In vitro reconstitution of sortase-catalyzed pilus polymerization reveals structural elements involved in pilin crosslinking

Chungyu Chang[a,1], Brendan R. Amer[b,c1], Jerzy Osipiuk[d,e], Scott A. McConnell[b,c], I-Hsiu Huang[f], Van Hsieh[b,c], Janine Fu[b,c], Hong H. Nguyen[b,c], John Muroski[b,c], Erika Flores[a], Rachel R. Ogorzalek Loo[b,c], Joseph A. Loo[b,c], John A. Putkey[g], Andrzej Joachimiak[d,e], Asis Das[h], Robert T. Clubb [b,c,2], and Hung Ton-That[a,2]

[a]*Department of Microbiology & Molecular Genetics, University of Texas Health Science Center, Houston, TX 77030;* [b]*Department of Chemistry and Biochemistry*, *University of California, Los Angeles, CA 90095;* [c]*University of California, Los Angeles-US Department of Energy Institute of Genomics and Proteomics, University of California, Los Angeles, CA 90095;* [d]*Center for Structural Genomics of Infectious Diseases, Argonne National Laboratory, Argonne, IL 60439;* [e]*Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637;* [f]*Department of Microbiology and Immunology, College of Medicine, National Cheng Kung University, Tainan 701, Taiwan;* [g]*Department of Biochemistry and Molecular Biology, University of Texas Health Science Center, Houston, TX 77030; and* [h]*Department of Molecular Biology and Biophysics, University of Connecticut Health Center, Farmington, CT 06030*

[1]Contributed equally to this work

[2]To whom correspondence should be addressed:

Robert T. Clubb
Department of Chemistry and Biochemistry, University of California, Los Angeles, 611 Charles Young Drive East, Los Angeles, CA 90095, USA
Tel. (+1) 310 206 2334; Fax (+1) 310 206 4779; Email: rclubb@mbi.ucla.edu

Hung Ton-That
Department of Microbiology and Molecular Genetics, University of Texas Health Science Center, 6431 Fannin Street, R224/MSE, Houston, TX 77030, USA.
Tel. (+1) 713 500 5468; Fax (+1) 713 500 5499; E-mail: ton-that.hung@uth.tmc.edu

Running title: *In vitro assembly of Gram-positive pilus polymers*

Keywords: *Corynebacterium diphtheriae*; sortase; pilus polymerization; protein ligation; transpeptidation

SUPPORTING INFORMATION

SI Materials and Methods

Recombinant Plasmids

pSUMO-SrtA^{WT} and its derivatives – To generate a recombinant plasmid expressing His-tagged SrtA^{WT} (residues 37 to 257), the *srtA* gene sequence without N-terminal signal peptide and C-terminal membrane spanning domains was PCR-amplified from the genomic DNA of *C. diphtheriae* NCTC 13129 with appropriate primers (Table S2) and inserted into the pE-SUMO (LifeSensors) expression vector using the Gibson assembly method (New England BioLabs). pSUMO-SrtA^{WT} was used as a template to generate D81A and W83A mutations (pSUMO-SrtA^{2M}), as well as Y225A, S229A or N228A mutation, using site-directed mutagenesis carried out by QuickChange method (Agilent) (see Table S2 for primers). Resulting plasmids were then transformed into XL10 for amplification prior to DNA sequence confirmation. Similarly, pSUMO-$^{\Delta}$SrtA^{2M} were generated using pSUMO-SrtA2M as a template, in which H1 helix (residues 37-54) was removed. The resulting plasmid was introduced into *E. coli* BL21 (DE3) after verification by DNA sequencing.

pMCSG-SrtA^{WT} and its derivatives – For protein crystallization, the same *srtA* fragment as the above was cloned into the pMCSG7 expression vector by ligation-independent cloning (LIC) as previously reported (1). The resulting plasmid was introduced into *E. coli* DH5α for selection and DNA sequencing, and then *E. coli* BL21 (DE3). To generate pMCSG-SrtA^{2M}, pMCSG-SrtA^{WT} was used as a template for inverse PCR amplification with a pair of phosphorylated primers carrying the intended mutation (Table S1) as previously described (1). The resulting linear PCR product was ligated before introduced into *E. coli* DH5α. The generated plasmid was verified by DNA sequencing prior to introduce into *E. coli* BL21 (DE3).

