Most of Gram-positive bacteria anchor surface proteins to the peptidoglycan cell wall by sortase, a cysteine transpeptidase that targets proteins displaying a cell wall sorting signal. Unlike other bacteria, *Clostridium difficile*, the major human pathogen responsible for antibiotic-associated diarrhea, has only a single functional sortase (SrtB). Sortase’s vital importance in bacterial virulence has been long recognized, and *C. difficile* sortase B (Cd-SrtB) has become an attractive therapeutic target for managing *C. difficile* infection. A better understanding of the molecular activity of Cd-SrtB may help spur the development of effective agents against *C. difficile* infection. In this study, using site-directed mutagenesis, biochemical and biophysical tools, LC-MS/MS, and crystallographic analyses, we identified key residues essential for Cd-SrtB catalysis and substrate recognition. To the best of our knowledge, we report the first evidence that a conserved serine residue near the active site participates in the catalytic activity of Cd-SrtB and also SrtB from *Staphylococcus aureus*. The serine residue indispensable for SrtB activity may be involved in stabilizing a thioacyl-enzyme intermediate because it is neither a nucleophilic residue nor a substrate-interacting residue, based on the LC-MS/MS data and available structural models of SrtB–substrate complexes. Furthermore, we also demonstrated that residues 163–168 located on the β6/β7 loop of Cd-SrtB dominate specific recognition of the peptide substrate PPKTG. The results of this work reveal key residues with roles in catalysis and substrate specificity of Cd-SrtB.

*Clostridium difficile* infection (CDI) is a global healthcare problem associated with morbidity and mortality for hospitalized patients (1, 2). CDI frequently occurs in older and severely ill patients who are in long-term care facilities (2). The mortality rate of CDI patients ranges from 5 to 40% (3, 4), and the rate of recurrence occurring within 30 days after treatment is to to 30% (5). *C. difficile* is a Gram-positive, spore-forming, and anaerobic bacterium (6), which is the causing agent of a multitude of intestinal diseases ranging from mild diarrhea to severe inflammatory bowel perforations or pseudomembranous colitis (7, 8). CDI is transmitted through bacterial spores or from person to person by the fecal–oral route. Patients infected with *C. difficile* spores are mostly by direct contact with contaminated surfaces and symptomatic patients in the hospital setting (9). Current treatment of CDI mainly relies on the administration of antibiotics such as metronidazole, vancomycin, and fidaxomicin to alleviate immediate symptoms for patients (2, 10, 11). Furthermore, alternative treatment options are considered, such as fecal microbiota transplantation as a means of re-establishing a normal microbiota profile for patients with recurrent CDI (12). Nevertheless, fecal microbiota transplantation is still not widely applied, and broad-spectrum antibiotic therapy remains the first choice in managing CDI (6). However, antibiotic use is a major risk factor for recurrent CDI and *C. difficile* superinfection (5), because of the disruption of the normal gut microbiota (6). The incidence rate of multiple recurrent CDI and frequency of antimicrobial treatment failures have significantly increased (13). Moreover, *C. difficile* 630 is a multidrug-resistant strain whose genome was isolated from a patient with pseudomembranous colitis (14). Given that the prevalence of antibiotic-resistant bacteria is rising, nonconventional antimicrobial therapies are in demand, and efforts searching for developing alternative anti-infective drugs for the treatment of CDI patients are growing (15).

This work was supported by Taiwan Protein Project Grant AS-KPQ-109-TPP2 and by Grants 106-2320-B-006-010, 107-2320-B-006-048, and 108-2320-B-006-012 from the Ministry of Science and Technology of Taiwan (to S. W.). The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Table S1 and Figs. S1–S3.

**The atomic coordinates and structure factors (codes 6KYC and 6KXY) have been deposited in the Protein Data Bank (http://wwpdb.org/).**

1 These authors contributed equally to this work.

2 To whom correspondence should be addressed: Dept. of Microbiology and Immunology, College of Medicine, National Cheng Kung University, Tainan 701, Taiwan. Tel.: 886-6-2353535, Ext. 5634; Fax: 886-6-2082705; E-mail: sswang23@mail.ncku.edu.tw.

This is an Open Access article under the CC BY license.

