PNAS
www.pnas.org

Supplementary Information for

Interplay of a secreted protein with type IVb pilus for efficient enterotoxigenic *Escherichia coli* colonization

Hiroya Oki, Kazuki Kawahara, Takahiro Maruno, Tomoya Imai, Yuki Muroga, Shunsuke Fukakusa, Takaki Iwashita, Yuji Kobayashi, Shigeaki Matsuda, Toshio Kodama, Tetsuya Iida, Takuya Yoshida, Tadayasu Ohkubo, and Shota Nakamura

Tadayasu Ohkubo, and Shota Nakamura
Email: ohkubo@phs.osaka-u.ac.jp, nshota@gen-info.osaka-u.ac.jp

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.1805671115

Materials and Methods (SI)

Construction of CFA/III-positive *E. coli* HB101 strain and its mutant Strains. By nucleotide sequence analysis of the *cof* gene cluster, we found useable restriction enzyme sites (*Kpn*I and *Stu*I) for the cloning of CFA/III. The *Kpn*I (at position 1 in GenBank accession number AB049751) and the *Stu*I (at position 13,790) sites are located upstream of the putative promoter of the *cof* genes and downstream of *cofP*, the last gene in the cof gene cluster, respectively. pSH1134 (1), containing the *cof* gene cluster, was digested with *Stu*I (New England Biolabs), ligated to phosphorylated *Kpn*I linker (Takara), and subsequently digested with *Kpn*I (New England Biolabs). The resulting 13.8-kb *Kpn*I fragment was isolated by agarose gel electrophoresis and inserted into the *Kpn*I site of pMW119 (Nippon Gene), a pSC101-derived low-copy-number plasmid vector. The recombinant plasmid thus generated was designated pTT240. The *E. coli* K-12 derivative, HB101 (TaKaRa Bio), harboring recombinant plasmid, pTT240, containing the entire *cof* gene cluster, was used for CFA/III functional analyses. The abilities of CFA/III-positive *E. coli* HB101 (cof+ strain) and wild-type *E. coli* 31-10 to produce CFA/III and adhere to Caco2 cells were similar (1). For *cofJ* deletion plasmid construction, plasmid pTT240 was utilized as a template. The *cofJ* open reading frame (11769–12382) was digested with *Nco*I (New England Biolabs). The digestion product was gel-purified and self-ligated using a Ligation High Ver.2 kit (Toyobo). The ligated plasmid (Δ*cofJ*) was transformed into *E. coli* HB101 (Δ*cofJ* strain). The Δ*cofB* plasmid was constructed as previously reported (2). Briefly, the open reading frame (3,766–4,676) of cofB was digested with *NgoM*IV (New England Biolabs) and *Nhe*I-HF (New England Biolabs). The digested product was ligated with the adaptors containing *Xho*I restriction sites (5'-CCGGCTCGAGGCTAAG-3' and 5'-CTAGCTTCCGCTCGAG-3') using a Ligation High Ver.2 kit (Toyobo). The ligated plasmid (Δ*cofB*) was transformed into *E. coli* HB101 (Δ*cofB* strain). For *cofA* deletion plasmid construction, plasmid pTT240 was utilized as a template. The open reading frame of *cofA* (at position 3,373) and the *Spe*I site of the pMW119 vector were digested with *Spe*I-HF (New England Biolabs). The digested products were ligated with two adaptors, one containing the TAA stop codon for *cofA* gene deletion (5′-GTGATAACTGGGAGTATGTGGCAGTTGGTACTCTCGAGTAATCCTTCTGGTTCTTATGATGCTCTGTCTGC-3′) and the other restoring the digested pMW119 vector to the original condition (5′-TACAGATACCCACAACTCAAAGGAAAAGGACTAGTAATTATCATTGACTAGCCCATCTCAATTG-3′), using a NEBuilder HiFi DNA Assembly master mix (New England Biolabs). The ligated plasmid (Δ*cofA*) was transformed into *E. coli* HB101 (Δ*cofA* strain). For *cofD* deletion

plasmid construction, the open reading frame of *cofD* (at position 6,318) and the *Aat*II site of the pMW119 vector were digested with *Aat*II (New England Biolabs). The digested products were ligated with two adaptors, one containing the TAA stop codon for *cofD* gene deletion (5′-TAAAAGACGTATCTGTTCTTGCATTTATTGGACTCGAGTAACCGTCAGAATATCGTATGCTTGATAATGTAC-3′) and the other restoring the digested pMW119 vector to the original condition (5′-CGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACC-3′), using a NEBuilder HiFi DNA Assembly master mix. The ligated plasmid (Δ*cofD*) was transformed into *E. coli* HB101 (Δ*cofD* strain).

Cloning, Expression, and Purification of Recombinant Proteins. The truncated CofA construct (ΔN28-CofA) lacking the N-terminal hydrophobic segment was cloned, expressed, and purified as previously reported (3). Briefly, PCR products containing truncated CofA were inserted into a modified pET32b expression vector. *E. coli* expression strain Origami B (DE3) (Novagen) harboring the plasmid was grown in LB medium containing ampicillin (50 μg/mL). Cells collected from the induced culture were lysed by sonication. The expressed fusion protein was purified using a HiTrap Chelating HP column (GE Healthcare Biosciences) and dialyzed against 10 mM PBS, pH 7.4. The fusion protein thioredoxin tag was cleaved with thrombin protease, dialyzed, again purified using a HiTrap Chelating HP column, and further purified using a Superdex 75 26/60 (GE Healthcare Biosciences) pre-equilibrated with buffer (20 mM Tris-HCl pH 8.0 and 150 mM NaCl).

The N-terminally truncated CofB (ΔN28-CofB) construct was cloned, expressed, and purified as previously described (2). Briefly, a PCR fragment containing the *cofB* gene was inserted into a pET48b vector. *E. coli* SHuffle T7 Express LysY cells (New England Biolabs) were grown in LB medium containing kanamycin (20 μg/mL) and chloramphenicol (10 μg/mL). Cells collected from the induced culture were lysed by sonication and the supernatant was applied to a HiTrap Chelating HP column. The fusion protein thioredoxin tag was cleaved with HRV 3C protease. The protein was further purified by HiTrap Q HP anion-exchange column and by Superdex 200 size-exclusion column (both GE Healthcare Biosciences) pre-equilibrated with buffer as described above.

The DNA sequence encoding CofJ was amplified from plasmid pTT240 using forward and reverse primers, 5′-GCGCCCGGGTCGCCATCCTCTTCAGAAGG-3′ and 5′-CAAGAATTCTTATTAATCAAGGCCACAAGCCTTC-3′, respectively. The PCR product was digested with *Sma*I and *Eco*RI and ligated into the pET48b vector. The DNA sequence of N-

terminally truncated CofJ (ΔN24-CofJ) was amplified from the CofJ expression plasmid by inverse PCR using forward and reverse primers: 5′-CAAAGGTTTCAACCCGTTGG-3′ and 5′-GAGAAAGTCCCTGGGCCCAGACCCA-3′. The amplified linear DNA was self-ligated. These plasmids, containing full-length CofJ and ΔN24-CofJ, were transformed into *E. coli* expression strain SHuffle T7 Express LysY. The cells were grown in LB medium containing kanamycin (35 μg/mL) until an optical density at 660 nm and 310 °C of about 0.6, induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and incubated for 18 h at 20°C. The cells were centrifuged at 4,000 × *g* for 7 min, resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 1 M NaCl, 0.2% Triton X-100, and 0.2 mg/mL lysozyme), and sonicated on ice for 10 s using a Branson Sonifier 150 (Branson Ultrasonics) with a 50 s interval between each sonication. This procedure was repeated 15 times. The lysate was centrifuged at 24,000 × *g* for 1 h at 4°C. The supernatant was loaded onto a nickel-chelating column pre-equilibrated with buffer (50 mM Tris-HCl pH 8.0, 1 M NaCl). CofJ containing a thioredoxin and His6 tag was eluted with a linear gradient from 15 to 500 mM imidazole in 50 mM Tris-HCl, pH 8.0. The eluted protein was cleaved with 30 U Turbo3C protease (Accelagen), dialyzed against buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 15 mM imidazole) at 10°C overnight, and again purified using a nickel-chelating column. The cleaved CofJ protein contained two extra residues (Gly-Pro) at the N-terminus. The protein solution was concentrated and further purified using a Superdex 200 size-exclusion column pre-equilibrated with buffer as described above. Finally, the purified CofJ was checked by SDS-PAGE. Purified CofJ that was left to stand for 2 months at 4°C was degraded and used as degraded CofJ for pull-down analysis.