Recombinant SpaA plasmids – A plasmid expressing recombinant His-tagged SpaA protein of *C. diphtheriae* lacking the N-terminal signal peptide and C-terminal transmembrane domain (residues 30 to 500) was generated using the LIC method described above (see Table S2 for primers). The resulting plasmid (pMCSG-SpaA) was introduced into *E. coli* BL21 (DE3) after verification by DNA sequencing. Using the above site-directed mutagenesis method, pMCSG-SpaA was then used as a template to generate pMCSG-SpaA^{K190A} that expresses the same SpaA molecule with lysine 190 replaced by alanine. To generate plasmids pSUMO-^{N}SpaA and pMBP-^{C}SpaA, which express the N-terminal (residues 30 to 194) and C-terminal (residues 350 to 500) domains of SpaA, pE-SUMO and pE-MAPLE were used, respectively, in the same cloning protocol as described for pSUMO-SrtA^{WT} above.

Recombinant SpaB plasmids – Similar to the construction of the recombinant SpaA plasmids, pMCSG was used to clone a recombinant SpaB protein (residues 25 to 180), lacking the N-terminal signal peptide and the C-terminal hydrophobic domain and the charged tail. This generated plasmid was used as a template for site-directed mutagenesis to produce the SpaB-K139A mutant protein.

pSrtA and its derivatives – For *srtA* expression in *C. diphtheriae*, the *E. coli*/*Corynebacterium* shuttle vector pCGL0243 was used (2). pSrtA (3), a pCGL0243 derivative that expresses *C. diphtheriae srtA* under control of the *spaA* promoter, served as a template for site-directed mutagenesis, as described above, to generate various SrtA variants used in this study (Table S1 and see Table S2 for primers). The resulting plasmids were introduced into *E. coli* DH5α for DNA sequencing prior to electroporation into *C. diphtheriae* strains.

Protein Expression and Purification

For in vitro pilus polymerization, His-tagged proteins were purified according to a published procedure (4). Briefly, *E. coli* BL21 (DE3) cells harboring pSUMO-SrtA^{WT}, pSUMO-SrtA2M, pMCSG-SpaA, pMCSG-SpaA^{K190A}, pSUMO-^{N}SpaA, or pMBP-^{C}SpaA were grown in LB supplemented with ampicillin at 100 µg ml^{-1} at 37^{0}C until OD_{600} of ~ 0.6. Cells were equilibrated to 17°C and treated with 1 mM IPTG before they were allowed to grow overnight at 17°C to induce protein expression. Cells were then harvested by centrifugation (8,000 RPM for 20 min) and stored at -80°C for further processing. SrtA and SpaA-derived proteins were purified as a His6x-SUMO- fusion using HisPure Co^{2+} IMAC resin (Thermo) per the manufacturer's instructions. Briefly, cell pellets were resuspended in 50 mM Tris-HCl pH 8.0, 300 mM NaCl, and 5 mM $CaCl_2$ (lysis buffer) and lysed by sonication. Subsequent cell lysate was then fractionated by centrifugation (15,000 RPM for 1 hr) and the supernatant was loaded onto HisPure Co^{2+} IMAC resin. Proteins were then eluted from the resin using lysis buffer supplemented with 200 mM Imidazole. The His6x-SUMO tag was removed by the addition of His6x-Ulp1 protease, and subsequent HisPure Co^{2+} purification. Protein purity was determined by SDS-PAGE analysis.

For crystallization, recombinant proteins were purified according to a published procedure (5). Briefly, *E. coli* BL21 (DE3) cells harboring pMCSG-SrtA^{WT} or pMCSG-SrtA^{2M} was cultured in 2x YT broth containing ampicillin at 100 µg ml^{-1} at 37°C with shaking until OD_{600} of ~ 1.0. The culture was later induced with 0.4 mM IPTG and allowed to grow overnight at 18°C with shaking. Cells were harvested and disrupted by sonication. The lysate containing SrtA was purified by using Ni-NTA (Qiagen) affinity chromatography with the addition of 5 mM β-mercaptoethanol in all buffers. The N-terminal His-tag and TEV restriction sequence of the protein was removed by the TEV protease (0.15 mg for 20 mg purified protein) incubated for 16 h at 4°C, and then passed through a Ni-NTA column to remove both the TEV protease and cleaved N-terminal tags. The final step of purification was gel-filtration on HiLoad 16/60

Superdex 200pg column (GE Healthcare) in crystallization buffer 10 mM HEPES buffer pH 7.5, 200 mM NaCl and 1 mM DTT. The protein was concentrated on Amicon Ultracel 10K centrifugal filters (Millipore) up to 24 mg/ml concentration.

