays for assessing the reactivity of wildtype spySrtA *versus* the T207A and R216A mutants were conducted using a Biotek Synergy H1 plate reader as previously described (9). Briefly, *Abz*-LPATXGK(Dnp)-NH_2 peptide substrates (50 µM final concentration) were incubated with hydroxylamine nucleophile (5 mM) and sortase enzyme (5 µM) at room temperature. All reactions contained 10% (v/v) 10× sortase reaction buffer (500 mM Tris [pH 7.5] and 1500 mM NaCl) and small amounts of residual dimethyl sulfoxide (≤0.9% v/v) from the peptide stock solutions. The fluorescence intensity of each reaction well was measured at 2 min time intervals over a 2 h period (λ_{ex} = 320 nm; λ_{em} = 420 nm; and detector gain = 75). All reactions were performed in triplicate, and fluorescence intensity (in relative fluorescence units) over time was plotted using Kaleidagraph 5.01 (Synergy software).

Crystallization of spySrtA complex structures

The C208A spySrtA protein was crystallized at approximately 20 mg/ml or 1.1 mM. Peptide (LPATA, LPATS, or LPAT-LII), at 1 mM final concentration, was incubated with protein in a 1:1 ratio at room temperature for approximately 1 h prior to crystallization by hanging drop vapor diffusion method using a 500 µl well solution to protein solution ratio of 1:1, for a final drop volume of 4 µl (2 µl + 2 µl). Crystallization conditions were optimized using those for the wildtype apo protein (17). The crystallization conditions for the crystals used for data collection were (for all, containing C208A spySrtA): LPATA (0.1 M sodium acetate, 34% [w/v] PEG 8000, 0.1 M Tris [pH 6]), LPATS (0.1 M sodium acetate, 30% [w/v] PEG 8000, 0.1 M Tris [pH 6]), LPAT-LII "Thr-in" (0.15 M sodium acetate, 26% [w/v] PEG 8000, 0.1 M Tris [pH 6]), and LPAT-LII "Thr-out" (same conditions as LPAT-LII "Thr-in"). Microseeding was used to obtain crystals of suitable diffraction quality for structure determination with the LPAT-LII-bound complexes, using initial crystals that grew in conditions of higher PEG 8000 concentration (>30% [w/v], consistent with the other conditions described). As described in the main text, glycerol was used as a cryoprotectant for the C208A spySrtA–LPATA crystal (cryo: crystallization conditions plus 12% [w/v] glycerol), but for the other crystals, PEG 400 was used (cryo: 0.15 M sodium acetate, 10% [w/v] PEG 8000, 40% [w/v] PEG 400, 0.1 M Tris [pH 6]). The crystals were flash cooled by plunging into liquid nitrogen.

Data collection, structure determination, and protein analyses

Data were collected at the Advanced Light Source at Lawrence Berkeley National Laboratory on beamline 5.0.1 and 5.0.2, at λ= 1.00004 nm or 0.97741 nm over 360°, with ΔΦ = 0.25° frames and an exposure time of 0.5 s per frame. Data were processed using the XDS package (Max Planck Institute for Medical Research) (Table 1) (52, 53). Molecular replacement was performed using Phenix with spySrtA (PDB ID: 3FN5) used as the search model. Refinement was performed using Phenix (The Phenix Industrial Consortium), manual refinement was done using Coot (MRC Laboratory of Molecular Biology), and model geometry was assessed using MolProbity (Duke University) and the PDB validation server (54–56). Coordinates for the Abz moiety in C208A spySrtA–LPATA and LPAT-LII were initially determined using phenix.eLBOW from the SMILES (Simplified Molecular Input Line Entry System) strings rendered using ChemDraw (PerkinElmer Informatics) (57). All crystal data and refinement statistics are provided in Table 1. Structural analyses and figure rendering were done using PyMOL (Schrödinger software). PDB accession codes for the structures presented here are provided in Table 1 and are (for all, containing C208A spySrtA): LPATA (PDB ID: 7S51), LPATS (PDB ID: 7S40), LPAT-LII "Thr-in" (PDB ID: 7T8Y), and LPAT-LII "Thr-out" (PDB ID: 7T8Z).