For CofJ-CofB complex formation, purified CofJ was mixed with the purified N-terminal truncated CofB and subsequently left to stand for 1 h on ice and then applied onto a Superdex 200 size-exclusion column pre-equilibrated with buffer as described above. The purified CofJ-CofB complex was checked by SDS-PAGE.

For construction of the *cofJ* expression plasmid for crystallization, the plasmid pTT240 was used as a template. The DNA fragment containing the *cofJ* gene was digested with *Sal*I (Roche) and *Hind*III (New England Biolabs). The resulting 1.74-kb DNA fragment was isolated by agarose gel electrophoresis and inserted into the pGEM-4Z vector (Promega). *E. coli* expression strain BL21 (DE3) pLysS (Novagen) was grown in 2xYT medium containing ampicillin (100 μg/mL) and chloramphenicol (20 μg/mL) until an optical density at 660 nm and 37°C of approximately 0.45, induced with 2.0 mM IPTG, and further incubated for 16 h at 27°C. The culture was centrifuged at 8,000 rpm for 30 min and the supernatant was collected, saturated with ammonium sulfate up to 60%, left to stand for 12 h at 4°C, centrifuged at 10,000 rpm for 30 min

to pellet crude CofJ, which was redissolved in 10 mM phosphate buffer, pH 7.2, mixed with ammonium sulfate up to 1.0 M, and centrifuged at 6,000 rpm for 10 min. The supernatant was applied onto a Phenyl Sepharose column (GE Healthcare Biosciences) pre-equilibrated with 10 mM phosphate buffer, pH 7.2, containing 1.0 M ammonium sulfate. CofJ was eluted with a step-wise gradient from 1.0 to 0 M ammonium sulfate in 10 mM phosphate buffer, pH 7.2. The eluent was dialyzed against 10 mM phosphate buffer, pH 7.2, loaded onto a DEAE Sepharose column (GE Healthcare Biosciences), and CofJ was eluted with a linear gradient from 0 to 2.0 M NaCl in 10 mM phosphate buffer, pH 7.2. The eluted CofJ was concentrated and further purified using a Superdex 75 (GE Healthcare Biosciences) pre-equilibrated with 10 mM phosphate buffer, pH 7.2, and finally checked by SDS-PAGE.

Generation of Fab Fragments of the anti-CofJ IgG Antibody. Anti-CofJ antibody was purified from rabbit serum by means of ammonium sulfate precipitation and a Hitrap rProtein A FF column (GE Healthcare Biosciences), according to manufacturer's instructions. The purified anti-CofJ was dialyzed against PBS. To generate Fab fragments, 2 mg/mL purified antibody was digested with 0.02 mg/mL papain (Nacalai Tesque) in a digestion buffer containing PBS and 20 mM L-cysteine for 3 h at 37°C. The mixture was passed over a Hitrap rProtein A FF column to remove undigested IgG and Fc fragments. Fab fragments were further purified using a Superdex200 size-exclusion column and checked by SDS-PAGE.

Bacterial Adherence Assay. *E. coli* strains harboring pTT240, Δc*ofB*, and Δc*ofJ* were cultured in Luria-Bertani (LB) medium containing ampicillin (100 μg/mL) at 25°C overnight, then plated on a CFA (1% casamino acids, 0.15% yeast extract, 0.005% $MgSO_4$, 0.0005% $MnCl_2$, and 2% agar) plate (1) containing ampicillin (100 μg/mL) to induce CFA/III expression at 37°C for 24 h. Suspensions of 10^9 bacteria/mL cultured on CFA agar in PBS were prepared. Recombinant CofJ and ΔN24-CofJ were mixed with Δ*cofJ* strain at 100 μg/mL final concentrations and left for 1 h at 25°C. The Fab fragment of anti-CofJ antibody was mixed with cof+ strain at 2.5, 5.0, and 10 mg/mL final concentration and left for 1 h at 25°C. Caco-2 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 50 μg/mL gentamicin and 10% fetal bovine serum at 37°C in 5% CO_2. Caco-2 cells were seeded on coverslips in 6-well tissue-culture plates. The cultures were used at post-confluence after 2 weeks of incubation—the condition for well-matured cells—washed in PBS twice, and added to 2 mL DMEM supplemented with 1% D-mannose and 100 μg/mL ampicillin. Then, 200 μL bacterial suspension was added and the mixture was incubated for 3 h at 37°C in 5% CO_2. The samples were washed three times with

PBS, fixed in methanol, stained with 20% Giemsa solution, and examined by oil-immersion light microscopy to observe bacterial adherence. For quantitative analysis, Caco-2 cells were seeded in 24-well plates and incubated for 2 weeks. Prior to the adhesion assay, Caco-2 cells were washed in PBS twice and added to 500 µL DMEM supplemented with 1% D-mannose and 100 µg/mL ampicillin. A suspension of 10^9 bacteria/mL cultured on CFA agar in PBS was prepared, of which 50 µL was added to the cells, incubated for 3 h at 37°C in 5% CO_2, washed three times with PBS, lysed with 0.1% Triton X-100 in PBS (500 µL), serially diluted, and plated onto LB agar plates with 100 µg/mL ampicillin for CFU counts. Recovery rates were calculated by dividing the number of adhered CFUs by that of the inoculum. Data represent the mean ± SD for five individual experiments. Statistical analysis was performed by two-tailed Welch's *t*-test. $^{**}P < 0.005$. $^{*}P < 0.05$.

Western Blot Analysis. Each strain was grown on CFA agar plates for CFA/III pilus expression. Cells were collected, resuspended in PBS, normalized on the basis of optical density at 660 nm (termed whole-cell culture [WCC]), centrifuged at 20,000 × *g* for 10 min, and the supernatant filtered (0.22-µm) to remove bacterial cells (Sup). These samples were denatured by incubation at 95°C for 10 min in SDS sample buffer. Proteins were separated by SDS-PAGE, electrophoretically transferred to PVDF membranes (Amersham) using a semidry blotting apparatus, and analyzed by western blotting. Membranes were blocked for 1 h in PBS-T (0.05% Tween-20 in PBS) containing 5% skim milk, washed with PBS-T three times, incubated for 1 h with rabbit IgG (1:20,000 dilution) against recombinant CofJ protein (1 mg/mL) and anti-DnaK antibody (Abcam) in Can Get Signal Immunoreaction Enhancer Solution 1 (TOYOBO), washed with PBS-T three times, and incubated for 1 h with a 1:2,000 dilution of Western BLoT Rapid Detect v2.0 (TaKaRa-Bio) in Can Get Signal Immunoreaction Enhancer Solution 2 (TOYOBO). Protein bands were visualized by enhanced chemiluminescence using the Western BLoT Chemiluminescence HRP Substrate (TaKaRa-Bio) and digitized using an LAS4000 imager (Fujifilm).

Transmission Electron Microscopy (TEM). *E. coli* strains harboring pTT240, Δc*ofB*, and Δc*ofJ* were cultured in LB medium containing ampicillin (100 µg/mL) for 20 h at 25°C and plated on CFA agar containing ampicillin (100 µg/mL) to induce CFA/III pilus expression for 28 h at 37°C. *E. coli* cells on agar plates were transferred to glow-discharged carbon-coated cupper grids by touching, then washed with five droplets of buffer (40 mM Tris-HCl (pH 7.4), 150 mM NaCl) prior to negative staining with 2% uranyl acetate. The prepared mesh was observed using a JEM-

1400 (JEOL Ltd.) operated at 120 kV accelerating voltage and recorded using the built-in CCD camera.