---

3 The abbreviations used are: CDI, *Clostridium difficile* infection; SrtA, sortase A; SrtB, sortase B; Cd-SrtB, *C. difficile* sortase B; Sa-SrtB, *S. aureus* sortase B; Cd-SrtB<sub>Δ<sub>Y263N,LS</sub></sub>; loop swap mutant (S163T, D164K, Y165D, D166N, and L168I) in *C. difficile* sortase B; RMSD, root-mean-square deviation; PDB, Protein Data Bank; Edans, 5-((2-aminoethyl)amino)naphthalene-1-sulfonic acid; Dabcyl, 4-([4-(dimethylamino)phenyl]azo) benzoic acid.
Functional residues of C. difficile sortase B

Ser-207 is essential for the catalytic activity of Cd-SrtB

A sequence alignment of SrtB from Gram-positive bacteria including B. anthracis, S. aureus, Streptococcus pyogenes, and other Gram-positive bacteria showed that a highly conserved serine residue is located near the catalytic cysteine residue (Fig. S1, A and B). Moreover, superimposition of structures of SrtB from C. difficile (PDB code 5GY1) (40), S. aureus (PDB code 1NG5) (39), B. anthracis (PDB code 1RZ2) (39), and S. pyogenes (PDB code 3PSQ) (43) also revealed the structural conservation of the serine residue (Fig. S2). These suggest that the conserved His–Cys–Asp triad of Cys/Ser protease (39). It remains to be explored whether other residues located near the active site also contribute to catalytic activity.

Crystallographic structures of Cd-SrtB determined by our group (40) and Chambers et al. (41) reveal that the overall structure of Cd-SrtB conforms the canonical sortase fold. In addition, our previous study also constructed and validated an in silico model of a Cd-SrtB–PPKTG complex and elucidated the molecular interaction governing the PPKTG recognition (40). It was suggested that all sortases form similar sorting signal–binding grooves. The direct evidences came from the currently available structures of sortase–substrate analog complexes: Sa-SrtA-LPAT* (32), Sa-SrtB-NPQT* (33), and Bacillus anthracis sortase A complexed with LPAT* (34). The binding grooves are primarily formed by strands β4 and β7 and loops connecting β2/β3, β3/β4, β6/β7, and β7/β8. The β6/β7 loop plays a substantial role in interacting and discriminating sorting motif because studies showed that the replacement of the β6/β7 loop in Sa-SrtA with the corresponding site from Sa-SrtB results in converting the specificity profile of Sa-SrtA to Sa-SrtB (33). Our previous work has defined two residues located within β6/β7 loop of Cd-SrtB, Ser163 and Tyr167, that are to be in the direct contact with the substrate peptide PPKTG (40). The structural variation in β6/β7 loop is significant between class A and class B sortase. It draws our attention whether the less dissimilar β6/β7 loop in class B sortase also acts as a determinant for bacteria-specific sorting signal recognition.

Significant efforts have been made to seek the novel therapeutics for CDI, and sortase is one of the most considered targets (28, 38, 41, 42). Therefore, a better understanding of the molecular basis of Cd-SrtB could provide insightful information to facilitate the development of Cd-SrtB–based agents against CDI. Cd-SrtB is not as extensively studied as Sa-SrtA. This work was initiated by a sequence alignment and structural superimposition of class B sortases showing a conserved serine residue near the active site that drew our attention to investigate whether this serine residue participates in catalytic activity. Moreover, we also performed the β6/β7 loop swap between Cd-SrtB and Sa-SrtB to study whether the β6/β7 loop also dictates substrate specificity of class B sortases among different bacteria. Our results demonstrated that the conserved serine residue in proximity to the active site is indispensable for the catalytic activity of Cd-SrtB and Sa-SrtB and that the β6/β7 loop dominates the molecular interactions governing the specific motif recognition.

Results

Ser–207 is essential for the catalytic activity of Cd-SrtB

A sequence alignment of SrtB from Gram-positive bacteria including B. anthracis, S. aureus, Streptococcus pyogenes, and other Gram-positive bacteria showed that a highly conserved serine residue is located near the catalytic cysteine residue (Fig. S1, A and B). Moreover, superimposition of structures of SrtB from C. difficile (PDB code 5GY1) (40), S. aureus (PDB code 1NG5) (39), B. anthracis (PDB code 1RZ2) (39), and S. pyogenes (PDB code 3PSQ) (43) also revealed the structural conservation of the serine residue (Fig. S2). These suggest that the conserved

