STRUCTURAL BIOLOGY

Volume 75 (2019)

Supporting information for article:

Structures of major pilins in *Clostridium perfringens* demonstrate dynamic conformational change

Eiji Tamai, Seiichi Katayama, Hiroshi Sekiya, Hirofumi Nariya and Shigehiro Kamitori

Table S1. Plasmids used.

pGST-cppA	Expression vector for a gultathione-S-transferase fusion CppA-St13
pGALK	Suicide vector for in-frame deletion (Nariya *et al.*, 2011)
p*ΔcppA*	Suicide vector for preparing *cppA*-deficient mutant strain. Upstream and downstream (1 kb) of *cppA*-St13 were cloned into *Bam*HI - *Pst*I site of pGALK.
p*ΔcppBA*	Suicide vector for preparing *cppBA*-deficient mutant strain. Upstream and downstream (1 kb) of *cppBA*-St13 were cloned into *Bam*HI - *Pst*I site of pGALK.
p*ΔsrtC*	Suicide vector for preparing *srtC*-deficient mutant strain. Upstream and downstream (1 kb) of *srtC*-St13 were cloned into *Bam*HI – *Pst*I site of pGALK.
pJIR418	*C. perfringens-E. coli* shuttle vector (Sloan *et al.*, 1992)
p*cppA*	*cppBA* promoter and *cppA*-St13 gene were cloned into *Eco*RI - *Sph*I site of pJIR418
p*srtC*	s*rtC* -St13 gene with its promoter were cloned into *Eco*RI - *Bam*HI site of pJIR418
pColdICppA-St13	Expression vector for N-terminal His-tagged CppA-St13 lacking N-terminal signal peptide and C-terminal transmembrane domain. The gene was cloned into *Nde*I - *Xho*I site of pColdI (TaKaRa).
pColdICppA-SM	Expression vector for N-terminal His-tagged CppA-SM101 lacking N-terminal signal peptide and C-terminal transmembrane domain. The gene was cloned into *Nde*I - *Xho*I site of pColdI.
pColdICppA-D1D2-St13	Expression vector for D1-D2 domain of CppA-St13. The gene was cloned into *Nde*I - *Xho*I site of pColdI.
pETSrtC-SM	A vector expressing SrtC-SM101 fused with His-Tag at the C terminus. The synthesized *srtC* DNA, the region excluding N-terminal signal peptide and C-terminal transmembrane region, was cloned into the *Nde*I - *Xho*I site of pET22b.

In the vector containing the mutated gene, the mutation site is described after the name of the plasmid.

Table S2. Data collection statistics of SeMetCppA-D1D2-St13.

Values in parentheses are for the highest resolution bin.

	Peak	Edge	Remote
Beamline/Radiation source		PF BL5A	
Detector		PILATUS3-S-2M	
Temperature (K)		100	
Wavelength (Å)	0.97920	0.97934	0.96411
Resolution range (Å)		47.80 – 2.64 (2.71 – 2.64)	
No. of measured refs.	157,099 (12,162)	157,322 (12,172)	156,864 (12,229)
No. of unique refs.	22,275 (1,667)	22,293 (1,667)	22,213 (1,672)
Redundancy	5.3 (5.4)	6.5 (6.5)	3.4 (3.6)
Completeness (%)	99.9 (100.0)	99.9 (100.0)	99.9 (100.0)
Mean $Io/\sigma(Io)$	24.6 (4.7)	24.9 (4.7)	23.7 (4.1)
R_{merge} (%)†	6.0 (47.5)	6.1 (48.3)	6.6 (55.4)
$R_{\text{p.i.m.}}$ (%)‡	2.5 (19.0)	2.5 (19.4)	2.7 (23.1)
$CC_{1/2}$	0.999 (0.972)	0.999 (0.970)	0.999 (0.961)
Space group		*C*222	
Unit cell parameters	a = 89.14 Å	a = 89.16 Å	a = 89.05 Å
	b = 142.04 Å	b = 142.08 Å	b = 141.93 Å
	c = 61.75 Å	c = 61.76 Å	c = 61.68 Å

† $R_{\text{merge}} = \Sigma_{hkl} \Sigma_i | I_i(hkl) - \langle I(hkl) \rangle | / \Sigma_{hkl} \Sigma_i I_i(hkl)]$, where I_i (hkl) is the intensity value of the ith measurement of reflection hkl and $\langle I(hkl) \rangle$ is the mean value of $I(hkl)$ for all i measurements.