Pull-down Assay. Purified recombinant CofJ, degraded CofJ, and ΔN24-CofJ (20 μM) were incubated with Trx-His-CofA or Trx-His-CofB (10 μM) at 37°C for 1 h, loaded onto Ni Sepharose 6 Fast Flow beads equilibrated with buffer (20 mM Tris-HCl pH 8.0 and 150 mM NaCl), then washed with buffer containing 15 mM imidazole to remove unbound protein. Bound protein was eluted from the beads with buffer containing 500 mM imidazole and visualized by SDS-PAGE and Coomassie blue staining.

Isothermal Titration Calorimetry (ITC). ITC experiments were performed with iTC200 (GE Healthcare Biosciences). All samples were dissolved in a buffer solution containing 20 mM Tris-HCl (pH 8.0) and 150 mM NaCl. In the CofJ/CofB andΔN24-CofJ/CofB experiments, CofB solution (9.7 μM as a trimer) was loaded into the cell and CofJ or ΔN24-CofJ solution was loaded into the syringe (197 μM). In the CofJ(1–24)/CofB, CofJ(4–16)/CofB, CofJ(4–16)_F10A/CofB experiments, CofB solution (29 μM as a monomer) was loaded into the cell and each peptide solution was loaded into the syringe (570 μM). In the CofJ/CofA experiment, CofA (29 μM as a monomer) and CofJ solutions were loaded into the cell or syringe (197 μM), respectively. In the GalNAc/CofB and GalNAc/CofJ-CofB experiments, CofB solution (29 μM as a monomer) or CofJ-CofB solution (66 μM) was loaded into the cell and GalNAc solution was loaded into the syringe (570 μM). All titration was performed by sequentially titrating 1 μL for the first and 2 μL for all subsequent titration points with the syringe solution at intervals of 120 or 180 s at 37°C. The thermograms were analyzed with the NITPIC program (4) and the binding isotherms were fit with the A+B⇌AB interacting model using the SEDPHAT program (5). The dissociation constant and its standard deviation was determined from triplicate measurements.

Crystallization, Data Collection, and Structure Determination. Purified N-terminally truncated CofB was diluted to 5 mg/mL in 20 mM Tris-HCl (pH 8.0) and 150 mM NaCl, then mixed with CofJ(1–24) peptide at a molar ratio of 1:5 for co-crystallization on ice. Crystallization of CofJ(1–24)-CofB complex was performed by the sitting-drop vapor-diffusion method at 20°C. Diffraction quality crystals were obtained from the crystallization drop containing 1.6 μL protein solution (5 mg/mL) and 1.6 μL reservoir solution consisting of 100 mM CAPS, 20% PEG 8000, and 200 mM NaCl, pH 10.5. A crystal was flash frozen in a nitrogen gas stream at -173°C. X-ray diffraction data were collected on beamline BL26B1 at SPring-8 (Hyogo, Japan) using a Saturn

A200 CCD detector. All collected diffraction data were indexed and processed, and scaled using HKL2000 (6). Data collection and processing statistics are indicated in SI Table 2. The initial phases were calculated by molecular replacement (MR) using PHASER in the *PHENIX* software package (7). CofB domain 1 monomer (29–255) and domain 2 and domain 3 trimer (270–518) were employed as search models (PDB accession number: 5AX6). The MR solution contained three molecules of domain 1 and one trimer of domain 2 and domain 3. The linker between domain 1 and domain 2 (256–269) was subsequently modeled using the modeling tool Coot (8). Several refinement cycles using *phenix.refine* in the *PHENIX* software package (7) and Coot produced interpretable electron-density maps and revealed an additional well-defined density in the interface between two molecules of the CofB trimer. The bulky Phe10 electron density of the CofJ(1–24) peptide was observed in the hydrophobic pocket formed by Ala469, Thr484, and Val497 in one CofB monomer and Val435 and Leu462 in another. The CofJ(1–24) peptide was modeled from this Phe10 residue as a starting point using Coot. The CofJ(1–24)-CofB model was further refined using *phenix.refine* and Coot, resulting in a final model with R_{work} and R_{free} values of 0.192 and 0.232, respectively. The geometry of the final CofJ(1–24)-CofB model was checked with *MolProbity* (9). Phase determination and refinement statistics are shown in SI Table2.

To avoid CofJ digestion, the purified CofJ was immediately crystalized by the hanging-drop vapor-diffusion method with seeding at 20°C, and we checked the obtained crystals composed of full-length CofJ by SDS-PAGE. Diffraction-quality crystals were obtained from the crystallization drop containing 1.5 μL of protein solution (10 mg/mL) and 1.5 μL of reservoir solution consisting of 100 mM Tris-HCl, 14% PEG 3350, and 200 mM calcium acetate, pH 7.0. To improve the diffraction quality of crystals, the native CofJ crystals obtained from the crystallization mixture were dehydrated by transferring the hanging-drop coverslip to a new well and equilibrating to fresh reservoir solution of the same composition for 3 days. Single crystals of CofJ were soaked in cryoprotectant solution consisting of 20% glycerol, 100 mM Tris-HCl, 14% PEG 3350, and 200 mM calcium acetate, pH 7.0, and washed several times. Crystals were flash frozen in a nitrogen gas stream at -173°C. X-ray diffraction data of native CofJ were collected on beamline BL-17A at Photon Factory using an ADSC Quantum 270 CCD detector. The diffraction data of the $DyCl_3$ derivative CofJ were collected on beamline BL38B1 at SPring-8 using an ADSC Quantum 210 CCD detector. The collected diffraction data of native crystal were indexed and processed using XDS (10), and scaled using Aimless from CCP4 program suits (11). The data of $DyCl_3$ derivative crystal were indexed and processed, and scaled using HKL2000. Data collection and processing statistics are indicated in SI Table 2.

Initial phase calculations were performed by the single anomalous dispersion method using the $DyCl_3$ derivative crystal form of CofJ with the programs SOLVE/RESOLVE (7), which produced interpretable electron-density maps. Subsequent model building and refinement were performed by using the programs ARP/wARP in the CCP4 program suits (11), Coot, and REFMAC (12). The resulting model was used as the search model for MR with the program Phaser in the PHENIX software package using the data from the native CofJ crystal. The CofJ model was further refined using the program *phenix.refine*, resulting in a final model with R_{work} and R_{free} values of 0.180 and 0.212, respectively. The geometry of the final model of CofJ was checked with *MolProbity*. Phase determination and refinement statistics are shown in SI Table 2.

Analytical Ultracentrifugation. Sedimentation velocity experiments were performed using a Beckman Optima XL-I (Beckman Coulter) analytical ultracentrifuge equipped with a 4-hole An60 Ti rotor at 20°C using 12-mm double-sector charcoal-filled epon centerpieces with quartz windows. The CofJ-CofB complex was dissolved in a buffer solution containing 20 mM Tris-HCl (pH 8.0) and 150 mM NaCl. Sedimentation data were collected at 42,000 rpm using absorbance optics at 280 nm with a radial increment of 0.003 cm. The distribution of the sedimentation coefficient was analyzed using the *c(s)* method in the program SEDFIT (13). The range of sedimentation coefficients for fitting was 0–15 S, with a resolution of 300. The buffer density and viscosity calculated by the program SEDNTERP were 1.00500 g/mL and 1.0214 cP, respectively. Sedimentation equilibrium experiments were also performed using a Beckman Optima XL-I analytical ultracentrifuge equipped with a 4-hole An60 Ti rotor at 20°C using 12-mm six-sector charcoal-filled epon centerpieces with quartz windows. The rotor speeds were set to 5,000, 8,000, and 10,000 rpm, respectively. Data were collected at each rotor speed as the average of 4 scans at 0.001 cm intervals after the sedimentation equilibrium was reached, and were subjected to global analysis using SEDPHAT (5).