For Gram-positive bacteria, the attachment of virulence-associated surface proteins to the peptidoglycan cell wall is mediated by sortase enzymes (16). Sortases are cysteine transpeptidases that function in covalently anchoring of surface proteins to the cell wall envelope (17) and in constructing pili (18, 19). Based on the primary sequences and biological roles, sortases are classified into six classes (A–F) (20–22). All the characterized sortases possess a signal sequence that enables their translocation across the membrane via the Sec apparatus and target proteins consisting of a cell wall sorting signal in the C-terminal region (23, 24). The characteristic five-residue sortase-recognition sequence motif located within cell wall sorting signal of substrate proteins is class- and/or bacteria-specific. Class A sortase enzymes anchor many surface proteins in cell wall and play a housekeeping role. Sortase A from Staphylococcus aureus (Sa-SrtA), the best studied sortase, recognizes the LPXTG motif of its substrates and initiates catalysis by employing the thiolate of the active site cysteine residue to cleave the peptide bond between Thr and Gly residues (25). This process results in generation of an acyl-enzyme intermediate by a thioester linkage between the cysteine residue of sortase and its substrate (26, 27). Subsequently, a secondary substrate (lipid II or pilin) is recognized by sortase that catalyzes a reaction in which the amine group from lipid II or a lysine residue within a pilin subunit nucleophilic attacks the thioacyl bond and relieves the sortase-protein thioacyl intermediate (23). Class B sortase can perform distinct tasks including heme uptake and pilus polymerization. In contrast to many other bacteria that typically have multiple sortases, C. difficile possesses only sortase B (Cd-SrtB), which attaches seven proteins to cell wall and appears to play a general role (28). Cd-SrtB recognizes a sorting signal containing a (S/P)PXTG motif (29) that differs from the conserved class A sortase-recognition LPXTG motif and the S. aureus sortase B (Sa-SrtB) NP(Q/K)TN sorting motif (30, 31). The molecular origin of how Cd-SrtB discriminates (S/P)PXTG from LPXTG or NP(Q/K)TN is not well-understood.

Structural and computational studies have provided in-depth insights into the molecular mechanism of sortase-mediated catalysis and have advanced our understanding on the complicated process of substrate recognition (32–34). Structures of the catalytic domains of sortases share a conserved eight-stranded β-barrel core harboring a His–Cys–Arg triad essential for catalysis (23, 33). Characteristic structural features and variations among different classes of sortases are observed, and the structural differences dictating the class-specific function and substrate specificity are described (23, 24). The key catalytic residues His, Cys, and Arg are structurally equivalent in the family of sortase and are in proximity to one another within the active site located at the edge of β-barrel. The His residue functions as a general acid/base during acyl and decacyl process (35, 36); the Arg residue is believed to play an important role in stabilizing the acyl-enzyme intermediate by forming an oxyanion hole (21, 33, 36, 37). Although the residues that constitute the active site are believed to be His–Cys–Arg, studies also reported that other residues also participate in catalysis (23, 38, 39). The crystal structure of Sa-SrtB revealed that Asp196 also constitutes the catalytic site, similar to the catalytic His–Cys–Asp triad of Cys/Ser protease (39). It remains to be explored whether other residues located near the active site also contribute to catalytic activity.
serine residue in proximity to the active site may have a functional role in SrtB catalysis.

To investigate whether the equivalent conserved serine residue Ser-207 in C. difficile is involved in the catalytic activity of Cd-SrtB, a mutant replacing Ser-207 with Ala was generated by site-directed mutagenesis. Purified recombinant WT Cd-SrtB and S207A mutant with a deletion of 26 residues at the N-terminal transmembrane region are designated as Cd-SrtB\textsubscript{N26,WT} and Cd-SrtB\textsubscript{N26,S207A}. The catalytic activity of Cd-SrtB\textsubscript{N26,S207A} was assessed by FRET-based assay using PPKTG-containing peptide (Dabcyl-PVPPKTGDSTTIIGE-Edans) as a substrate described previously (40). The results showed that Cd-SrtB\textsubscript{N26,S207A} exhibited reduced cleavage activity compared with Cd-SrtB\textsubscript{N26,WT} (Fig. 1, A and B), indicating the participation of Ser-207 in catalysis. To study whether two single-site mutations on Ser-207 and Cys-209 produce additive effect, the double mutant Cd-SrtB\textsubscript{N26,S207A/C209A} was generated. The results from FRET-based assay showed a nonadditive effect of the double mutations on catalytic activity of Cd-SrtB (Fig. 1, C and D). To confirm the substrate specificity of recombinant Cd-SrtB\textsubscript{N26,WT}, LC-MS/MS was performed to identify the substrate cleavage site. The PPKTG-containing peptide (Dabcyl-PVPPKTGDSTTIIGE-Edans) and scrambled peptide (H-PVGSSTPDSTTIIGE-OH) were incubated with Cd-SrtB\textsubscript{N26,WT} for LC-MS/MS studies. The MS/MS spectrum of reaction products revealed a precursor ion at \( m/z \) 570.7421 with \( z = 2 \) that corresponds to the predicted peptide mass of GDSSTTIIGE-Edans (Fig. 2, A, upper panel), confirming that Cd-SrtB\textsubscript{N26,WT} cleaved PPKTG-containing peptide between Thr and Gly residues (Fig. 2, A and B). In addition, the MS/MS spectrum of reaction mixture of Cd-SrtB\textsubscript{N26,WT} with scrambled peptide showed a predominant ion at \( m/z \) 730.8409 with \( z = 2 \), corresponding to the mass of the intact scrambled peptide (Fig. 2 C). Furthermore, the LC-MS/MS results showed that the mutant Cd-SrtB\textsubscript{N26,C209A} did not cleave the substrate peptide between Thr and Gly (Fig. 2A, lower panel), verifying that Cys-209 is the active nucleophilic residue of Cd-SrtB.