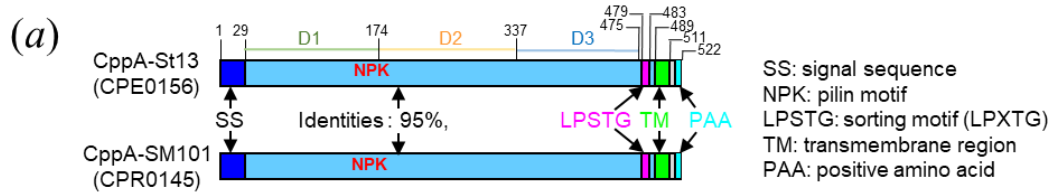
‡ $R_{p.i.m.} = \Sigma_{hkl}\{1/[N(hkl) - 1]\}^{1/2} \Sigma_i | I_i(hkl) - \langle I(hkl) \rangle | / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $N(hkl)$ is the multiplicity of reflection hkl.

Table S3. Phi - Psi angles of the residues in the loop connecting D1 and D2 in CppA-St13 and CppA-SM101.

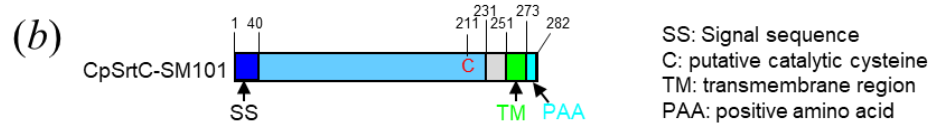
	CppA-St13		CppA-SM101	
	Phi (º)	Psi (º)	Phi (º)	Psi (º)
Ile171	-131	127	-132	114
Asn172	-98	122	-98	103
Pro173	-47	145	-70	154
Lys174	-97	78	-61	133
Asp175	-137	171	-84	98
Asn176	-128	157	-76	141
Thr177	-122	109	-108	157
Pro178	-74	152	-60	136
Ile179	-126	153	-104	130
Leu180	21	146	28	133