Protein crystallization, Data Collection, Structure Determination and Refinement

The initial crystallization condition was determined with a sparse crystallization matrix at 4°C and 16°C temperatures using the sitting-drop vapor-diffusion technique as reported (5). The best crystallization condition for the $SrtA^{WT}$ protein was found in 0.1 M MES:NaOH buffer pH 6.5, 1.6 M ammonium sulfate and 10% dioxane at 4°C. The SrtA2M protein was crystallized in 0.2 M sodium chloride and 20% PEG 3350 at 4°C. The $SrtA^{WT}$ and $SrtA^{2M}$ protein crystals selected for data collection were soaked in the crystallization buffer supplemented with either 28% sucrose or 25% glycerol, respectively, and flash-cooled in liquid nitrogen.

Single-wavelength X-ray diffraction data were collected at 100 K temperature at the 19-ID beamline using the program SBCcollect. The intensities were integrated and scaled with the HKL3000 suite (6). The SrtA structures were determined by molecular replacement using the HKL3000 suite incorporating following programs: MOLREP (7), SOLVE/RESOLVE (8), and ARP/wARP (9). The coordinates for the *A. oris* sortase SrtC-1 (10) (PDB:2XWG) were used as the starting model for the $SrtA^{WT}$ structure. Several rounds of manual adjustments of structure models using COOT (11) and refinements with Refmac program (12) from CCP4 suite (13) were performed. The stereochemistry of the structure was validated with PHENIX suite (6) incorporating MOLPROBITY(14) tools. Data collection and refinement statistics are summarized in Tables 1 and 2. Atomic coordinates and structure factors of $SrtA^{WT}$ and $SrtA^{2M}$ were deposited into the Protein Data Bank as 5K9A and 6BWE, respectively.

Nuclear magnetic resonance (NMR) analysis of $SrtA^{2M}$ and $^{\Delta}SrtA^{2M}$ sortases

^{1}H-^{15}N HSQC NMR spectra of $SrtA^{2M}$ and $^{\Delta}SrtA^{2M}$ were acquired using the following samples: 150 μM [^{15}N] $^{\Delta}SrtA^{2M}$ and 500 μM [^{15}N] $SrtA^{2M}$ proteins dissolved in 50 mM sodium phosphate, 100 mM sodium chloride, 0.01% sodium azide, 5% D_2O, pH 6.5. NMR spectra were acquired at 298 K using a Bruker Avance 600 MHz spectrometer equipped with a triple resonance cryogenic probe. A total of 16 and 32 scans were used to acquire the $SrtA^{2M}$ and $^{\Delta}SrtA^{2M}$ spectra, respectively.

SI Figure Legends

Figure S1: Visualization of *C. diphtheriae* pili assembled by the SrtA^{2M} enzyme. Corynebacterial cells of indicated strains were immobilized on nickel grids and subjected to negative staining with 0.75% uranyl formate (A) or immune-gold labeling with antibodies against SpaA prior to negative staining (B). Scale bars in (A) indicate 100 nm; (B), 200 nm.

Figure S2: SrtA containing the wild type lid is inactive in catalyzing crosslinking of the isolated domains. Fusion proteins between SUMO and the N-terminal SpaA domain (^{N}SpaA; residues 30 to 194) and between maltose-binding protein MBP and the C-terminal SpaA domain (^{C}SpaA; residues 350 to 500) were reacted with either the SrtA^{WT} or SrtA^{2M} enzyme at a 3:1 ratio, respectively. The reaction samples were analyzed by SDS-PAGE and Coomassie staining after 0, 24 h, 48 h, and 72 h of incubation.

Figure S3: Sequence alignment of Gram-positive sortases showing the class C signature sequence. Shown is an alignment of amino acids encompassing strands β7 to β8 of class C sortases. All aligned class C enzymes have been experimentally demonstrated to assemble pili by either cellular or biochemical methods. The conserved TP(Y/L)XIN(S/T)H signature sequence in class C enzymes is shaded in light blue, while the catalytic cysteine and arginine residues are colored red. The bottom of the figure shows representative class A, B, D and E enzymes sortases that do not assemble pili, but instead attach proteins to the cell wall. The GenBank accession codes are as follows: WP_010935503 (*Corynebacterium diphtheriae* SrtA); WP_010934130 (*Corynebacterium diphtheriae* SrtB); WP_010934133 (*Corynebacterium diphtheriae* SrtC); WP_010935679 (*Corynebacterium diphtheriae*, SrtD); WP_010935678 (*Corynebacterium diphtheriae*, SrtE); WP_002307920 (*Enterococcus faecium*, SrtC); WP_014569086 (*Lactobacillus rhamnosus*, SrtC1); WP_060958109 (*Actinomyces oris*, SrtC1); WP_060956887 (*Actinomyces oris*, SrtC2); WP_000047114 (*Streptococcus pneumoniae*, SrtC1); WP_050148456

(*Streptococcus agalactiae*, SrtC1); WP_000746885 (*Streptococcus agalactiae*, SrtC2); WP_000828081 (*Bacillus cereus*, SrtD); WP_037276992 (*Ruminococcus albus*, SrtC).