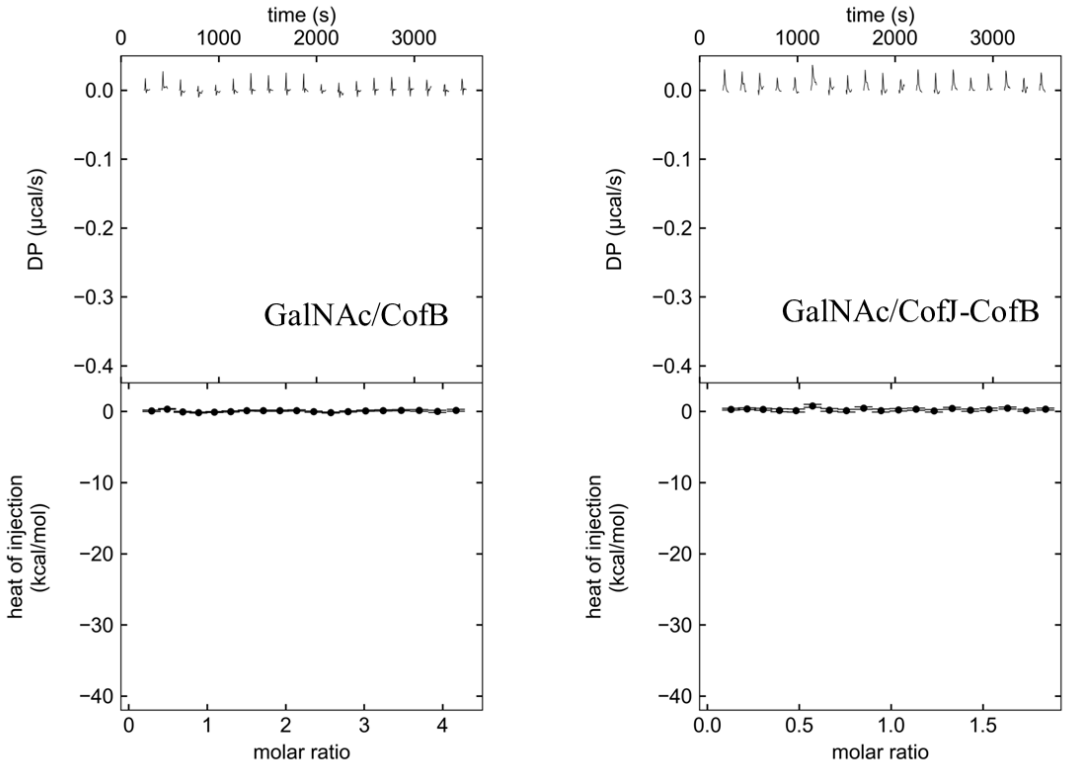


Fig. S1. Characterization of GalNAc-binding ability of minor pilin CofB. ITC profiles of the titration of GalNAc molecule with the pilin subunit CofB in the absence (left panel) or presence (right panel) of CofJ. For each case, raw titration data of GalNAc molecule injected into CofB or CofJ-CofB (top panel) and integrated heat measurements for the titration (bottom panel) are shown.

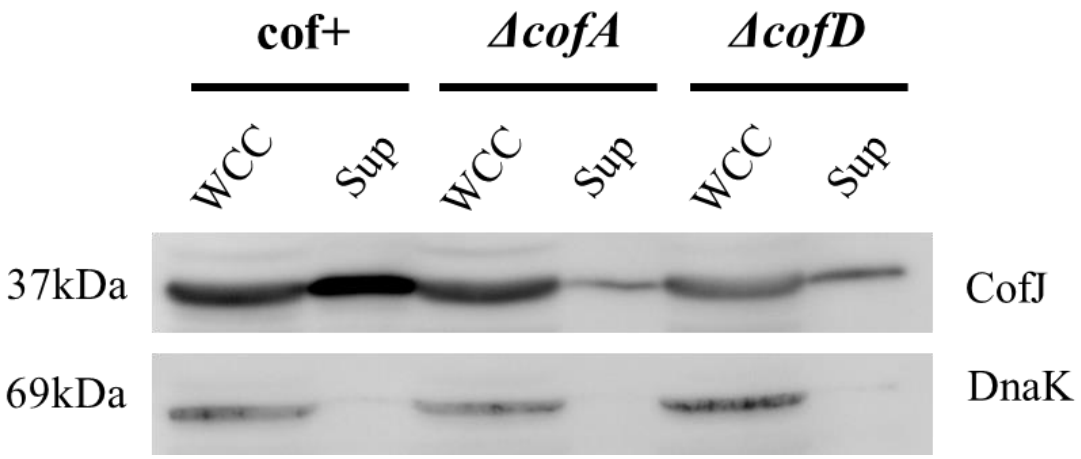


Fig. S2. Western blot analysis of whole-cell culture (WCC, left) and supernatant (Sup, right) of *E. coli* strains, cof+, Δ*cofA*, and Δ*cofD*, using anti-CofJ antibody. DnaK was detected with Anti-DnaK antibody, as a loading control.

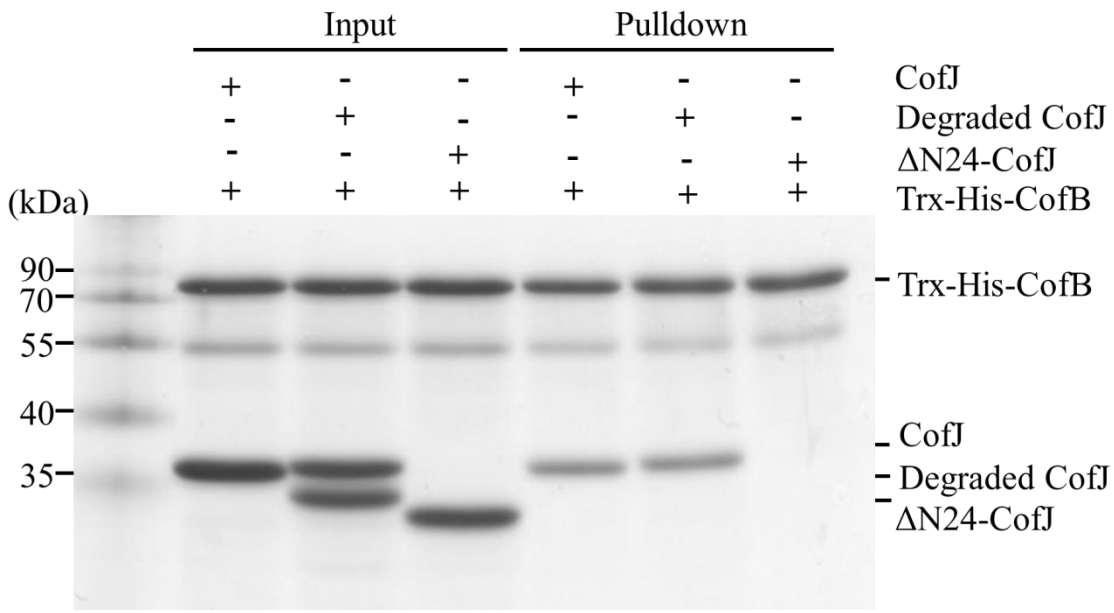
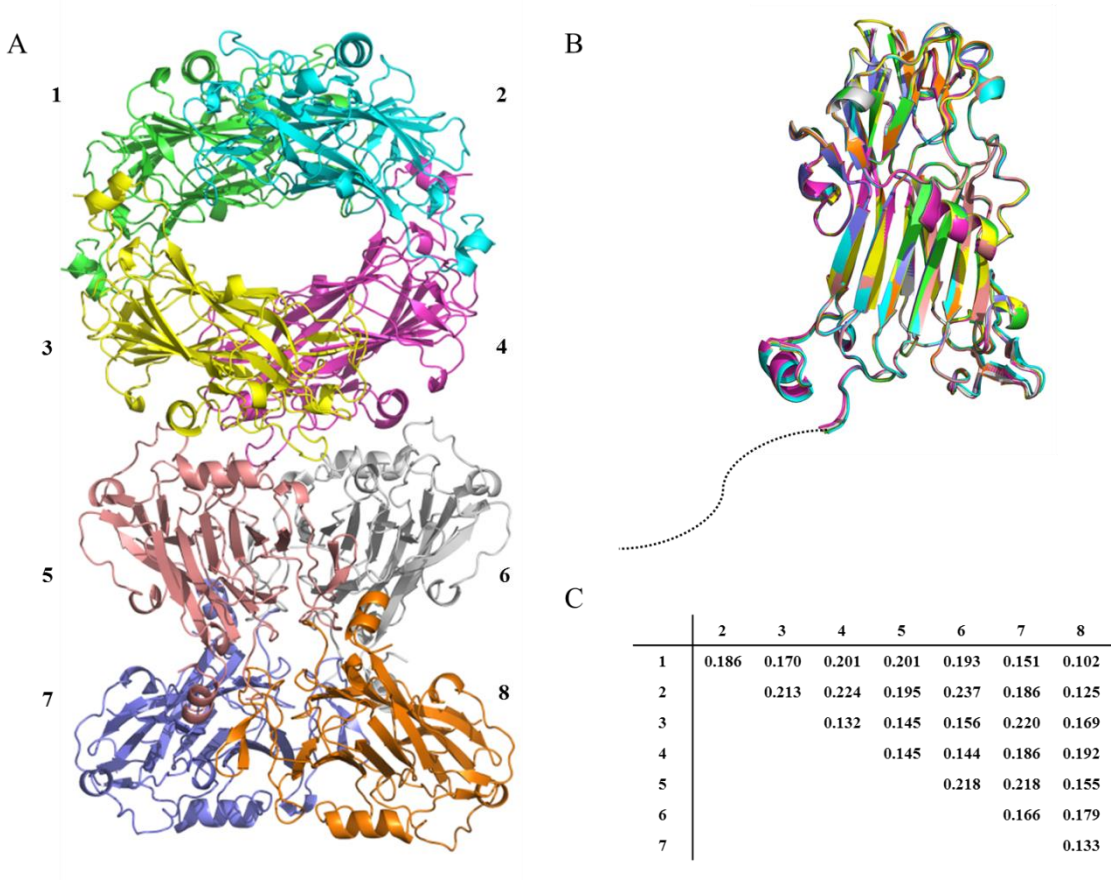