Figure S1

Amino acid sequences of CppA-St13, CppA-SM101, and CpSrtC-SM101. (*a*) Schematic drawings and alignment of amino acid sequences of CppA-St13 and CppA-SM101. The amino acids residues for X-ray structure determination are indicated by blue boxes. Unconserved residues are shown in yellow. (*b*) Schematic drawing of the amino acid sequence of CpSrtC-SM101. The amino acids residues for X-ray structure determination are indicated by blue boxes.




```
CppA-St13     1' MKNNILKKFS FMMVFIFVFI TTSIPVFAAT PSISKDAPIK GSITISKKGA TFTAYKLLDA
CppA-SM101    1" MKNNILKKFS FMMVFIFVFI TTSIPVFAAT PSISKDAPIK GSITISKKGA TFTAYKLLDA

CppA-St13    61' IKSGDAYEYS VNSDLKDFFN NSNYGSYSQE SIQKLFGEEV KEFAINLHKY ILENKKSGIE
CppA-SM101   61" IKSGDAYEYS VNSDLKDFFN NSNYGSYSQE SIQKLFGEQV KEFAVFLHKY VLDNKKSGIE

CppA-St13   121' LRDGQKNTVD LGYYLVTEFS SDSEGAAVAS TPIIVSVPQV SGDSWNYDVT INPKDNTPIL
CppA-SM101  121" LRDGQKNTVD LGYYLVTEAS SDSEGAAVAS TPIIVSVPQV SGDSWNYDVT INPKDNTPIL

CppA-St13   181' EKNIVKENQR VKTSSENIGD VVKYEVKASI PVYQKNAQMI MYKFTDTMSK GLTYDEKTGF
CppA-SM101  181" EKNIVKENQR VKTSSENIGD VVKYEVKASI PVYQKNAQDI MYKFTDTMSK GLTYDEKTGF

CppA-St13   241' KVTSGDKVFA KDIDYTVDVK KQEDGSTVIT INFVYENIKA YAETGITLNY QATLNKDAVI
CppA-SM101  241" KVTSGDKVFA KDMDYTVEVK KQEDGETVIT INFVYENIKA YAETGITLNY QATLNKDYVI

CppA-St13   301' SNKENLGNTN NIQLDYTNNP HVKDSYKKLT DKVTTYTFGF GITKVDSELN SKLLQGAEFS
CppA-SM101  301" SNKENLGNTN NIQLDYTNNP HVKDSYKKLT DKVTTYTFGF GITKVDSELN SKLLQGAEFS

CppA-St13   361' VKDAGGKEVA KYTYDEKGQV VELSGNGVTN SKGITTFLGL KEGKYSITEE VAPSGYSLLK
CppA-SM101  361" VKDRSGKEVA KYTYDEKGQV VELSGNGVTN SKGITTFVGL KEGKYEITEE VAPSGYSLLK

CppA-St13   421' NPVEVTITAN KDESGAYTGA ATIEISNGNK AGQIINDISE PDGNILFNVQ IENHAGFSLP
CppA-SM101  421" NPVEVTITAN KDESGEYTGA ATIEISNGNK AGQIINDISE MDGNILFNVQ IENHAGFSLP

CppA-St13   481' STGGLGNTGF IKIAIILLSI VCVLAILGLG YTKFENSRKT KN
CppA-SM101  481" STGGLGNTGF IKIAIILLSI VCVLSILGLG YTKFENSRKT KN
```




```
CpSrtC-SM101   1  MNFDIALDFL KRLKRIYFSI LILITLIALG FLLYPSFSNY INNKFAVSTI SDYTEKINNV

CpSrtC-SM101  61  KDEEVDDLIK NINKYNYDLF NGTAENQLPD YLNIHEGDVL GYIEIPSINI KLPIYYGTSV

CpSrtC-SM101 121  DILKKGVGVL EGTSLPVGGE NTHSVLSAHT GLANQKLFTD IDKLKDGDVF YLHILKKDLA

CpSrtC-SM101 181  YKVNQIKVVH PDEIDELKIS DDKDYVTLLT CYPYGINTER LLVRGERTDL SPSNVEQVQK

CoSrtC-SM101 241  EISTFNHSNE NLILIVIILN SVLIIIFLLF LIMKFKGKNK SR
```

Figure S2

Spatial arrangements of Lys residues in pilin motifs and CWSSs. (*a*) Structures of Lys174 residues and CWSSs of Mol-A/A' and Mol-B/B' of CppA-St13 are illustrated with the distances between Thr482(C)-Lys174(Nζ) presented as dotted lines. Omit map at the 3.5 σ contour level is shown for Phe477 – Thr482. (*b*) Structures of Lys174 residues and CWSSs of Mol-A/B and Mol-C/D of CppA-SM101 are illustrated with the distances between Thr482(C)-Lys174(Nζ) presented as dotted lines. Omit map at the 3.5 σ contour level is shown for Phe477 – Thr482. (*c*) Structures of Lys183 residues and CWSSs of Mol-A/A', Mol-B/B', and Mol-C/C' of RrgB (PDB: 2Y1V) are illustrated with the distances between Thr631(C)-Lys183(Nζ) presented as dotted lines. Note that Lys162 is between Lys183 and Thr631 to prevent them from contacting each other. (*d*) Structures of Lys203 residues and CWSSs of Mol-A/A' and Mol-B/B' of PitB (PDB: 4S3L) are illustrated with the distances between Thr383(C)-Lys203(Nζ) presented as dotted lines. Note that Lys203(Nζ) does not direct to Thr383(C).