Figure S4: Comparison of the ^{1}H-^{15}N HSQC spectra of SrtA^{2M} and $^{\Delta}$SrtA^{2M}. The ^{1}H-^{15}N HSQC spectrum of SrtA^{2M} is shown (red). The expanded images are selected regions from the HSQC spectrum overlaid with the corresponding spectrum of $^{\Delta}$SrtA^{2M} (blue). The cross-peaks originate from backbone amides located within structured portions of each protein based on their down-field ^{1}H chemical shifts. Similar chemical shifts are observed indicating that the proteins adopt similar atomic structures.

SI Tables

Table S1: Bacterial strains and plasmids used

Strain & Plasmid	Description	Reference
Strain		
NCTC 13129	A *C. diphtheriae* type strain	(15)
HT2	Δ*srtA*; an isogenic derivative of NCTC 13129	(15)
HT52	Δ*spaA*/Δ*srtA*; an isogenic derivative of NCTC 13129	(3)
Plasmid		
pCGL0243	*C. diphtheriae*/*E. coli* shuttle vector, Kan^R	(16)
pSrtA	A pCGL0243 derivative expressing *C. diphtheriae* SrtA under control of the *spaA* promoter	(3)
$pSrtA^{2M}$	A pSrtA derivative, expressing SrtA with D81G and W83G mutations	This study
$pSrtA^{D81A}$	A pSrtA derivative, expressing SrtA with D81G mutation	This study
$pSrtA^{W83A}$	A pSrtA derivative, expressing SrtA with W83A mutation	This study
$pSrtA^{DW2A}$	A pSrtA derivative expressing SrtA with D81A and W83A mutations	This study
$pSrtA^{C222A}$	A pSrtA derivative expressing SrtA with C222A mutation	This study
$pSrtA^{H160A}$	A pSrtA derivative expressing SrtA with H160A mutation	This study
$pSrtA^{R231A}$	A pSrtA derivative expressing SrtA with R231A mutation	This study
pSpaA-SrtA	A pCGL0243 derivative expressing *C. diphtheriae* SpaA and wild type SrtA under control of the *spaA* promoter	This study
pSpaA-$SrtA^{2M}$	A pSpaA-SrtA derivative expressing *C. diphtheriae* SpaA and $SrtA^{2M}$ under control of the *spaA* promote	This study
$pSpaA^{K190A}$-$SrtA^{2M}$	A pSpaA-$SrtA^{2M}$ derivative with K190A mutation in SpaA	This study
pMCSG7	Ligation-independent cloning vector; Amp^R	(1)
pMCSG-SrtA	*srtA* gene encoding a.a. 37 to 257cloned into pMCSG7	This study
pMCSG-$SrtA^{2M}$	A pMCSG7 derivative expressing His-tagged SrtA with D81G and W83G mutations	This study
pMCSG-SpaA	A pMCSG7 derivative expressing His-tagged SpaA (residues 30 to 500)	This study
pMCSG-$SpaA^{K190A}$	A pMCSG-SpaA derivative expressing SpaA with K190A mutation	This study
pMCSG-SpaB	A pMCSG7 derivative expressing His-tagged SpaB (residues 25 to 180)	This study

pMCSG-SpaB^{K139A}	A pMCSG-SpaB derivative expressing SpaB with K139A mutation	This study
pE-SUMO	Expression vector	LifeSensors
pSUMO-SrtA^{WT}	*srtA* gene encoding a.a. 37 to 257 cloned into pE-SUMO	This study
pSUMO-SrtA^{2M}	Same as pSUMO-SrtA^{WT}, but carrying D81G and W83G mutations	This study
pSUMO-$^{\Delta}$SrtA^{2M}	The gene sequence encoding H1 helix (a.a. 37-54) was deleted from pSUMO-SrtA^{2M}	This study
pSUMO-SrtA^{Y225A}	Same as pSUMO-SrtA^{2M}, but carrying Y225A mutation	This study
pSUMO-SrtA^{N228A}	Same as pSUMO-SrtA^{2M}, but carrying N228A mutation	This study
pSUMO-SrtA^{S229A}	Same as pSUMO-SrtA^{2M}, but carrying S229A mutation	This study
pSUMO-^{N}SpaA	N-terminal sequence of SpaA (residues 30-194) cloned into pE-SUMO	This study
pE-MAPLE	Expression vector	
pMBP-^{C}SpaA	C-terminal sequence of SpaA (residues 350-500) cloned into pE-MAPLE	This study