MD simulations of spySrtA

All MD simulations were performed using GROMACS 2020.4 (GROMACS development teams at the KTH Royal Institute of Technology and Uppsala University) with the AMBER99SB*-ILDN force fields (43, 58–60). Additional details and relevant references are in the Supporting information section (61–70).

For energy minimization of the spySrtA–LPAT model, a steepest descent energy minimization was performed on the solvated system with a maximum force tolerance of 500 kJ/mol/nm. The steepest descent converged in 2998 steps.

Data availability

All data are contained in the article and the supporting information.

References

1. Ton-That, H., and Schneewind, O. (1999) Anchor structure of staphylococcal surface proteins. IV. Inhibitors of the cell wall sorting reaction. *J. Biol. Chem.* **274**, 24316–24320
2. Mazmanian, S. K., Liu, G., Ton-That, H., and Schneewind, O. (1999) Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285**, 760–763
3. 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
4. Spirig, T., Weiner, E. M., and Clubb, R. T. (2011) Sortase enzymes in Gram-positive bacteria. *Mol. Microbiol.* **82**, 1044–1059
5. Antos, J. M., Truttmann, M. C., and Ploegh, H. L. (2016) Recent advances in sortase-catalyzed ligation methodology. *Curr. Opin. Struct. Biol.* **38**, 111–118
6. Zhang, J., Liu, H., Zhu, K., Gong, S., Dramsi, S., Wang, Y. T., *et al.* (2014) Antiinfective therapy with a small molecule inhibitor of Staphylococcus aureus sortase. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 13517–13522
7. Hendrickx, A. P. A., 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
8. Jacobitz, A. W., Kattke, M. D., Wereszczynski, J., and Clubb, R. T. (2017) Sortase transpeptidases: structural biology and catalytic mechanism. *Adv. Protein Chem. Struct. Biol.* **109**, 223–264
9. Piper, I. M., Struyvenberg, S. A., Valgardson, J. D., Johnson, D. A., Gao, M., Johnston, K., *et al.* (2021) Sequence variation in the β7-β8 loop of bacterial class A sortase enzymes alters substrate selectivity. *J. Biol. Chem.* **297**, 100981
10. Perry, A. M., Ton-That, H., Mazmanian, S. K., and Schneewind, O. (2002) Anchoring of surface proteins to the cell wall of Staphylococcus aureus. III. Lipid II is an *in vivo* peptidoglycan substrate for sortase-catalyzed surface protein anchoring. *J. Biol. Chem.* **277**, 16241–16248
11. Clancy, K. W., Melvin, J. A., and McCafferty, D. G. (2010) Sortase transpeptidases: Insights into mechanism, substrate specificity, and inhibition. *Biopolymers* **94**, 385–396
12. 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
13. Chan, A. H., Yi, S. W., Terwilliger, A. L., Maresso, A. W., Jung, M. E., and Clubb, R. T. (2015) Structure of the Bacillus anthracis sortase A enzyme bound to its sorting signal: a flexible amino-terminal appendage modulates substrate access. *J. Biol. Chem.* **290**, 25461–25474
14. Suree, N., Liew, C. K., Villareal, V. A., Thieu, W., Fadeev, E. A., Clemens, J. J., *et al.* (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
15. Nikghalb, K. D., Horvath, N. M., Prelesnik, J. L., Banks, O. G. B., Filipov, P. A., Row, R. D., *et al.* (2018) Expanding the scope of sortase-mediated ligations by using sortase homologues. *Chembiochem* **19**, 185–195
16. Zong, Y., Bice, T. W., Ton-That, H., Schneewind, O., and Narayana, S. V. L. (2004) Crystal structures of Staphylococcus aureus sortase A and its substrate complex. *J. Biol. Chem.* **279**, 31383–31389
17. Race, P. R., Bentley, M. L., Melvin, J. A., Crow, A., Hughes, R. K., Smith, W. D., *et al.* (2009) Crystal structure of Streptococcus pyogenes sortase A: implications for sortase mechanism. *J. Biol. Chem.* **284**, 6924–6933
18. Antos, J. M., Chew, G. L., Guimaraes, C. P., Yoder, N. C., Grotenbreg, G. M., Popp, M. W., *et al.* (2009) Site-specific N- and C-terminal labeling of a single polypeptide using sortases of different specificity. *J. Am. Chem. Soc.* **131**, 10800–10801
19. Popp, M. W., Dougan, S. K., Chuang, T.-Y., Spooner, E., and Ploegh, H. L. (2011) Sortase-catalyzed transformations that improve the properties of cytokines. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3169–3174
20. Guimaraes, C. P., Witte, M. D., Theile, C. S., Bozkurt, G., Kundrat, L., Blom, A. E., *et al.* (2013) Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. *Nat. Protoc.* **8**, 1787–1799
21. Dai, X., Böker, A., and Glebe, U. (2019) Broadening the scope of sortagging. *RSC Adv.* **9**, 4700–4721
22. Freund, C., and Schwarzer, D. (2021) Engineered sortases in peptide and protein chemistry. *Chembiochem* **22**, 1347–1356
23. Gao, M., Johnson, D. A., Piper, I. M., Kodama, H. M., Svendsen, J. E., Tahti, E., *et al.* (2022) Structural and biochemical analyses of selectivity determinants in chimeric Streptococcus Class A sortase enzymes. *Protein Sci.* **31**, 701–715
24. Raz, A., Tanasescu, A. M., Zhao, A. M., Serrano, A., Alston, T., Sol, A., *et al.* (2015) Streptococcus pyogenes sortase mutants are highly susceptible to killing by host factors due to aberrant envelope physiology. *PLoS One* **10**, e0140784
25. Litou, Z. I., Bagos, P. G., Tsirigos, K. D., Liakopoulos, T. D., and Hamodrakas, S. J. (2008) Prediction of cell wall sorting signals in gram-positive bacteria with a hidden markov model: application to complete genomes. *J. Bioinform. Comput. Biol.* **6**, 387–401
26. Theile, C. S., Witte, M. D., Blom, A. E., Kundrat, L., Ploegh, H. L., and Guimaraes, C. P. (2013) Site-specific N-terminal labeling of proteins using sortase-mediated reactions. *Nat. Protoc.* **8**, 1800–1807