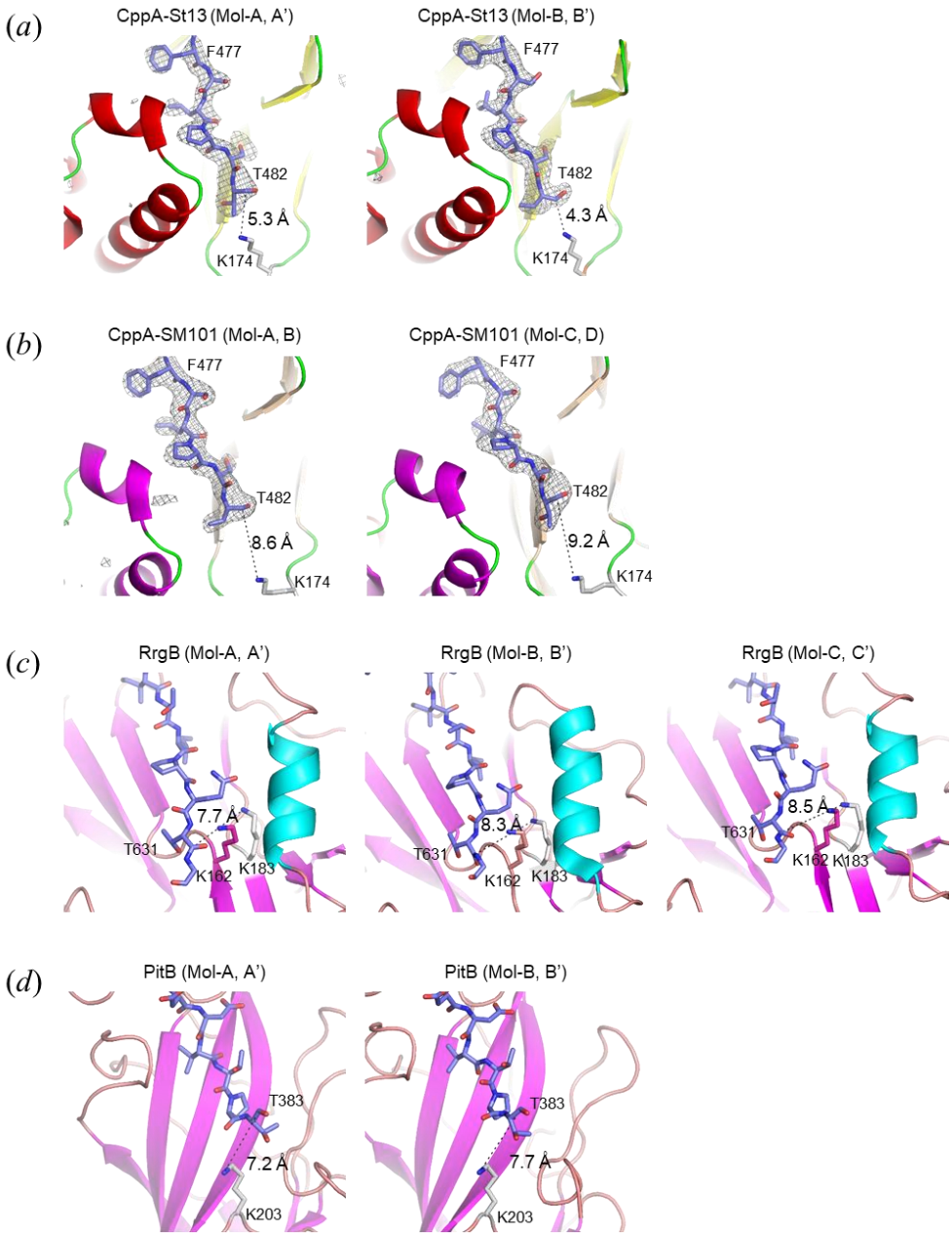


Figure S3
The interface between D1 and D2 of CppA-SM101. The structure of the interface between D2 and D3 of CppA-SM101 is illustrated. The amino acid residues of D2 and D3 are shown in yellow and blue, respectively.