To further understand whether the mutation at Ser-207 influences the structure of Cd-SrtB that leads to the reduced enzymatic activity, the crystal structure of Cd-SrtB\textsubscript{N26,S207A} was determined at 2.6 Å resolution (Fig. 3A). Structural superposition of Cd-SrtB\textsubscript{N26,S207A} with Cd-SrtB\textsubscript{N26,WT} showed that Cd-SrtB\textsubscript{N26,S207A} mutant has a nearly identical overall structure with a root-mean-square deviation (RMSD) of 0.174 Å for 176 C\textsubscript{\textalpha} coordinates (Fig. 3B). These data revealed that the mutation at Ser-207 did not alter the
Functional residues of C. difficile sortase B
main-chain conformation on the Cd-SrtB structure but had impact on the catalytic function of Cd-SrtB.

The conserved serine residue in Sa-SrtB is correspondingly indispensable for activity

To explore whether the catalytic contribution of the conserved serine residue is specific to *C. difficile*, we extended our studies to Sa-SrtB. The analogous experiments were performed with Sa-SrtB to study whether the equivalent serine residue Ser-192 in Sa-SrtB also participates in catalytic activity. The Sa-SrtB<sub>H9004N29</sub>,WT, Sa-SrtB<sub>H9004N29,S192A</sub>, Sa-SrtB<sub>H9004N29,C194A</sub>, and Sa-SrtB<sub>H9004N29,S192A C194A</sub> were generated, purified, and subjected to FRET-based assay using NPQTN-containing peptide as a substrate. Similar to the results from Cd-SrtB, the data showed that the Sa-SrtB<sub>H9004N29,S192A</sub> displayed reduced enzymatic activity comparable with the catalytic mutant Sa-SrtB<sub>H9004N29,C194A</sub>, and no additive effect was observed on double mutations (Fig. 4). These studies demonstrated that a conserved serine residue plays a role in catalysis in both Cd-SrtB and Sa-SrtB.

Characterize effects of mutations at His–Cys–Arg triad on Cd-SrtB activity

The conserved serine residue was found to be essential for the catalytic activity of Cd-SrtB and Sa-SrtB in this study that led us to be keen on reconfirming the functional contribution of His-116 and Arg-217 in the active site of Cd-SrtB. The mutational impact of His–Cys–Arg catalytic triad on sortase activity has been studied in *S. aureus* (19, 33), *B. anthracis* (39), and *S. pyogenes* (44) but has never been assessed in *C. difficile*. Therefore, we created mutants Cd-SrtB<sub>H9004N26,H116A</sub> and Cd-SrtB<sub>H9004N26,R217A</sub> in addition to Cd-SrtB<sub>H9004N26,C209A</sub>. The recombinant purified Cd-SrtB<sub>H9004N26,H116A</sub> and Cd-SrtB<sub>H9004N26,R217A</sub> were subjected to FRET-based assay. To our surprise, the Cd-SrtB<sub>H9004N26,H116A</sub> mutant exhibited a comparable level of enzymatic activity as Cd-SrtB<sub>H9004N26,WT</sub>, indicating that the mutation of His-116 to alanine has no detectable effect on the catalytic activity. The Cd-SrtB<sub>H9004N26,R217A</sub> mutant exhibited significantly decreased enzymatic activity as expected (Fig. 5). It is believed that the
conserved active-site arginine residue plays a critical role in stabilizing the oxyanion transition state of the enzyme via electrostatic interactions (33). To confirm that the mutation of Arg to alanine has no effect on the global structure, we determined the crystal structure of Cd-SrtB\textsubscript{\textit{H9004N26,R217A}} to 3.1 Å resolution. Superimposition of structure of Cd-SrtB\textsubscript{\textit{H9004N26,Wt}} with that of Cd-SrtB\textsubscript{\textit{H9004N26,R217A}} revealed very small structural variation (RMSD = 0.239 for 176 C\textalpha atoms), and the structures differ mostly in the replacement of the Arg-217 side chain with Ala methyl group (Fig. 6). The results suggest that the absence of the positively charged Arg-217 guanidine group results in the disruption of the electrostatic interactions, which is essential for stabilization of the tetrahedral oxyanion intermediate.