Table S2: Primers used for cloning and site-directed mutagenesis (SDM)

Primer	Sequence[(a)]	Used for
SrtA-SDM-DW2G-5	cgGggctgaacagggcgtcgaaaaacag	pCGL-SrtA^{2M}; pMCSG-SrtA^{2M}; pSpaA-SrtA^{2M}
SrtA-SDM-DW2G-3	gaCcgaggatcgggaagcctactactaccg	pCGL-SrtA^{2M}; pMCSG-SrtA^{2M}; pSpaA-SrtA^{2M}
SrtA-SDM-D81A-5	gCtccgtggctgaacagggcgtc	pCGL-SrtA^{D81A}
SrtA-SDM-D81A-3	gaggatcgggaagcctactaccgac	pCGL-SrtA^{D81A}
SrtA-SDM-W83A-5	GCgctgaacagggcgtcgaaaaacag	pCGL-SrtA^{W83A}
SrtA-SDM-W83A-3	cggatcgaggatcgggaagcc	pCGL-SrtA^{W83A}
SrtA-SDM-DW2A-5	cgGggctgaacagggcgtcgaaaaacag	pCGL-SrtA^{DW2A}
SrtA-SDM-DW2A-3	gaCcgaggatcgggaagcctactaccg	pCGL-SrtA^{DW2A}
SrtA-SDM-C222A-5	GCcacccccctacgccgtcaac	pCGL-SrtA^{C222A}
SrtA-SDM-C222A-3	ggtgatgagtgttatttggtc	pCGL-SrtA^{C222A}
SrtA-SDM-H160A-5	GCcagcggccttgccaacgc	pCGL-SrtA^{H160A}
SrtA-SDM-H160A-3	tccggtgatcacggggtg	pCGL-SrtA^{H160A}
SrtA-SDM-R231A-5	ctcctcgtacgagcccaccgc	pCGL-SrtA^{R231A}
SrtA-SDM-R231A-3	TGCgtgggagttgacggcgtagg	pCGL-SrtA^{R231A}
SrtA-LIC-N37-5	tacttccaatccaatgcaaacaacgcgcgccaagcacg	pMCSG-SrtA
SrtA-LIC-Q257-3	ttatccacttccaatgttactattgccagattttggtgccgg	pMCSG-SrtA
SpaB-LIC-Q25-5	tacttccaatccaatgcacaagaagcaaacacattggtcattgacctc	pMCSG-SpaB
SpaB-LIC-G180-3	ttatccacttccaatgttactatccgaggatgcttgcgcc	pMCSG-SpaB
SpaB-SDM-K139A-5	cttatcgacgccacccccgg	pMCSG-SpaB^{K139A}
SpaB-SDM-K139A-3	cGCggggcgcagggtcatgg	pMCSG-SpaB^{K139A}
SpaA-SDM-K190A-5	GCgcaccaggctttgtctgagcc	pMCSG-SpaA^{K190A}; pSpaA^{K190A}-SrtA^{2M}
SpaA-SDM-K190A-3	gggatacacgtgcacgtcttggag	pMCSG-SpaA^{K190A}; pSpaA^{K190A}-SrtA^{2M}
SpaA-LIC-E30-5	tacttccaatccaatgcagaagagtcacagtatcatgcagtccaac	pMCSG-SpaA
SpaA-LIC-I500-3’	ttatccacttccaatgttagatgcgccccgaaccac	pMCSG-SpaA
SpaA-SDM-K190A-5	gcgcaccaggctttgtctgagcc	pMCSG-SpaA^{K190A}
SpaA-SDM-K190A-3	gggatacacgtgcacgtcttggag	pMCSG-SpaA^{K190A}
SrtA-SDM-Y225A-5	catcacctgcacccccGCGgccgtcaactcccaccg	pSUMO-SrtA^{Y225A}
SrtA-SDM-Y225A-3	cggtgggagttgacggcCGCgggggtgcaggtgatg	pSUMO-SrtA^{Y225A}
SrtA-SDM-N228A-5	cccctacgccgtcGCctcccaccgactc	pSUMO-SrtA^{N228A}
SrtA-SDM-N228A-3	gagtcggtgggagGCgacggcgtagggg	pSUMO-SrtA^{N228A}
SrtA-SDM-S229A-5	cccctacgccgtcaacGcGcaccgactcctcgtacg	pSUMO-SrtA^{S229A}

SrtA-SDM-S229A-3	cgtacgaggagtcggtgCgCgttgacggcgtagggg	pSUMO-SrtA^{S229A}

[a] Upper case letters are the mutated nucleotides for site-directed mutagenesis.