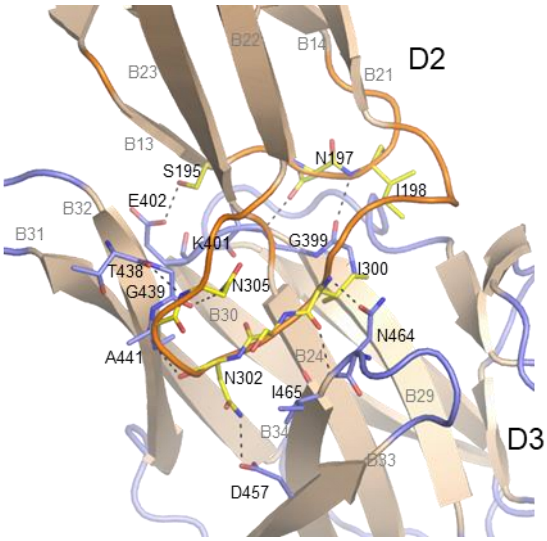


Figure S4

The procedure to build a model of a polymeric structure with a zigzag-shape of CppA-SM101. (*a*) Two dimers of CppA-SM101 (Mol-A, -B, -C, and -D) and two dimers with 150º rotation around the horizontal axis (Mol-A', -B', -C', and –D') are illustrated. Mol-A' and Mol-D' could be superimposed on Mol-B and Mol-C, respectively. (*b*) The superimposed model is illustrated. Note that two molecules (Mol-B' and Mol-C') are added to the model on the right side. (*c*) The model with the addition of two molecules (Mol-A' and Mol-D') on the left side is illustrated. The model includes eight molecules, and each chain has four molecules. (*d*) The polymeric structure model with 12 molecules, with each chain comprising six molecules, is illustrated.