The \textit{β6/β7} loop is the specificity determinant of Cd-SrtB

Our previous study provided a computational model of the Cd-SrtB\textsubscript{\textit{AN26 WT}}–PPKTG complex in which Ser-163 is hydrogen-bonded with the P2 lysine residue of PPKTG motif and Tyr-167 interacts with P4 proline noncovalently (40). The mutagenic study also confirmed that the abolishment of these specific interactions affected the cleavage activity of Cd-SrtB, confirming that Ser-163 and Tyr-167 play important roles in specific substrate-binding (40). The Ser-163 and Tyr-167 residues are located in the \textit{β6/β7} loop, which has been demonstrated to play a crucial role in substrate specificity between class A and B sortases (45–47). Larger structural variation in \textit{β6/β7} loop between SrtA and SrtB is observed than the differences among SrtB from different bacteria (33, 48). Whether the \textit{β6/β7} loop of bacteria-specific SrtB can also discriminate its specific motif has not yet been studied.

To assess the role of the \textit{β6/β7} loop on substrate recognition in \textit{C. difficile}, a loop swap mutant, designated as Cd-SrtB\textsubscript{\textit{AN26 LS}}, was generated by replacing residues Ser-163–Leu-168 (Ser-163–Asp-164–Tyr-165–Asp166–Tyr-167–Leu-168) with the corresponding residues Thr-177–Ile-182 (Thr-177–Lys-178–Asp-179–Asn-180–Tyr-181–Ile-182) from the Sa-SrtB\textsubscript{\textit{AN26 WT}} (Fig. 7). When the \textit{β6/β7} loop was
swapped in Cd-SrtB\textsubscript{ΔN26,Wt}, Cd-SrtB\textsubscript{ΔN26,LS} changed the substrate recognition profile from PPKTG to NPQTN (Fig. 8). These data demonstrated that the β6/β7 loop is the specificity determinant for both class-specific and bacteria-specific sortases.

Discussion

To the best of our knowledge, this work is the first report that reveals an essential serine residue located near the active site of Cd-SrtB contributing to the catalytic activity of Cd-SrtB. In addition, we demonstrated that the corresponding serine residue in Sa-SrtB also participates in the Sa-SrtB–catalyzed cleavage activity, suggesting that the role of the conserved serine residue in SrtB among Gram-positive bacteria is indispensable. Furthermore, we also demonstrated the β6/β7 loop of Cd-SrtB is the specificity determinant for substrate recognition of PPKTG motif, revealing that the β6/β7 loop governs the molecular interactions for class-specific and bacteria-specific motif recognition.
Our LC-MS/MS analysis confirmed the substrate specificity of Cd-SrtBΔN26,WT and showed that Cd-SrtBΔN26,C209A is catalytically inactive (Fig. 2), indicating that Ser-207 is not a nucleophilic residue. In addition, the location of Ser-207 is distant from substrate-binding pocket based on the structural models of Sa-SrtB–NPQT* (33) and Cd-SrtB–PPTKG (40) (Fig. S3), suggesting that Ser-207 does not directly interact with substrate peptide. Furthermore, crystallographic studies of the Cd-SrtBΔN26,S207A mutant suggest that Ser-207 does not play a structural role in Cd-SrtB (Fig. 3). Taken together, Ser-207 may be involved in stabilizing a thioacyl-enzyme intermediate during catalytic process.

Compared with SrtA structure, SrtB contains additional helices at the N terminus and an additional α-helix in the β6/β7 loop (23, 33). The β6/β7 loop in Sa-SrtA undergoes significant conformational change that the disordered β6/β7 loop transits to an ordered state upon substrate binding (47). In contrast, the β6/β7 loop in Ba-SrtA and Sa-SrtB forms a well-defined binding pocket for substrates (33). In addition, the replacement of the β6/β7 loop in Sa-SrtA with that of Sa-SrtB shifts the specificity profile of Sa-SrtA to Sa-SrtB, demonstrating that the β6/β7 loop plays a major role in distinguishing the sorting signal between class A and class B sortases (45). Nevertheless, whether the β6/β7 loop is also the specificity determinant for the same class of sortase enzymes has not been investigated. Our results established that the β6/β7 loop also plays a dominant role in recognizing the specific sorting signal of bacteria-specific SrtB enzymes by loop-swapping mutagenesis. The loop-swapped mutant Cd-SrtBΔN26,L3 is able to recognize the Sa-SrtB–specific NPQTN motif instead of the cognate PPKTG motif (Fig. 8). It is concluded here that the β6/β7 loop in sortase enzymes is the specificity determinant for the class-specific and bacteria-specific sorting motif.