Fig. S3. Interaction analysis of minor pilin CofB and secreted protein CofJ. Pull-down assay was carried out for the following pairs, Trx-His-CofB and whole CofJ, Trx-His-CofB and degraded CofJ, and Trx-His-CofB and ΔN24-CofJ. The mixed samples were loaded onto Ni Sepharose beads that were then washed with buffer containing a low concentration of imidazole. Bound proteins were eluted with buffer containing a high concentration of imidazole and detected by SDS-PAGE. The molecular weight markers are indicated on the left, and the position of each protein is indicated on the right.



	2	3	4	5	6	7	8
1	0.186	0.170	0.201	0.201	0.193	0.151	0.102
2		0.213	0.224	0.195	0.237	0.186	0.125
3			0.132	0.145	0.156	0.220	0.169
4				0.145	0.144	0.186	0.192
5					0.218	0.218	0.155
6						0.166	0.179
7							0.133

Fig. S4. Crystal structure of CofJ. (*A*) The CofJ crystal structure asymmetric unit comprises eight chains, illustrated in unique colors and in ribbon representation. (*B*) Structural superposition of CofJ molecules. Superposition of eight CofJ molecules in the crystal depicted as a ribbon model. Colors of molecules in this panel correspond to those in Fig. S3A. The disordered residues, Ser1 to Asp22, were represented by a dashed line. (*C*) Average Cα root-mean-square deviation values among eight monomers of CofJ.

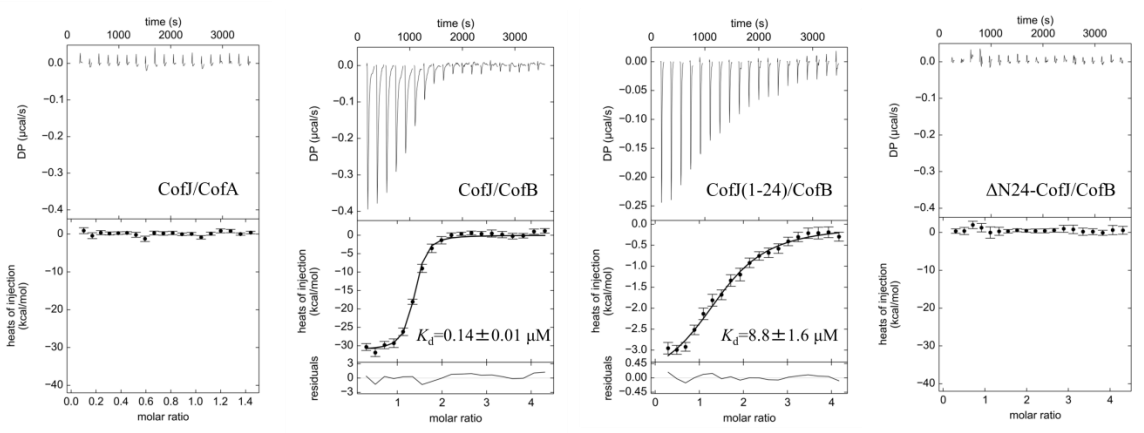


Fig. S5. ITC profiles of the titration of CofJ, CofJ (1–24) peptide, and ΔN24-CofJ with the pilin subunit CofA or CofB. For all experiments: Upper panel: representative thermogram of CofJ, CofJ (1–24) peptide, and ΔN24-CofJ injected into CofA or CofB. Middle panel: integrated heat plot for the titration. Bottom panel: residual plot of the fitting except for the case of ΔN24-CofJ/CofB and CofJ/CofA.

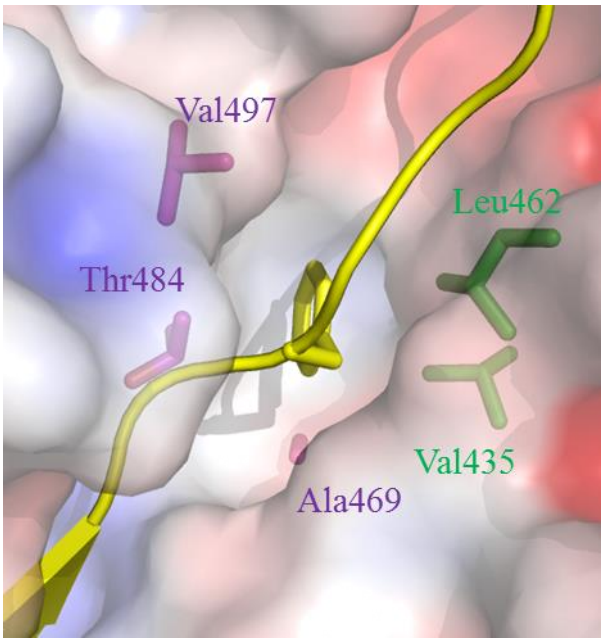


Fig. S6. Close-up view of the Phe10 of CofJ (1–24) peptide. Electrostatic surface potential representation of domain 3 of the CofB-binding groove and the Ser5–Pro15 fragment of CofJ (1–24) shown in yellow cartoon representation is shown. The electrostatic surface potential was generated by the Adaptive Poisson-Boltzmann Solver (APBS) (14). Red and blue surfaces represent negative and positive charges, respectively. The side chain of Phe10 fitted in the hydrophobic pocket formed by Ala469, Thr484, and Val497 in one CofB monomer (magenta) and Val435 and Leu462 in another CofB monomer (green) are shown. The side chains of these residues are shown in stick model.

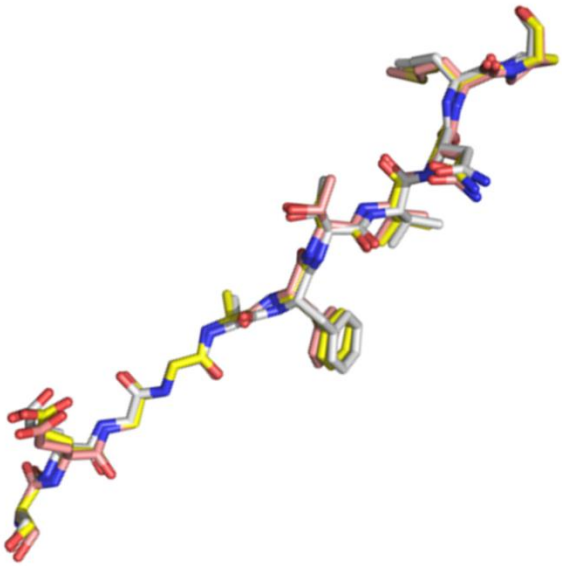


Fig. S7. Structural superposition of each CofJ (1–24) molecule in the CofJ(1–24)-CofB complex. Superposition of three CofJ (1–24) molecules in the crystal depicted as grey, yellow, and pink stick models.