References

1. Reardon-Robinson ME*, et al.* (2015) A thiol-disulfide oxidoreductase of the Gram-positive pathogen *Corynebacterium diphtheriae* is essential for viability, pilus assembly, toxin production and virulence. *Mol Microbiol* 98(6):1037-1050.
2. Chang C, Mandlik A, Das A, & Ton-That H (2011) Cell surface display of minor pilin adhesins in the form of a simple heterodimeric assembly in *Corynebacterium diphtheriae*. *Mol Microbiol* 79(5):1236-1247.
3. Guttilla IK*, et al.* (2009) Acyl enzyme intermediates in sortase-catalyzed pilus morphogenesis in gram-positive bacteria. *J Bacteriol* 191(18):5603-5612.
4. Bornhorst JA & Falke JJ (2000) Purification of proteins using polyhistidine affinity tags. *Methods Enzymol* 326:245-254.
5. Reardon-Robinson ME*, et al.* (2015) A Disulfide Bond-forming Machine Is Linked to the Sortase-mediated Pilus Assembly Pathway in the Gram-positive Bacterium *Actinomyces oris*. *J Biol Chem* 290(35):21393-21405.
6. Adams PD*, et al.* (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58(Pt 11):1948-1954.
7. Vagin A & Teplyakov A (1997) MOLREP: an automated program for molecular replacement. *Journal of Applied Crystallography* 30:1022-1025.
8. Terwilliger TC (2003) SOLVE and RESOLVE: automated structure solution and density modification. *Methods Enzymol* 374:22-37.
9. Morris RJ, Perrakis A, & Lamzin VS (2003) ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol* 374:229-244.
10. Persson K (Structure of the sortase AcSrtC-1 from *Actinomyces oris*. *Acta Crystallogr D Biol Crystallogr* 67(Pt 3):212-217.

11. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.

12. Murshudov GN, Vagin AA, & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53(Pt 3):240-255.

13. Collaborative Computational Project N (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760-763.

14. Davis IW, Murray LW, Richardson JS, & Richardson DC (2004) MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* 32(Web Server issue):W615-619.

15. Ton-That H & Schneewind O (2003) Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol Microbiol* 50(4):1429-1438.

16. Ankri S, Reyes O, & Leblon G (1996) Electrotransformation of highly DNA-restrictive corynebacteria with synthetic DNA. *Plasmid* 35(1):62-66.