The mutagenesis studies on His–Cys–Arg triad of SrtB from *B. anthracis*, *S. aureus*, and *S. pyogenes* have demonstrated the essential roles of the catalytic residues (19, 33, 39, 44). Studies have also reported that other residues are involved in catalysis of sortases (19, 23, 38, 39). Trp-194 in Sa-SrtA was shown to assist the thiolate-imidazolium ion-pair formation in active site (19). Asp-196 and Asp-234 were also shown to participate in catalytic site of Sa-SrtB and Ba-SrtB, respectively (39). The discovery of the catalytically essential serine residue near the active site led us to revisit the His–Cys–Arg triad in Cd-SrtB. Our results showed that mutation of Arg-217 significantly decreased the enzymatic activity of Cd-SrtB. The arginine residue in catalytic triad is proposed to be essential for stabilizing the oxyanion intermediate by hydrogen bonding and to facilitate catalysis (33). Structural studies and computational modeling of the Sa-SrtB–NPQT* and Ba-SrtA–LPAT* complexes showed that the active site arginine is hydrogen-bonded with P1 threonine (33). Our crystallographic study on Cd-SrtBΔN26,R217A showed that the major structural difference between Cd-SrtBΔN26,WT and Cd-SrtBΔN26,R217A is the side chain at position 217. This may explain how the absence of the arginine guanidino group abolishes the hydrogen-bonding interaction essential for intermediate stabilization. Surprisingly, point mutation on His-116 of Cd-SrtB did not affect the catalytic activity in our study. The corresponding histidine residue in Sa-SrtB, His-130, has been demonstrated to be catalytically essential, because H130A mutant enzyme exhibited no detectable activity (35). Various roles of the catalytically essential histidine residue have been proposed. It was originally thought that the histidine residue activates the nucleophilic cysteine to form a histidine–cysteine ion pair (49). It is now believed that the histidine functions as a general acid/base (36, 37). Our results suggest the possibility that Cd-SrtB employs a different functional residue in the active site.

In summary, our work defined key residues essential for Cd-SrtB catalysis and substrate recognition. Our studies provide information that may be useful for developing therapeutic strategies against CDI by manipulating the actions of Cd-SrtB without disrupting the beneficial bacteria in intestinal flora.

**Experimental procedures**

**Site-directed mutagenesis**

The primers listed in Table S1 were used to introduce a desired mutation in dsDNA by an overlapping and a back-to-back orientation. KOD FX DNA polymerase (Toyobo), the high-fidelity DNA polymerase enzyme, was used in PCR to prevent polymerase errors during PCR. Subsequently, the reactions were performed in the PCR machine (GeneAmp PCR system 2400, PerkinElmer), and the PCR products were treated by DpnI restriction enzyme (20 units/μl; New England Biolabs) to
digested methylation DNA from original DNA template and incubated at 37 °C for 2 h. Finally, all reaction products were directly transformed into *Escherichia coli* DH5α and confirmed by DNA sequencing.

**Protein overexpression and purification**

The gene encoding Cd-SrtB was cloned into pMCSG7 vector with an N-terminal His<sub>6</sub> tag described in our previous work (50). Residues 2–27 within the transmembrane domain of Cd-SrtB were deleted to improve solubility of the recombinant protein, designated as Cd-SrtB<sub>AN26,WT</sub>. The gene encoding Sa-SrtB amplifying from chromosomal DNA of *S. aureus* was cloned into pET21b<sup>+</sup> vector with a C-terminal His<sub>6</sub> tag. Residues 2–27 within the transmembrane domain of Sa-SrtB and Cd-SrtB were also deleted to improve solubility of the recombinant protein, designated as Sa-SrtB<sub>AN29,WT</sub> and Cd-SrtB<sub>AN26,WT</sub>. The genes encoding sortase mutants generated by site-directed mutagenesis were cloned to pMCSG7 and pET21b<sup>+</sup> for Cd-SrtB mutants and Sa-SrtB mutants, respectively. All plasmids were transformed into *E. coli* BL21(DE3). The overexpression of sortases were induced by adding 0.5 mM isopropyl-β-d-thiogalactopyranoside when cells density reached an A<sub>600</sub> of 0.5–0.6, at the temperatures of 37 °C (for Cd-SrtB<sub>AN26,WT</sub> and mutants) and 25 °C (for Sa-SrtB<sub>AN29,WT</sub> and mutants) and incubated for additional 4 h. The cell pellets were resuspended in lysis buffer (Cd-SrtB<sub>AN26,WT</sub> and mutants: 20 mM HEPES, pH 7.4, 200 mM NaCl, and 20 mM imidazole; Sa-SrtB<sub>AN29,WT</sub> and mutants: 20 mM HEPES, pH 8.0, 200 mM NaCl, and 20 mM imidazole), and disrupted by sonicator (digital sonifier; Branson). The crude extracts were harvested and centrifuged at 4 °C and 13,000 rpm for 30 min by using HITACHI CR22GiII with R20A2 rotor. Then the culture supernatant was filtered through 0.45-μm and 0.22-μm polyvinylidene difluoride membranes (Millipore). The mutants of recombinant *C. difficile* His<sub>6</sub>-tagged SrtB<sub>AN26 </sub>and C209A C. difficile His<sub>6</sub>-tagged SrtB<sub>AN26</sub>, R217A C. difficile His<sub>6</sub>-tagged SrtB<sub>AN26</sub>, H116A and S. aureus His<sub>6</sub>-tagged PPKTG- and NPQTN-containing substrate peptides for Cd-SrtB<sub>AN26,WT</sub>/Sa-SrtB<sub>AN29,WT</sub> and the reaction mixtures were subjected to LC-MS/MS studies. The peptide mixtures were desalted by C18 Zip-tip (Millipore, Bedford, MA) and evaporated to dryness using a SpeedVac. Dried peptides were dissolved in 5% acetonitrile and 0.1% formic acid, and 5 μL of the solution was loaded onto a manually packed precolumn (150-μm inner diameter × 30 mm, 5 μm, 200 Å) at a 10 μL/min flow rate. The peptides were analyzed with a 7-Tesla LTQ-FT Ultra mass spectrometer (linear quadrupole ion trap Fourier transform ion cyclotron resonance; Thermo Scientific, San Jose, CA) by using a Finnigan Series binary HPLC pump (Acquity Technologies, Palo Alto, CA), and a Famos autosampler (LC Packing, San Francisco, CA). Chromatographic separation was performed over 60 min on a manually packed reversed phase C18 nanocolumn (75-μm inner diameter × 200 mm, 3 μm, 200 Å) using 0.1% formic acid in water as mobile phase A, 0.1% formic acid in 80% acetonitrile as mobile phase B, and a split flow rate of 300 nl/min. The full-scan mass range was set from m/z 320 to 2000 with 100,000 resolution at m/z = 400. The top five most intense ions were sequentially isolated for CID MS/MS fragmentation and detection in the linear ion trap (AGC target at 10,000) with previously selected ions dynamically excluded for 15 s. Ions with
singly and unrecognized charge state were also excluded. The electrospray voltage was maintained at 1.7 kV, and the capillary temperature was set to 200 °C.