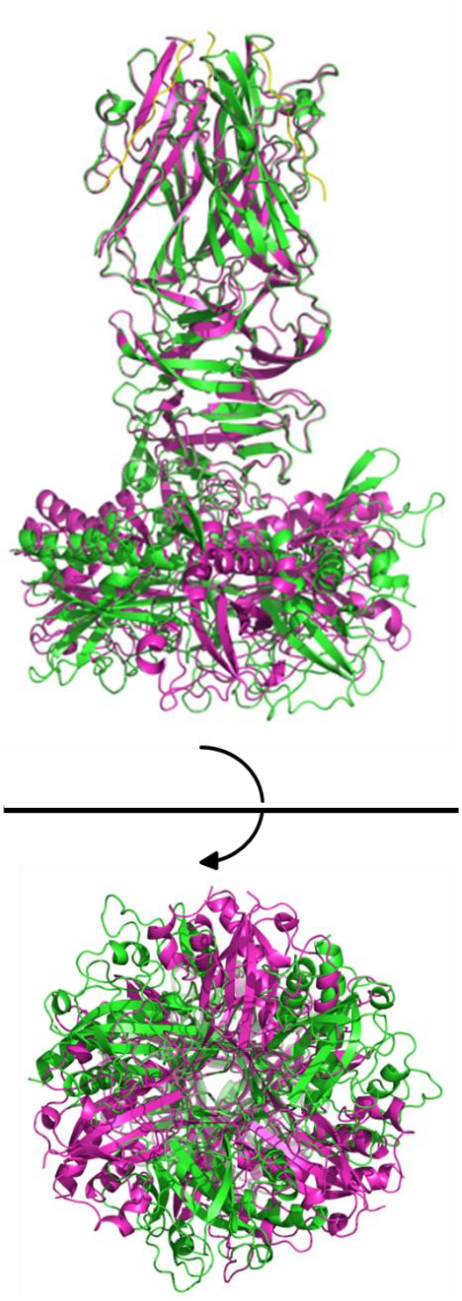


Fig. S8. Structural superposition of CofJ (1–24)-CofB complex with CofB alone. (top) Side view of the superposition of CofJ (1–24)-CofB depicted as green ribbons onto the CofB homo-trimer depicted as magenta ribbons. The structure was built by superimposing the crystal structure of CofJ (1–24)-CofB onto the CofB homo-trimer with PyMOL. (bottom) Bottom view of the superimposed structure showing the marked difference of the relative arrangement of domain1.

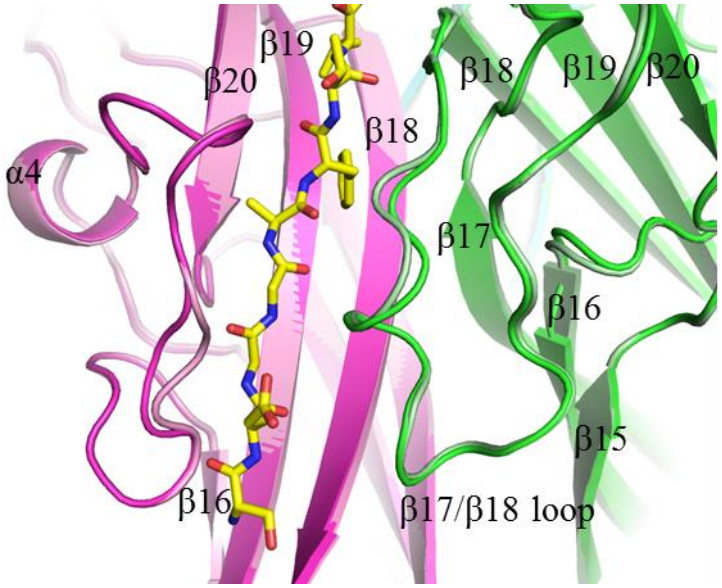


Fig. S9. Structural comparison of CofJ (1–24)-CofB (magenta, green) and CofB alone (pink, light green). CofJ (1–24) peptide is shown in yellow stick representation. Local conformational changes in the long β17/β18 loop of one monomer and the α4/β16 loop of another monomer were found by comparison with the structure of CofB alone.

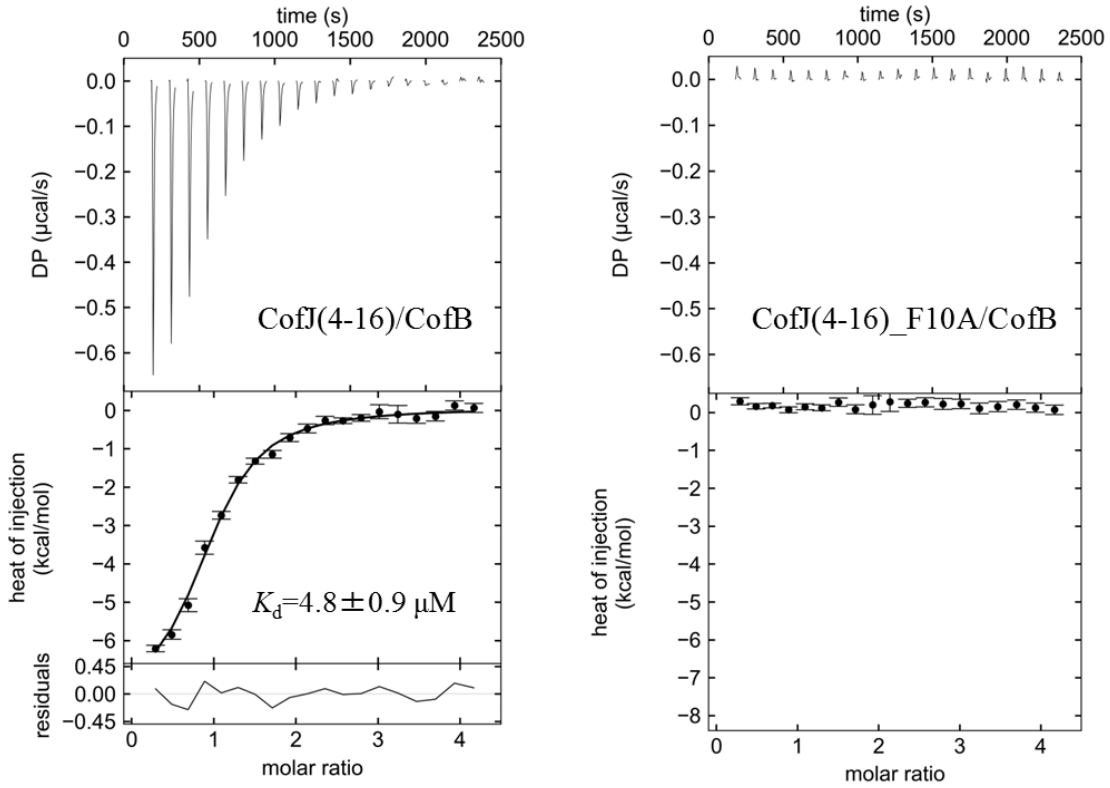


Fig. S10. ITC profiles of the titration of CofJ (4–16) and CofJ (4–16)_F10A peptides with the pilin subunit CofB. The upper panels show representative thermograms of CofJ(4–16) and CofJ (4–16)_F10A peptide injected into CofA or CofB. The middle panel (left) and lower panel on the right show an integrated heat plot for the titration. Bottom panel on the left: residual plot of the fitting except for the case of CofJ(4–16)_F10A peptide.

Fig. S11. N-terminal amino-acid sequence alignment among ETEC CofJ, ETEC LngJ, and *Citrobacter rodentium* CfcJ and *Vibrio cholerae* TcpF. The mature N-terminal sequences of CofJ, LngJ, CfcJ, and TcpF generated from N-terminal signal sequence cleavage are predicted by SignalP 4.1 (15). The black arrow indicates the conserved aromatic residue. In the cases of CofJ, LngJ, and CfcJ, the phenylalanine residue is situated at this position. TcpF has tyrosine residue at the corresponding position.