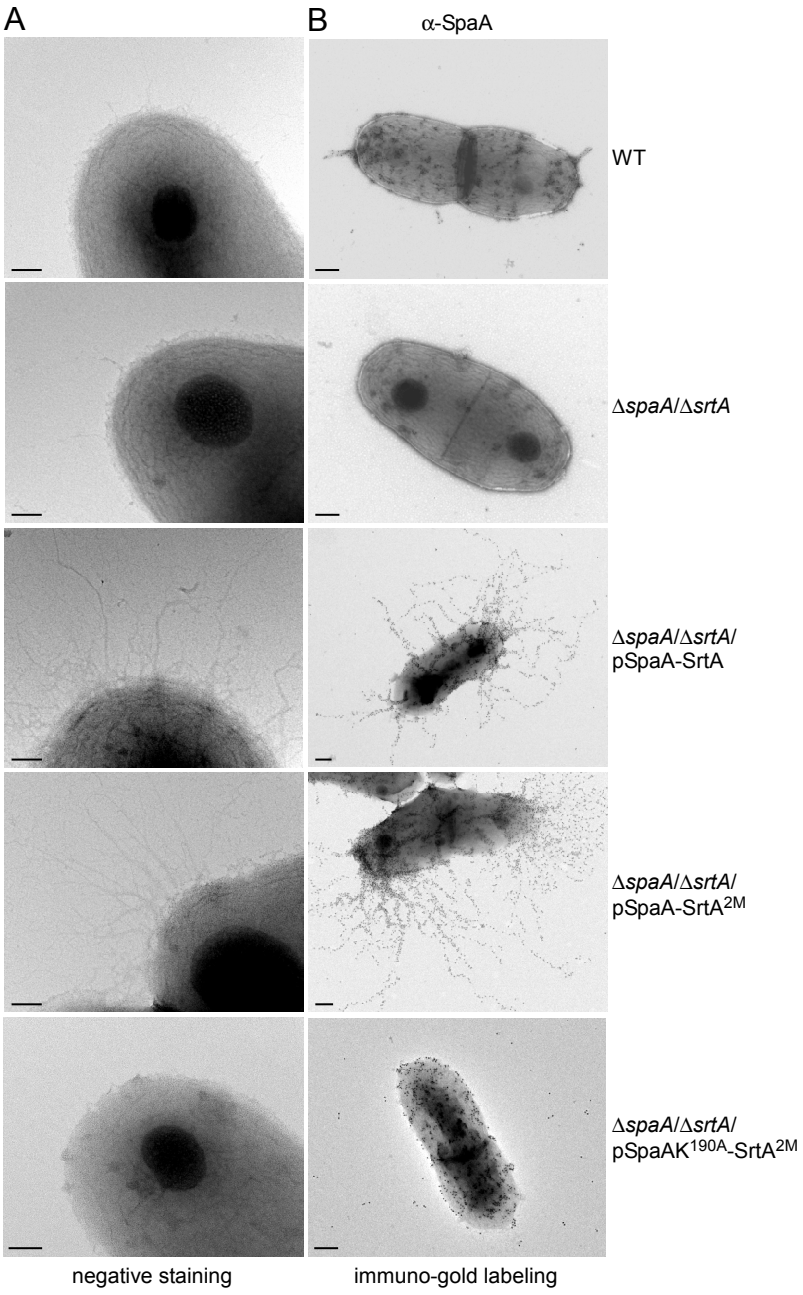

Figure S1: Chang et al.

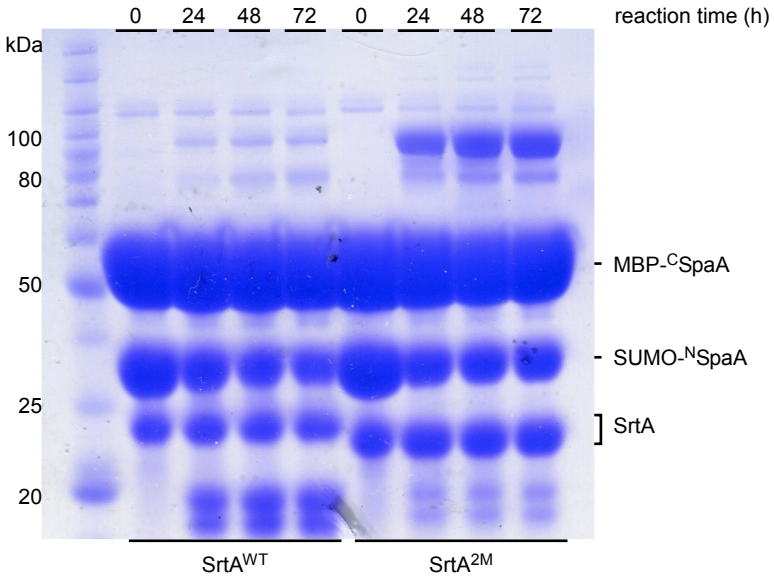

Figure S2: Chang et al.

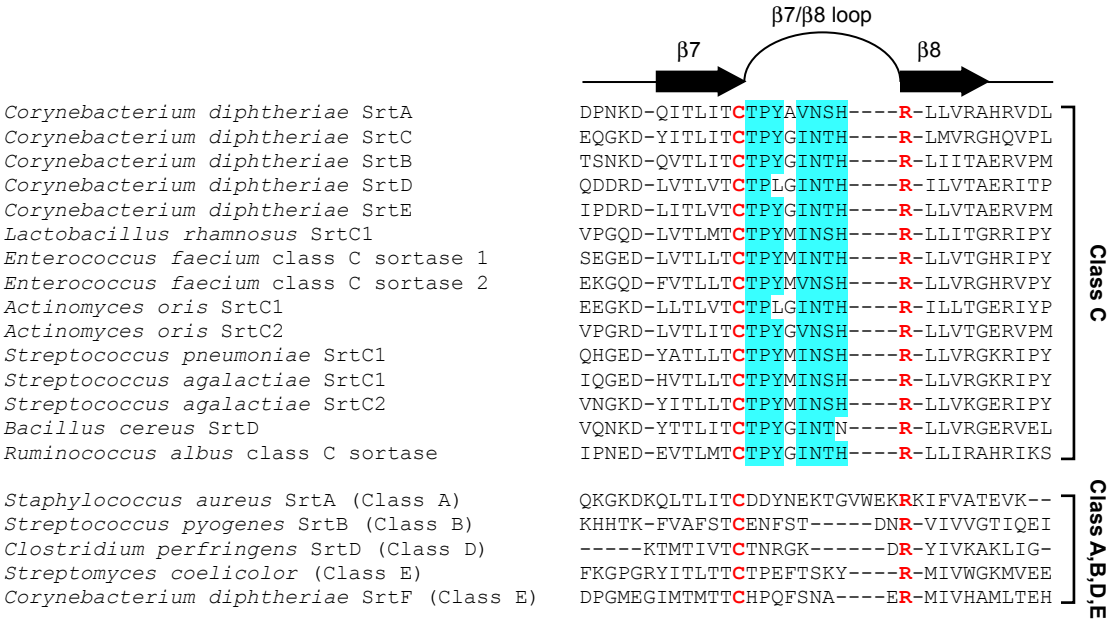

Figure S3: Chang et al.