All MS and MS/MS raw data were processed with Proteome Discoverer version 2.3 (Thermo Scientific), and the peptides were identified from the MS/MS data searched against the target substrate peptide sequence (PVPKPTGDSTTIIG) database using the Mascot search engine 2.6.2 (Matrix Science). Searches were limited to peptide mass tolerance of ± 1.0 Da and MS/MS ion mass tolerance of ± 1.0 Da. The variable modifications considered were N-terminal proline Dabcyl modification (peptides molecular + 252.1 Da, C15H12NO3S) and C-terminal glutamic acid Edans modification (peptides molecular + 250.05 Da, C14H22NO4S). The significant peptide hits defined as peptide score must be higher than Mascot significance threshold (p < 0.05) and therefore considered reliable, and that manual interpretation confirmed agreement between spectra and peptide sequence. The false discovery rate of the peptides and protein groups was set to 1% for the MS/MS spectra automatically quantification.

**Table 1**

| Structural data and refinement statistics |
|------------------------------------------|
| **Cd-SrtBΔN26,S207A** | **Cd-SrtBΔN26,R217A** |
| Data collection | | |
| Space group | I23 | 123 |
| Cell dimensions (Å) | a, b, c | 120.89, 120.89, 120.89 | 120.97, 120.97, 120.97 |
| Resolution (Å) | 30–2.6 (2.7–2.6) | 30–2.6 (2.7–2.6) |
| Wavelength (Å) | 1.0 | 1.0 |
| Rmerge (%) | 6.7 (35.3) | 5.2 (71.3) |
| I/I0 | 29.7 (5.1) | 31.7 (2.0) |
| Completeness (%) | 94.1 (100.0) | 99.4 (98.7) |
| Redundancy | 6.9 (7.3) | 7.1 (5.3) |
| Wilson B factor (Å²) | 75.76 | 41.32 |
| Refinement | | |
| No. reflections | 5100 | 8961 |
| Rwork/Rfree | 22.17/25.53 | 21.5/24.15 |
| B factors | | |
| Protein | 71.0 | 44.41 |
| RMSD(s) | | |
| Bond lengths (Å) | 0.007 | 0.002 |
| Bond angles (°) | 0.86 | 0.44 |
| Ramachandran plot statistics (%) | | |
| Favorable regions | 93.22 | 96.05 |
| Allowed regions | 6.78 | 3.95 |
| Outlier regions | 0 | 0 |

**Functional residues of C. difficile sortase B**

The crystallographic data of Cd-SrtBΔN26,S207A and Cd-SrtBΔN26,R217A proteins were crystallized in the same space group I23. The unit cell parameters were a = b = c = 120.969 Å and α = β = γ = 90° for Cd-SrtBΔN26,S207A and a = b = c = 120.895 Å and α = β = γ = 90° for Cd-SrtBΔN26,R217A. All refinement statistics of SrbB mutants were listed in Table 1. The structures of Cd-SrtBΔN26,S207A and Cd-SrtBΔN26,R217A were solved by molecular replacement by using the program of Cd-SrtBΔN26 (PDB code 5GYJ) (40) as a search model. The manual model rebuilding was performed by using COOT (52) with the guidance of 2Fo – Fc and Fo – Fc density maps. The iterative refinement was performed by using program CCP4 (53) and PHENIX (54). The structures of Cd-SrtB mutants S207A and R217A were solved at resolutions of 2.6 and 3.1 Å, respectively. The refinement statistics are shown in Table 1. Coordinates and structure factors with the identifier (PDB code 6KYD-R217A) and (PDB code 6KYC-S207A) have been deposited in the PDB.