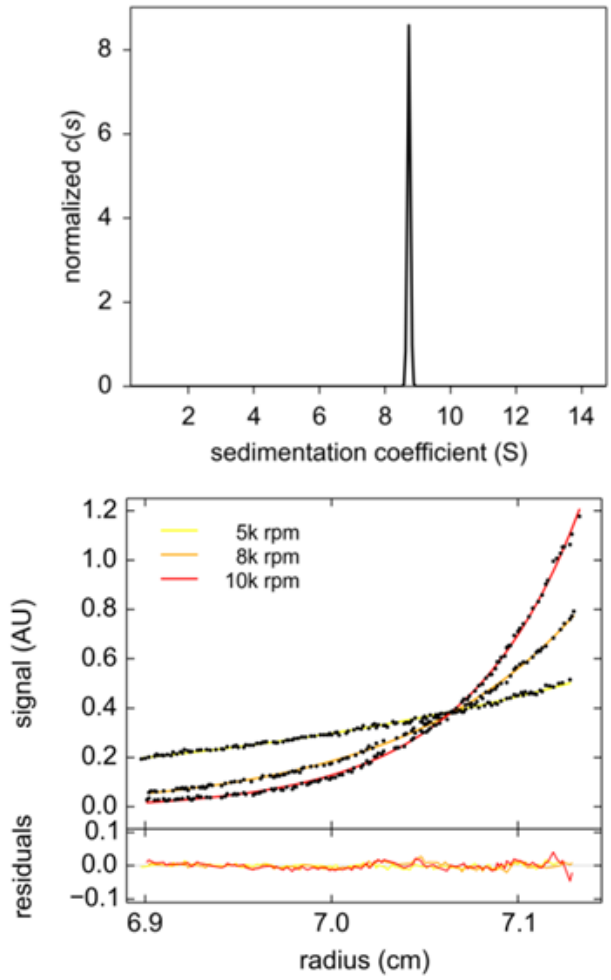


Fig. S12. Analytical ultracentrifugation results of CofJ-CofB complex. Top panel, *c(s)* distribution profile for the CofJ-CofB complex based on sedimentation velocity data collected at 42,000 rpm, 20°C. Bottom panel, plot of the sedimentation equilibrium data with the residuals from the best fit to a single ideal species.

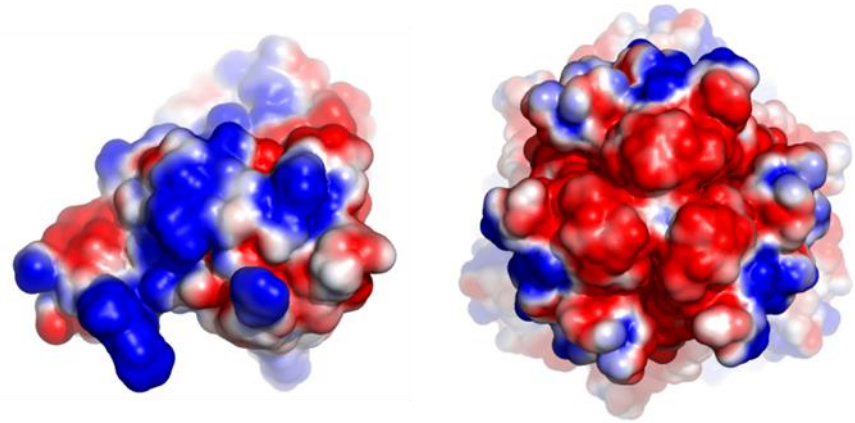


Fig. S13. Electrostatic surface potential of CofJ and CofB. The electrostatic surface potential was generated by APBS (14). Left and right panels, electrostatic surface potential representations of CofJ bottom region and CofB trimer top region, respectively.

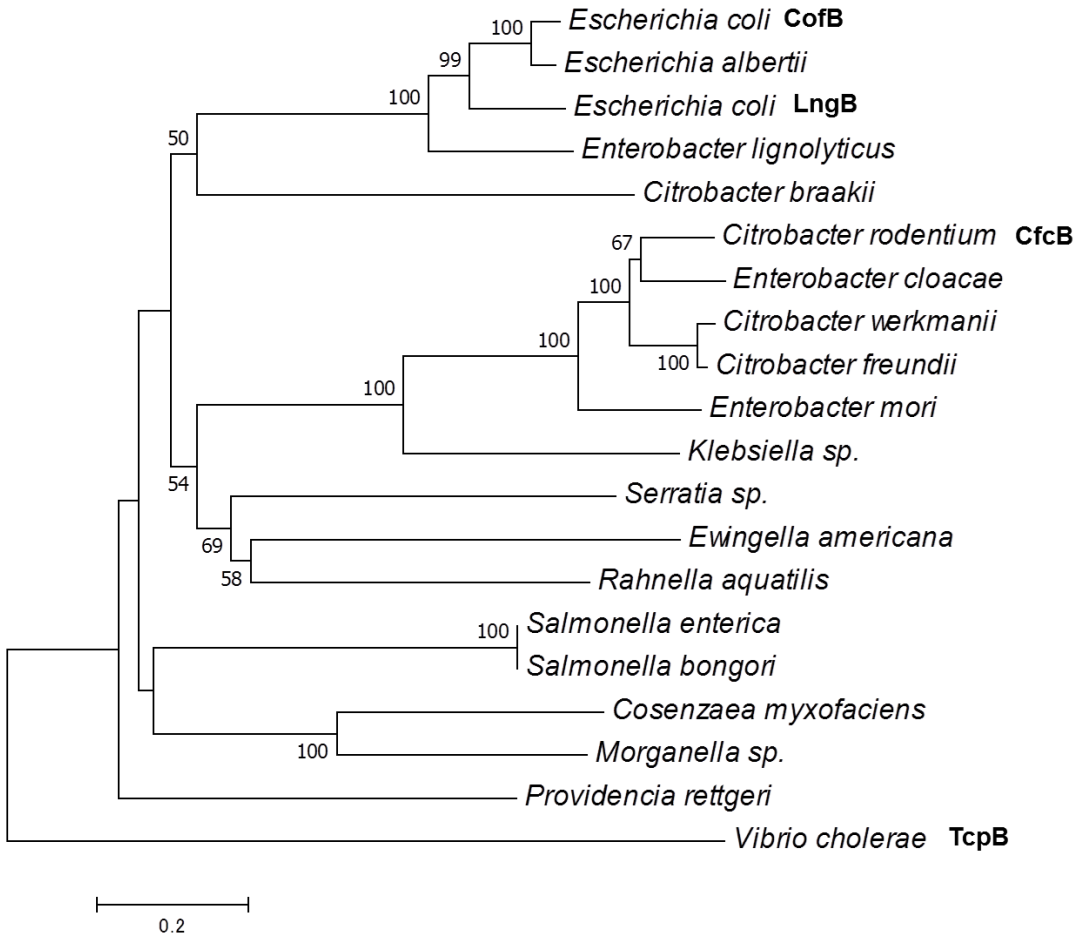


Fig. S14. Evolutionary relationships of CofB and other (predicted) minor pilins. To investigate the presence of minor pilins of T4Pb in other bacteria, BLASTP (15) analysis was performed using domain 3 of the CofB amino-acid sequence (354–518). Selected protein sequences from different bacteria shown in Table 2 (e-value < 10^{-3} cut-off) were used for multiple amino-acid sequence alignment with CofB. TcpB from *Vibrio cholerae* is included for comparison. A multiple alignment was generated using the MEGA7 (16) implementation of MUSCLE (17). The evolutionary history was inferred using the neighbor-joining method (18). The optimal tree with sum of branch length = 7.60618082 is shown. The percentages of bootstrap support above 50% are shown next to the branches (19). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (20) and are in the units of the number of amino-acid substitutions per site. The analysis involved 20 amino-acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 322 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