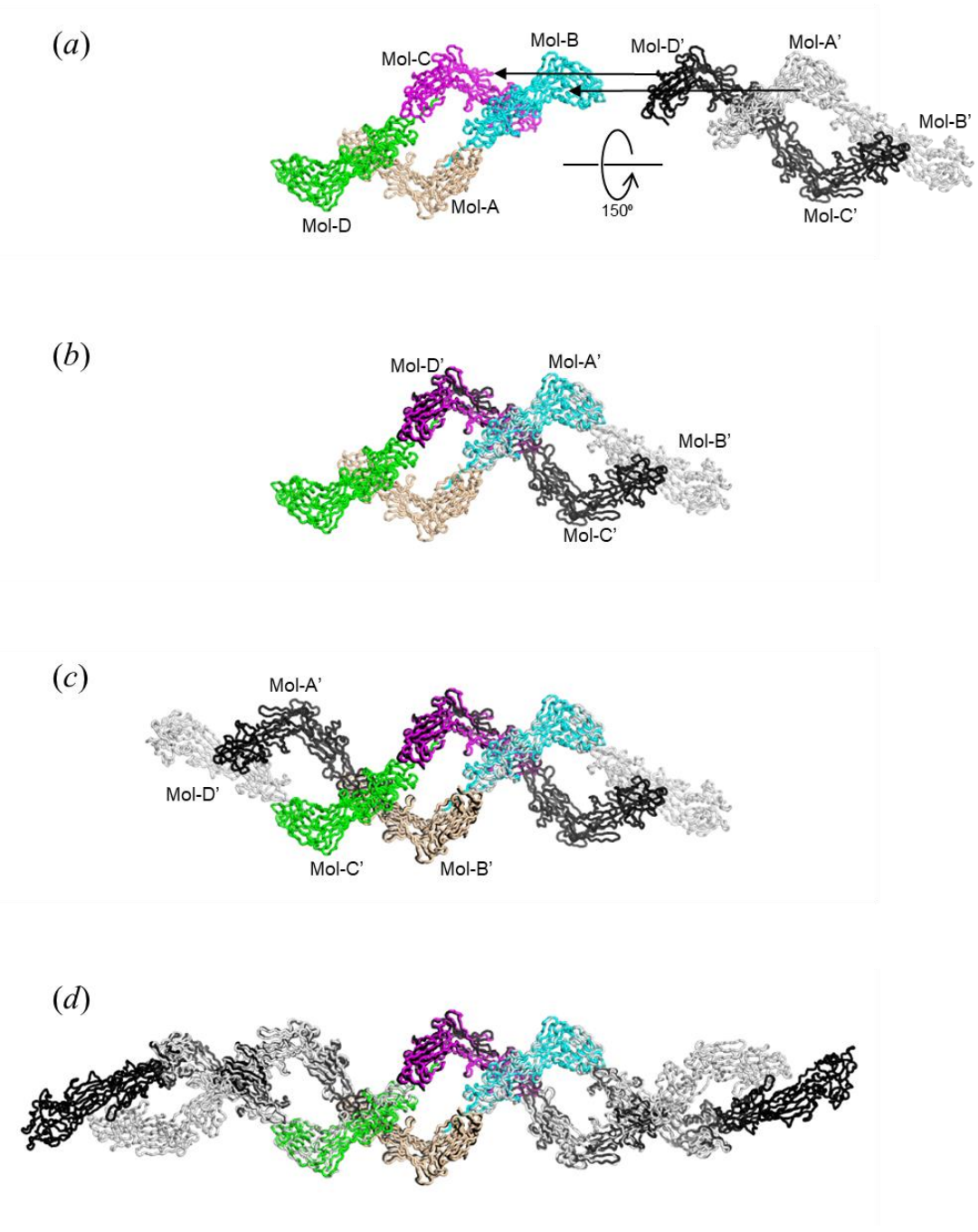


Figure S5

The procedure to build the preliminary docking model of CpSrtC-SM101 and CppA-SM101. (*a*) Using the program Coot (37), CpSrtC-SM101 was manually moved close to the C-terminal loop of CppA-SM101 (Mol-B) from the opposite side of Mol-A in Fig. 5*g*, so that CWSS of CppA-SM101 fitted into the CWSS-binding groove of CpSrtC-SM101. (*b*) A structural energy minimization calculation was carried out by CNS (39) with Engh and Huber stereochemical parameters using a model-minimization protocol. Only the structure of CpSrtC-SM101 was included in the calculation as a rigid group, while the structure of CppA-SM101 (Mol-B) was fixed. The energy-minimized model of CpSrtC-SM101 is shown in pink. (*c*) The docking model of CpSrtC-SM101 and CppA-SM101 (Mol-B) is illustrated. This structure is shown in Fig. 7*b*. (*d*) The structure of CppA-SM101 (Mol-A) is superimposed in the docking model of CpSrtC-SM101 and CppA-SM101 (Mol-B). (*e*) To avoid unusual short contacts between molecules, a structural energy minimization calculation was carried out. The structures of CpSrtC-SM101 and CppA-SM101 (Mol-A) were included in the calculation, while the structure of CppA-SM101 (Mol-B) was fixed. The energy-minimized model of CpSrtC-SM101 is shown in blue and violet. (*f*) The docking model of CpSrtC-SM101 and dimer of CppA-SM101 (Mol-A and Mol-B) is illustrated. This structure is shown in Fig. 8*c*.