**Author contributions**—C.-Y. K., C.-C. C., Y.-Y. H., I.-H. H., J.-C. C., Y.-H. K., and S. W. conceptualization; I.-H. H. and S. W. supervision; C.-Y. K., I.-H. H., J.-C. C., Y.-Y. H., and S. W. formal analysis; C.-Y. K., I.-H. H., J.-C. C., Y.-H. K., and S. W. writing-original draft; I.-H. H. and S. W. conceptualization; I.-H. H. and S. W. supervision; I.-H. H. and S. W. project administration; I.-H. H., C.-C. C., Y.-Y. H., and S. W. writing-review and editing; S. W. funding acquisition.

**Acknowledgments**—We thank the technical services provided by the Synchrotron Radiation Protein Crystallography Facility of the National Core Facility Program for Biotechnology, Ministry of Science and Technology, and the National Synchrotron Radiation Research Center, a national user facility supported by the Ministry of Science and Technology, Taiwan.

**References**

1. Peery, A. F., Dellon, E. S., Lund, J., Crockett, S. D., McGowan, C. E., Bulsiewicz, W. J., Gangarosa, L. M., Thiny, M. T., Stizenberg, K., Morgan, D. R., Ringel, Y., Kim, H. P., DiBonaventura, M. D., Carroll, C. F., Allen, J. K., et al. (2012) Burden of gastrointestinal disease in the United States: 2012 update. *Gastroenterology* 143, 1179–1187.e3 CrossRef Medline

2. Asemota, T. E., and Nicolau, D. P. (2017) *Clostridium difficile* infection in the elderly: an update on management. *Clin. Interv. Aging* 12, 1799–1809 CrossRef Medline

3. Garey, K. W., Sethi, S., Yadav, Y., and DuPont, H. L. (2008) Meta-analysis to assess risk factors for recurrent *Clostridium difficile* infection. *J. Hosp. Infect.* 70, 298–304 CrossRef Medline

4. Honda, H., Yamazaki, A., Sato, Y., and Dubberke, E. R. (2014) Incidence and mortality associated with *Clostridium difficile* infection at a Japanese tertiary care center. *Anaerobe* 25, 5–10 CrossRef Medline
40. Yin, J. C., Fei, C. H., Lo, Y. C., Hsiao, Y. Y., Chang, J. C., Nix, J. C., Chang, Y. Y., Yang, L. W., Huang, I. H., and Wang, S. (2016) Structural insights into substrate recognition by Clostridium difficile sortase. Front. Cell Infect. Microbiol. 6, 160 PubMed
41. Chambers, C. J., Roberts, A. K., Shone, C. C., and Acharya, K. R. (2015) Structure and function of a Clostridium difficile sortase enzyme. Sci. Rep. 5, 9449 PubMed
42. Cascioferro, S., Totsika, M., and Schillaci, D. (2014) Sortase A: an ideal target for anti-virulence drug development. Microb. Pathog. 77, 105–112 PubMed
43. 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 PubMed
44. 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 PubMed
45. Bentley, M. L., Gaweska, H., Kielec, J. M., and McCafferty, D. G. (2007) Engineering the substrate specificity of Staphylococcus aureus Sortase A. The B6/B7 loop from SrtB confers NPQTN recognition to SrtA. J. Biol. Chem. 282, 6571–6581 PubMed
46. 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 PubMed
47. 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 PubMed
48. 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 PubMed
49. 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 PubMed
50. Aslanidis, C., and de Jong, P. J. (1990) Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res. 18, 6069–6074 PubMed
51. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326 PubMed
52. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 PubMed
53. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 50, 760–763 PubMed
54. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. I., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. 58, 1948–1954 PubMed

**Functional residues of C. difficile sortase B**

48. 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 PubMed
49. 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 PubMed
50. Aslanidis, C., and de Jong, P. J. (1990) Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res. 18, 6069–6074 PubMed
51. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326 PubMed
52. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 PubMed
53. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 50, 760–763 PubMed
54. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. I., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. 58, 1948–1954 PubMed

**J. Biol. Chem. (2020) 295(11) 3734–3745 3745**