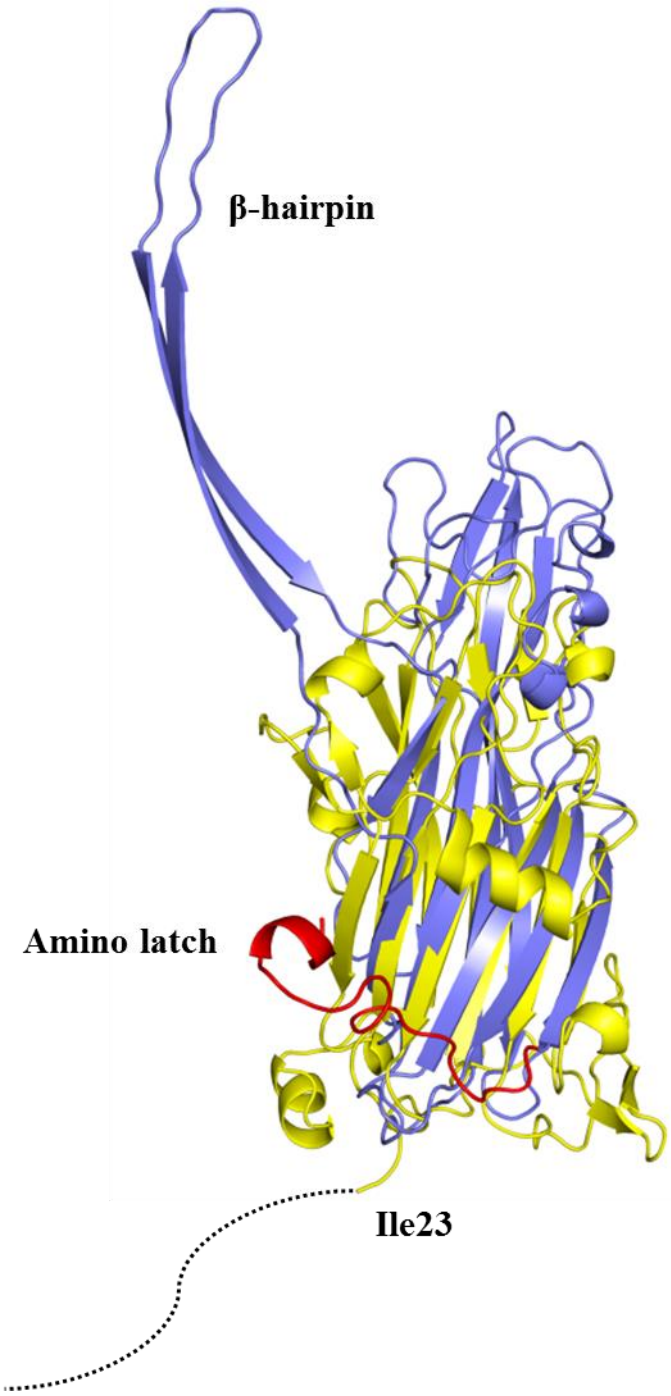


Fig. S15. Structural superposition of CofJ (yellow) onto α-hemolysin (medium purple, PDB code: 7AHL). The disordered residues Ser1 to Asp22 of CofJ are represented by a dashed line, and the corresponding "amino-latch" N-terminal segment, Ala1 to Val20, of α-hemolysin is colored red.

Table S1. BLAST search of the amino-acid sequence of H-type lectin domain (domain 3) of CofB.

Organism	Accession number	*E*-value
Escherichia albertii	WP_059259580.1	3e-91
Escherichia coli (LngB)	EMV54823.1	1e-76
Enterobacter lignolyticus	WP_013365029.1	5e-71
Serratia sp.	WP_056781109.1	1e-22
Salmonella enterica	WP_088546469.1	1e-19
Salmonella bongori	WP_000680636.1	1e-19
Ewingella americana	WP_084674026.1	2e-17
Cosenzaea myxofaciens	WP_074388166.1	5e-17
Rahnella aquatilis	AFE60214.1	2e-14
Morganella sp.	WP_036419502.1	3e-11
Citrobacter freundii	WP_038634225.1	3e-07
Providencia rettgeri	WP_071548095.1	5e-07
Enterobacter cloacae	WP_095450582.1	2e-06
Citrobacter werkmanii	WP_085048648.1	5e-06
Enterobacter mori	WP_089599804.1	2e-05
Citrobacter rodentium (CfcB)	AAO17806.1	5e-05
Klebsiella sp.	WP_049838909.1	2e-04
Citrobacter braakii	WP_080860199.1	4e-04

Table S2. Data collection and refinement statistics of native CofJ, $DyCl_3$-derivative CofJ, and CofJ (1–24)-CofB complex. Values in parentheses refer to the highest resolution shell.

Data collection	CofJ		CofJ(1-24)-CofB
	Native	$DyCl_3$ derivative	
Beamline	Photon Factory BL-17A	SPring-8 BL38B1	SPring-8 BL26B1
Detector	ADSC Quantum 270	ADSC Quantum 210	Saturn A200
Wavelength (Å)	0.98000	1.58590	1.00000
Space group	$P2_12_12_1$	$P2_12_12_1$	$P6_5$
Unit-cell parameters (Å, º)	a=142.2, b=147.5, c=149.3	a=140.9, b=147.2, c=148.8	a=b=157.2, c=118.1, γ=120
Total masured reflections	2325718 (109819)	1389942 (127088)	229612 (21871)
Unique reflections	307221 (14749)	203976 (19585)	20699 (2057)
Resolution (Å)	49.2-1.76 (1.79-1.76)	50.0-2.01 (2.08-2.01)	60.0-3.52 (3.58-3.52)
R_{merge}	0.120 (1.890)	0.114 (0.401)	0.290 (1.863)
$CC_{1/2}$	0.998 (0.521)	0.997 (0.921)	0.995 (0.500)
Completeness (%)	99.4 (96.3)	99.2 (97.9)	100.0 (100.0)
Willson B-factors (Å^2)	21.2	17.8	90.7
Average I/σ (I)	10.6 (1.3)	24.3 (5.3)	13.5 (2.0)
Redundancy	7.6 (7.4)	6.8 (6.5)	11.1 (10.6)
Refinement			
Resolution (Å)	49.2-1.76 (1.78-1.76)		54.2-3.52 (3.61-3.52)
Reflections	306946 (9801)		20668 (1459)
No. Atoms			
Protein	2403		1473
Water	3044		-
Ion	4		-
R_{work}	0.180 (0.340)		0.192 (0.302)
R_{free}	0.212 (0.367)		0.232 (0.365)
R.M.S.D. from ideal			
Bonds (Å)	0.008		0.004
Angles (º)	0.955		0.777
B-factors (Å^2)	26.8		106.5
Ramachandran plot analysis (%)			
Most favored	97.9		96.1
Allowed	2.1		3.7

SI References

1. Taniguchi T, *et al.* (2001) Gene cluster for assembly of pilus colonization factor antigen III of enterotoxigenic *Escherichia coli*. *Infect Immun* 69:5864–5873.
2. Kawahara K, *et al.* (2016) Homo-trimeric Structure of the Type IVb Minor Pilin CofB Suggests Mechanism of CFA/III Pilus assembly in human enterotoxigenic *Escherichia coli*. *J Mol Biol* 428:1209–1226.
3. Hill AVS (1991) HLA associations with malaria in Africa: Some implications for MHC evolution. *Molecular Evolution of the Major Histocompatibility Complex*, eds Klein J, Klein D (Springer, Heidelberg), pp 403–420.
4. Keller S, *et al.* (2012) High-precision isothermal titration calorimetry with automated peak-shape analysis. *Analyt Chem* 84:5066–5073.
5. Houtman JC, *et al.* (2007) Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Prot Sci* 16:30–42.
6. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in enzymology* 276:307–326.
7. Adams PD, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
8. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501.
9. Chen VB, *et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.
10. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66:125–132.
11. Winn MD, *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* 67:235–242.
12. Vagin AA, *et al.* (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr* 60:2184–2195.
13. Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* 78:1606–1619.
14. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* 98:10037–10041.
15. Nielsen H (2017) Predicting secretory proteins with SignalP. *Methods Mol Biol* 1611:59–73.
16. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410.
17. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874.
18. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
19. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425.
20. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791.
21. Zuckerkandl E, Pauling L (1965) Molecules as documents of evolutionary history. *J Theoret Biol* 8:357–366.