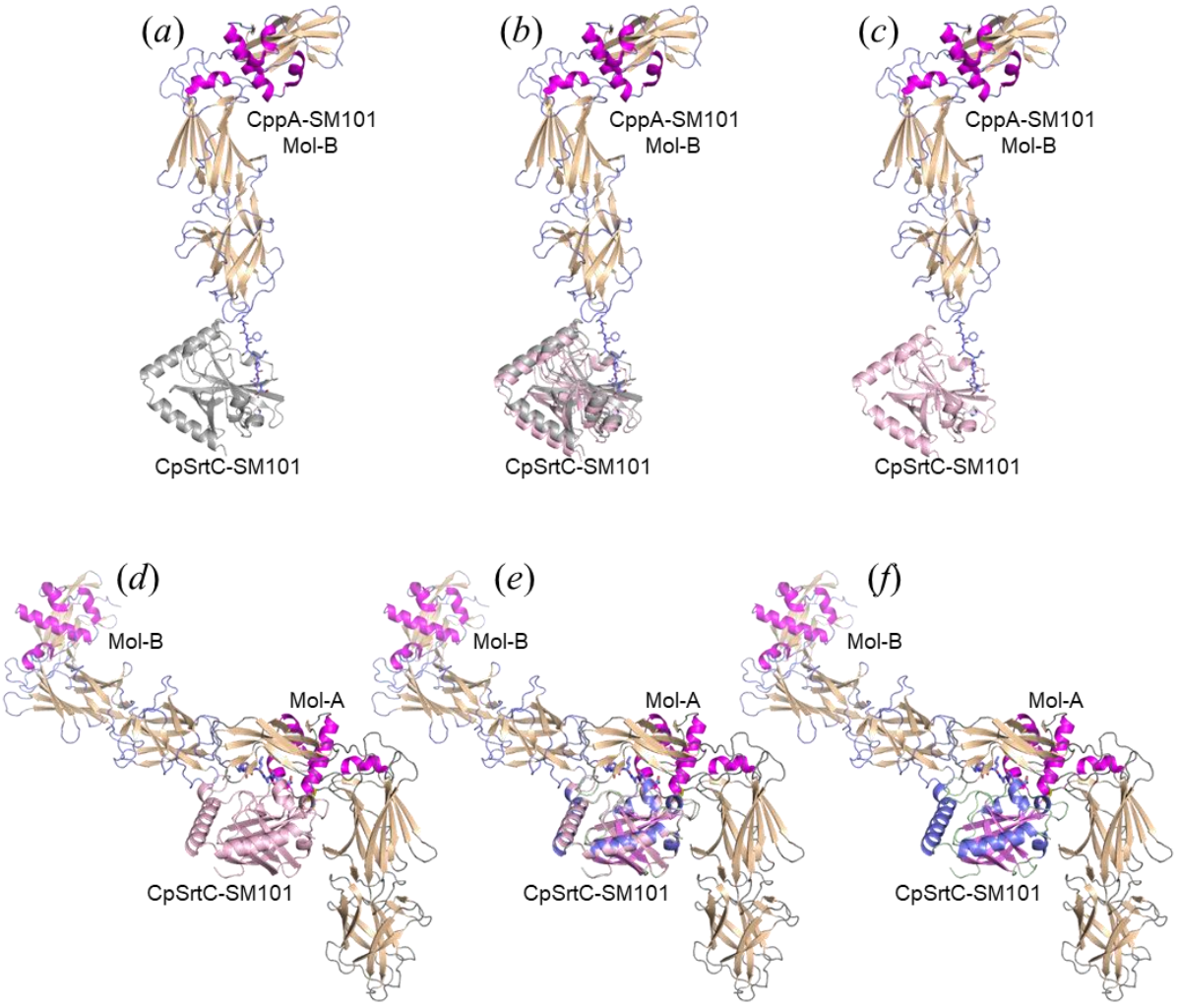


Figure S6

Structures of pilins from other bacteria. (*a*) The structure of BcpA (Mol-A), and its side view with another crystallographically independent molecule (blue) are illustrated. (*b*) The structure of FimP is illustrated. (*c*) The structure of SpaA is illustrated. (*d*) The structure of SpaD (Mol-A) and its side view with another crystallographically independent molecule (blue) are illustrated. (*e*) The structures of GGSpaD in an elongated structure and a bent structure (blue) are illustrated. Note that two drawings are viewed from the same direction to show the movement of D1 clearly.