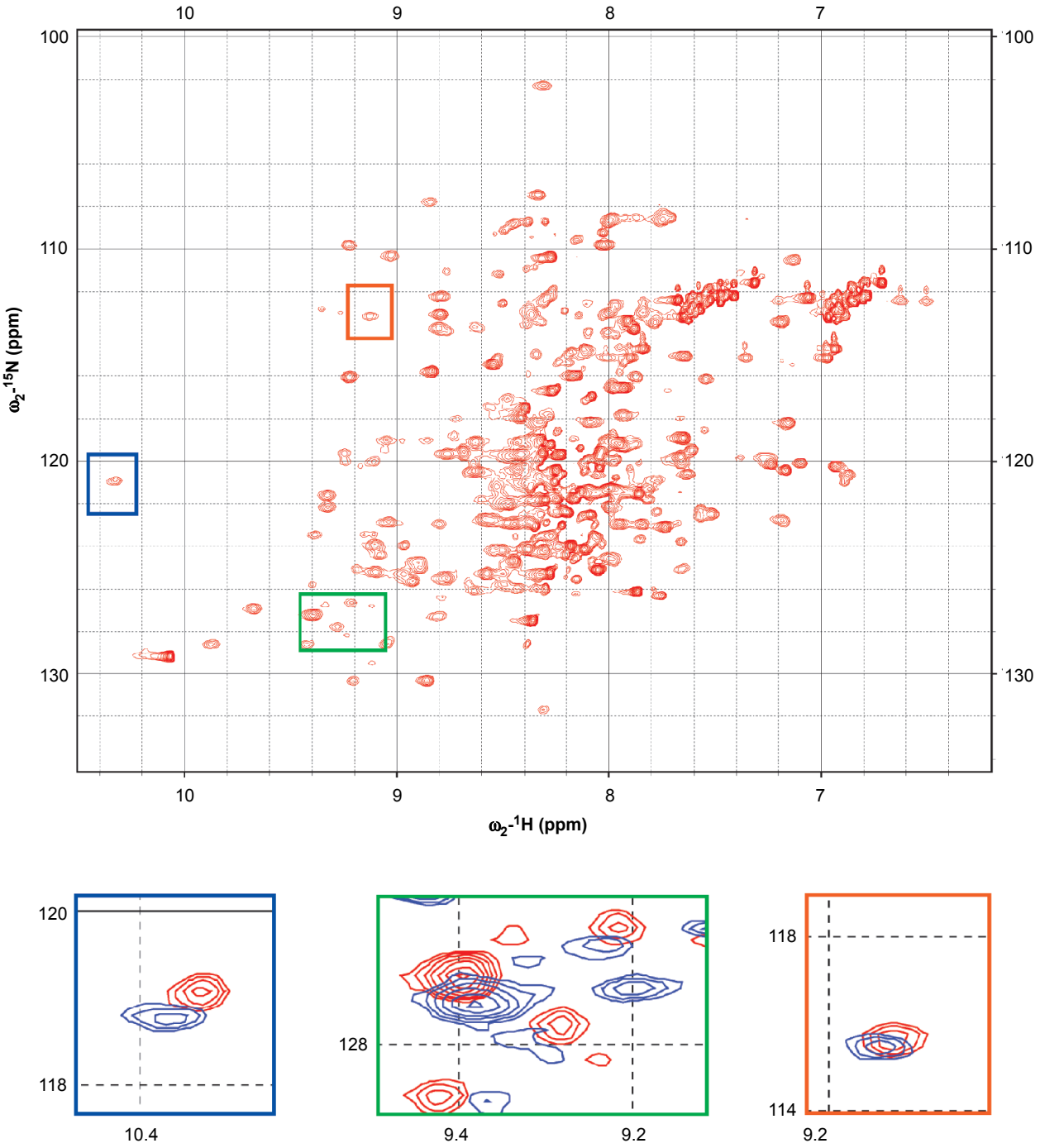

Figure S4: Chang et al.