27. Schmohl, L., Bierlmeier, J., von Kügelgen, N., Kurz, L., Reis, P., Barthels, F., *et al.* (2017) Identification of sortase substrates by specificity profiling. *Bioorg. Med. Chem.* **25**, 5002–5007

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

29. Wójcik, M., Vázquez Torres, S., Quax, W. J., and Boersma, Y. L. (2019) Sortase mutants with improved protein thermostability and enzymatic activity obtained by consensus design. *Protein Eng. Des. Sel.* **32**, 555–564

30. Zhang, Y., Huang, D., Zhang, C., Meng, J., Tan, B., and Deng, Z. (2021) IQF characterization of a cathepsin B-responsive nanoprobe for report of differentiation of HL60 cells into macrophages. *RSC Adv.* **11**, 16522–16529

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

32. Zwart, P. H., Grosse-Kunstleve, R. W., Lebedev, A. A., Murshudov, G. N., and Adams, P. D. (2008) Surprises and pitfalls arising from (pseudo) symmetry. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **64**, 99–107

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

34. 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

35. Jacobitz, A. W., Wereszczynski, J., Yi, S. W., Amer, B. R., Huang, G. L., Nguyen, A. V., *et al.* (2014) Structural and computational studies of the Staphylococcus aureus sortase B-substrate complex reveal a substrate-stabilized oxyanion hole. *J. Biol. Chem.* **289**, 8891–8902

36. Tian, B.-X., and Eriksson, L. A. (2011) Catalytic mechanism and roles of Arg197 and Thr183 in the Staphylococcus aureus sortase A enzyme. *J. Phys. Chem. B* **115**, 13003–13011