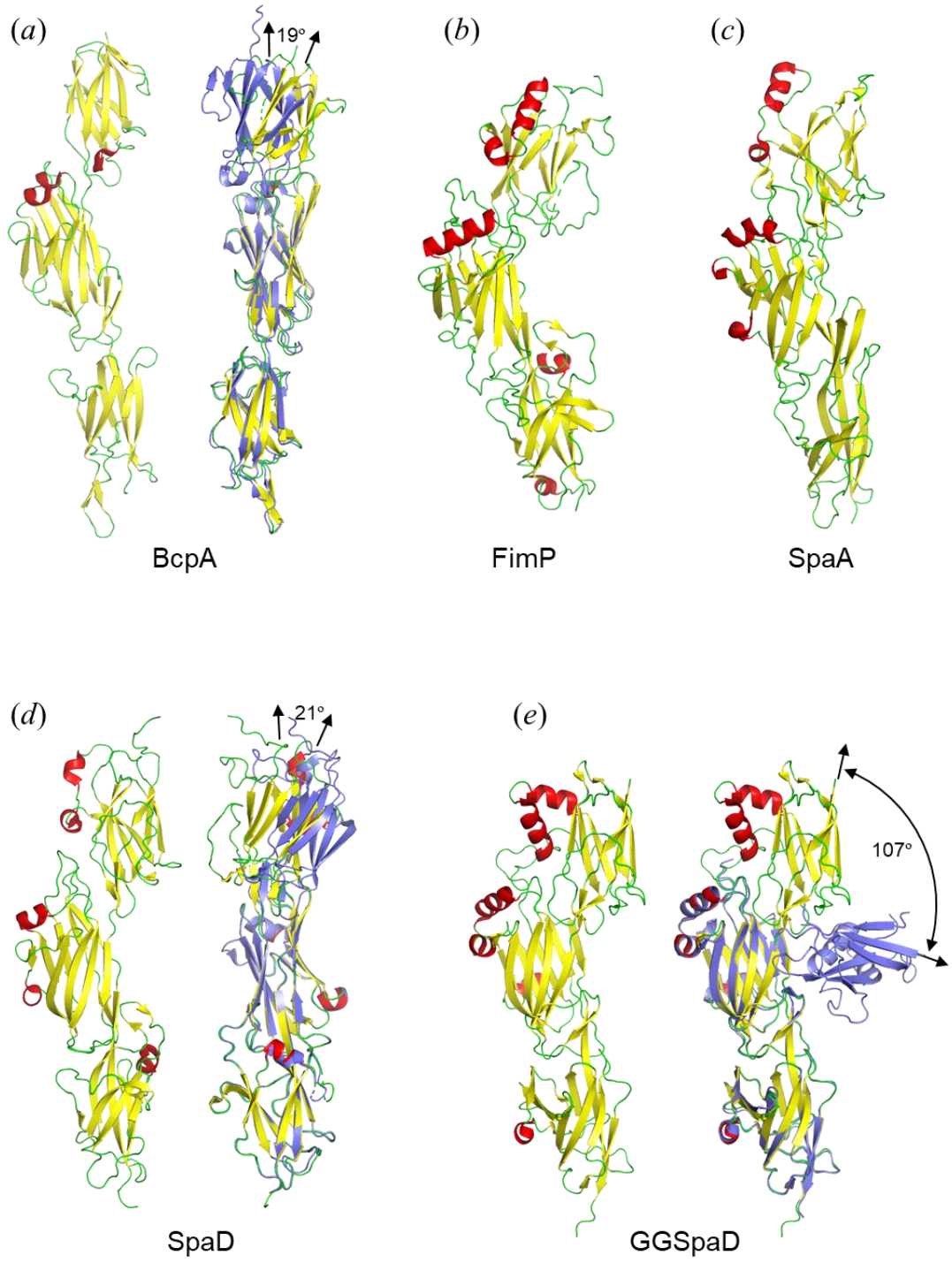


Figure S7

Structural comparison between CppA-SM101 and GGSpaD in the bent structure. (*a*) The structure of CppA-SM101 is illustrated viewing from three directions (front, side, and top). Lys174 is shown with a sphere model. (*b*) The bent structure of GGSpaD is illustrated viewing from three directions (front, side, and top). Lys180 is shown with a sphere model.