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

38. Schleifer, K. H., and Kandler, O. (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol. Rev.* **36**, 407–477

39. Münch, D., and Sahl, H.-G. (2015) Structural variations of the cell wall precursor lipid II in Gram-positive bacteria - impact on binding and efficacy of antimicrobial peptides. *Biochim. Biophys. Acta* **1848**, 3062–3071

40. Hadi, T., Hazra, S., Tanner, M. E., and Blanchard, J. S. (2013) Structure of MurNAc 6-phosphate hydrolase (MurQ) from Haemophilus influenzae with a bound inhibitor. *Biochemistry* **52**, 9358–9366

41. Vollmer, W., Blanot, D., and de Pedro, M. A. (2008) Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **32**, 149–167

42. Navarre, W. W., and Schneewind, O. (1999) Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol. Mol. Biol. Rev.* **63**, 174–229

43. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., *et al.* (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958

44. Ruzin, A., Severin, A., Ritacco, F., Tabei, K., Singh, G., Bradford, P. A., *et al.* (2002) Further evidence that a cell wall precursor [C(55)-MurNAc-(peptide)-GlcNAc] serves as an acceptor in a sorting reaction. *J. Bacteriol.* **184**, 2141–2147

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

46. Podracky, C. J., An, C., DeSousa, A., Dorr, B. M., Walsh, D. M., and Liu, D. R. (2021) Laboratory evolution of a sortase enzyme that modifies amyloid-β protein. *Nat. Chem. Biol.* **17**, 317–325

47. Schmohl, L., Bierlmeier, J., Gerth, F., Freund, C., and Schwarzer, D. (2017) Engineering sortase A by screening a second-generation library using phage display. *J. Pept. Sci.* **23**, 631–635

48. Zou, Z., Nöth, M., Jakob, F., and Schwaneberg, U. (2020) Designed Streptococcus pyogenes sortase A accepts branched amines as nucleophiles in sortagging. *Bioconjug. Chem.* **31**, 2476–2481

49. Hirakawa, H., Ishikawa, S., and Nagamune, T. (2012) Design of Ca2+-independent Staphylococcus aureus sortase A mutants. *Biotechnol. Bioeng.* **109**, 2955–2961

50. Pritz, S., Wolf, Y., Kraetke, O., Klose, J., Bienert, M., and Beyermann, M. (2007) Synthesis of biologically active peptide nucleic acid-peptide conjugates by sortase-mediated ligation. *J. Org. Chem.* **72**, 3909–3912

51. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., *et al.* (1999) Protein identification and analysis tools in the ExPASy server. *Met. Mol. Biol.* **112**, 531–552

52. Kabsch, W. (2010) Xds. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 125–132

53. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 133–144

54. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., *et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 12–21

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

56. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., *et al.* (2010) Phenix: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 213–221

57. Moriarty, N. W., Grosse-Kunstleve, R. W., and Adams, P. D. (2009) Electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **65**, 1074–1080

58. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005) Gromacs: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718

59. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., *et al.* (2015) Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25

60. Lindahl, H., and Spoel, V. D. (2020) *GROMACS 2020.4 Manual*, Zenodo, Geneva, Switzerland

61. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174

62. Cornell, W. D., Cieplak, P., Bayly, C. I., and Kollmann, P. A. (1993) Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.* **115**, 9620–9631

63. Barca, G. M. J., Bertoni, C., Carrington, L., Datta, D., De Silva, N., Deustua, J. E., *et al.* (2020) Recent developments in the general atomic and molecular electronic structure system. *J. Chem. Phys.* **152**, 154102

64. Lu, T., and Chen, F. (2012) Multiwfn: a multifunctional wavefunction analyzer. *J. Comput. Chem.* **33**, 580–592

65. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577

66. Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101

67. Parrinello, M. (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182

68. Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **4**, 116–122

69. Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., *et al.* (2014) XSEDE: accelerating scientific discovery. *Comput. Sci. Eng.* **16**, 62–74

70. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C., and Bussi, G. (2014) PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613

ASBMB